

2016 NATIONAL SURVEY ON DRUG USE AND HEALTH

METHODOLOGICAL RESOURCE BOOK SECTION 10: EDITING AND IMPUTATION REPORT

DISCLAIMER

SAMHSA provides links to other Internet sites as a service to its users and is not responsible for the availability or content of these external sites. SAMHSA, its employees, and contractors do not endorse, warrant, or guarantee the products, services, or information described or offered at these other Internet sites. Any reference to a commercial product, process, or service is not an endorsement or recommendation by SAMHSA, its employees, or contractors. For documents available from this server, the U.S. Government does not warrant or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed.

Substance Abuse and Mental Health Services Administration
Center for Behavioral Health Statistics and Quality
Rockville, Maryland

February 2018

This page intentionally left blank

2016 NATIONAL SURVEY ON DRUG USE AND HEALTH: EDITING AND IMPUTATION REPORT

Prepared for the 2016 Methodological Resource Book (Section 10)

Contract No. HHSS283201300001C
RTI Project No. 0213985.006.106.004.003
Deliverable No. 49

RTI Authors:

RTI Project Director:

Tricia Adams

Peter Frechtel

David Hunter

Kristen Gullledge Brown

Erica Hirsch

Lisa Carpenter

Larry Kroutil

Chuchun Chien

Jeff Laufenberg

SAMHSA Project Officer:

Susan Edwards

Peilan Chen Martin

Glynis Ewing

Andrew Moore

Peter Tice

Barbara Felts

Kayla Nowak

SAMHSA Author:

Rebecca Ahrnsbrak

For questions about this report, please e-mail Peter.Tice@samhsa.hhs.gov.

Prepared for Substance Abuse and Mental Health Services Administration,
Rockville, Maryland

Prepared by RTI International, Research Triangle Park, North Carolina

February 2018

Recommended Citation: Center for Behavioral Health Statistics and Quality.
(2018). *2016 National Survey on Drug Use and Health: Methodological
Resource Book (Section 10, Editing and Imputation Report)*. Substance
Abuse and Mental Health Services Administration, Rockville, MD.

Acknowledgments

This report was prepared for the Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, by RTI International (a registered trademark and a trade name of Research Triangle Institute). Special thanks are due to several current and former RTI staff members. Ralph Folsom, Eric Grau (formerly with RTI), and Avinash Singh (formerly with RTI) co-developed the predictive mean neighborhood (PMN) methodology, and Ralph Folsom and Phillip Kott co-developed the modified predictive mean neighborhood (modPMN) methodology. Claudia Clark edited the report, Debbie Bond formatted the report, Andrew Jessup and Valerie Garner assisted with graphics, and Amanda Lewis-Evans assisted with the flowcharts that appear in Chapters 3 and 7 ([Figures 3.1](#) through [3.4](#), [3.6](#), [3.7](#), and [7.1](#)).

Table of Contents

Chapter		Page
1.	Introduction.....	1
1.1	Organization of the NSDUH Questionnaire and Overview of Content.....	1
1.2	Overall Data Quality Issues with CAI	7
1.3	Organization of this Report.....	8
1.4	Changes from the 2015 Survey to the 2016 Survey	9
1.4.1	Changes to Accommodate Questionnaire Changes	9
1.4.2	Changes to Streamline Editing and Imputation Procedures	15
1.4.3	Changes Involving Minor Corrections and Enhancements	15
2.	Procedures and General Principles for Editing the Computer-Assisted Data.....	19
2.1	Introduction.....	19
2.2	Initial Processing of Transmitted NSDUH Interview Data	19
2.2.1	Assignment of Standard NSDUH Missing Data Codes.....	20
2.2.2	Remapping of Responses to "Enter All That Apply" Questions	21
2.2.3	Identification of Usable Cases	22
2.2.4	Creation of Cumulative Quarterly Unedited Data Files	25
2.3	Preliminary Editing and Coding of Processed Interview Data	26
2.3.1	Coding of "OTHER, Specify" Data	26
2.3.2	Investigation of Response Patterns in Usable Records.....	33
2.3.3	Edits to "Date-Dependent" Variables	37
2.4	General Issues and Principles for Editing NSDUH Data.....	39
2.4.1	Editing across Sections	39
2.4.2	Assignment of Relevant "Not Applicable" Codes	40
2.4.3	Additional Assignment of NSDUH Missing Data Codes as Part of Editing within a Section.....	42
2.4.4	Assignment of Codes to Edited "Enter All That Apply" Variables.....	43
2.4.5	"Back-Editing" Based on "OTHER, Specify" Data.....	45
2.4.6	"Lockout" Following Completion of Self-Administered Sections	46
2.4.7	"Flag and Impute" Principles for Substance Use Data Other than Prescription Drugs	47
3.	Imputation Procedures Used in the 2016 NSDUH	49
3.1	Introduction.....	49
3.1.1	Previously Used Hot-Deck Imputation Methods.....	53
3.2	Development of the Predictive Mean Neighborhood Methodology.....	55
3.3	Implementation for the Predictive Mean Neighborhood Methodology.....	57
3.3.1	Step 1: Response Propensity Adjustment	57
3.3.2	Step 2: Prediction Modeling	58
3.3.3	Step 3: Hot-Deck Imputation	61
3.4	Predictive Mean Neighborhood "Types"	66
3.4.1	Type 1: Single Response Propensity/Single Prediction.....	67
3.4.2	Type 2: Multiple Response Propensity/Multiple Prediction.....	68
3.4.3	Type 3: Single Response Propensity/Multiple Prediction	70

Table of Contents (continued)

Chapter	Page
3.5	Development of the Modified Predictive Mean Neighborhood Methodology 71
3.6	Procedural Steps for the Modified Predictive Mean Neighborhood Methodology 72
3.6.1	Step 1: Response Propensity Adjustment 72
3.6.2	Step 2: Prediction Modeling 73
3.6.3	Step 3: Imputation 74
3.7	Implementation of the Modified Predictive Mean Neighborhood Methodology 76
3.7.1	Sequencing of Variables for modPMN 77
3.7.2	Application of modPMN for a Single Variable 78
3.7.3	Cycling in modPMN 79
3.8	Special Auxiliary Variables: Age Group and State Rank 81
3.8.1	Age Groups 81
3.8.2	State Rank 81
4.	Editing and Imputation for the NSDUH Interviewer-Administered Demographics Variables 83
4.1	Introduction 83
4.2	Editing Selected Interviewer-Administered Demographics Variables 84
4.2.1	Creating the Edited Interview Date Variable (INTDATE) 84
4.2.2	Creating the Edited Age Variable (AGE) 85
4.2.3	Creating the Edited Current State Residency Variable (STATELOC) 88
4.2.4	Creating the Edited Gender Variable (IRSEX) 89
4.2.5	Creating the Edited Birth Date Variable (BRTHDATE) 89
4.2.6	Creating the Edited Marital Status Variables (MARITAL, EDMARIT, and NOMARR) 89
4.2.7	Creating the Edited Race and Hispanic/Latino (Origin and Group) Variables 90
4.2.8	Creating the Edited Military Service History and Military Veteran Status Variables 101
4.2.9	Creating the Edited Highest Grade Completed Variables (EDUHIGHST and EDEDUC) 105
4.2.10	Creating the Perceived Health Status Variable (HEALTH) 109
4.2.11	Creating the FI Debriefing Variables 109
4.3	Imputation for Selected Interviewer-Administered Demographics Variables 110
4.3.1	Marital Status (Imputation Set 1) 110
4.3.2	Race Variables (Imputation Set 2) 111
4.3.3	Hispanic/Latino Origin Variable (Imputation Set 3) 114
4.3.4	Hispanic/Latino Group Variable (Imputation Set 4) 115
4.3.5	Education Level Variable (Imputation Set 5) 116

Table of Contents (continued)

Chapter	Page
5. Editing and Imputation for the NSDUH Self-Administered Demographics Variables	119
5.1 Introduction.....	119
5.2 Editing the Self-Administered Demographics Variables.....	119
5.2.1 Editing for Patterned Responses and Questionable Interview Dates.....	119
5.2.2 Moves in the Past 12 Months.....	120
5.2.3 Immigrant Status.....	122
5.2.4 Adult Sexual Orientation	124
5.2.5 English Language Proficiency and Disability.....	125
5.2.6 Current Education	125
5.2.7 Military Families.....	133
5.2.8 Employment and Workplace.....	135
5.3 Imputation for Self-Administered Demographics Variables	139
5.3.1 Employment Status (Imputation Set 1).....	139
5.3.2 Immigrant Status, Born-in-U.S. Indicator (Imputation Set 2)	140
5.3.3 Immigrant Status, Age of Entry (Imputation Set 3).....	141
6. Editing and Imputation for the NSDUH Substance Use Variables	143
6.1 Introduction.....	143
6.1.1 Edited (but Not Imputed) Substance Use Variables	143
6.1.2 "Parent" and "Child" Drug Use Categories.....	144
6.2 Editing Procedures for the Substance Use Variables Other than Prescription Drugs.....	145
6.2.1 Edits of Lifetime Use Variables.....	147
6.2.2 Edits of Recency-of-Use Variables for Tobacco through Methamphetamine.....	152
6.2.3 Edits Involving Users of Only Child Drugs for Hallucinogens	160
6.2.4 Edits of Frequency-of-Use Variables.....	161
6.2.5 Initiation of Use (Age at First Use, Month of First Use, and Year of First Use)	167
6.2.6 Editing of Lifetime Daily Cigarette Use Status	168
6.2.7 Edits for Tobacco Variables that Do Not Undergo Imputation	168
6.3 Editing Procedures for the Prescription Drug Variables.....	174
6.3.1 Mentions of Over-the-Counter Drugs.....	176
6.3.2 Any Past Year or Lifetime Use of Specific Prescription Drugs	177
6.3.3 Most Recent Use of Prescription Drugs for Any Reason	180
6.3.4 Misuse of Specific Prescription Drugs in the Past 12 Months	181
6.3.5 Most Recent Misuse of Prescription Drugs	185
6.3.6 Initiation of Misuse of Prescription Drugs.....	186
6.3.7 Misuse of Prescription Drugs in the Past 30 Days.....	191
6.3.8 Ways of Misuse in the Past 12 Months.....	195
6.3.9 Injection of Prescription Stimulants in the Past 12 Months or Past 30 Days	196

Table of Contents (continued)

Chapter	Page
6.3.10	Identification of the Last Prescription Drug Misused in the Past 12 Months 198
6.3.11	Reasons for the Last Misuse of Prescription Drugs in the Past 12 Months 199
6.3.12	Sources of Prescription Drugs for the Last Misuse in the Past 12 Months 203
6.3.13	Prescription Drug Edit Rules 205
6.4	Additional Edits Applied During Imputation Processing 207
6.4.1	"Other" Hallucinogens 208
6.4.2	Respondents Imputed to Lifetime Use for Child Drug Variables..... 208
6.4.3	Age-at-First-Use Variables 209
6.5	Imputation of Substance Use Variables 210
6.5.1	Variables Imputed Using PMN Methodology 210
6.5.2	Variables Imputed Using the modPMN Methodology 236
6.5.3	Recodes for Additional Analyses..... 251
7.	Editing and Imputation for Variables from the Special Drugs and Substance Dependence and Abuse Sections 253
7.1	Introduction..... 253
7.2	Special Drugs 254
7.2.1	GHB and Over-the-Counter Cold Medicines 254
7.2.2	General Edits of Cocaine, Heroin, and Methamphetamine Variables for Nonusers 256
7.2.3	Edits for Heroin Users Who Did Not Report Smoking, Sniffing, or Injection of Heroin..... 256
7.2.4	Additional Edits for Smoking or Sniffing of Heroin 257
7.2.5	Injection Drug Use..... 258
7.3	Substance Dependence and Abuse..... 262
7.3.1	Editing of Substance Dependence and Abuse Variables 263
7.3.2	Imputation of Cigarette Dependence and Substance Dependence and Abuse Variables 272
8.	Editing of Additional Self-Administered Sections for Substance Use 281
8.1	Introduction..... 281
8.2	Blunts 281
8.2.1	Use of Blunts..... 281
8.2.2	Medical Marijuana Use..... 284
8.3	Prior Substance Use 287
8.3.1	Retrospective Use 288
8.3.2	Last Use of Drugs 288
8.4	Substance Use Treatment..... 291
8.4.1	General Skip Logic Edits 291
8.4.2	Lifetime, Past Year, or Current Treatment 295

Table of Contents (continued)

Chapter	Page
8.4.3	Last or Current Treatment Episode..... 298
8.4.4	Editing of Miscellaneous Treatment Variables that Existed Prior to 2004..... 300
8.4.5	Lifetime History of Treatment and Additional Self-Help Questions..... 301
8.4.6	Perceived Need for Substance Treatment..... 303
8.4.7	Reasons for Not Receiving Treatment..... 304
8.5	Consumption of Alcohol..... 306
8.5.1	General Skip Logic Edits..... 306
8.5.2	Number of Drinks for the Last Alcohol Use in the Past 30 Days..... 307
8.5.3	Source and Location of Last Underage Alcohol Use in the Past 30 Days..... 307
8.5.4	Use of Illicit Drugs with Alcohol in the Past 30 Days..... 308
8.5.5	Binge Alcohol History..... 309
9.	Editing of Self-Administered Sections for Risk and Protective Factors for Substance Use..... 311
9.1	Introduction..... 311
9.2	Risk/Availability..... 311
9.3	Special Topics..... 312
9.3.1	Editing of Variables for Criminal Justice Involvement..... 312
9.3.2	Editing of Variables for Driving under the Influence..... 314
9.3.3	Editing of Variables for State Marijuana Laws..... 317
9.4	Social Environment..... 318
9.5	Youth Experiences..... 318
9.6	Parenting Experiences..... 319
9.6.1	Editing of the Field Interviewer Checkpoint Variables..... 320
9.6.2	Editing of the Variables in the Parenting Experiences Section..... 323
10.	Editing of Self-Administered Sections for Physical and Mental Health..... 327
10.1	Introduction..... 327
10.2	Health..... 327
10.2.1	Pregnancy Status..... 327
10.2.2	Height and Weight..... 328
10.2.3	Utilization of Medical Services..... 329
10.2.4	Screening for Substance Use or Talks about Substance Use with a Health Care Professional..... 330
10.2.5	Specific Health Conditions..... 335
10.3	Adult Mental Health Service Utilization..... 337
10.4	Adult Mental Health..... 342
10.5	Adult and Adolescent Depression..... 343
10.6	Youth Mental Health Service Utilization..... 348

Table of Contents (continued)

Chapter	Page
11. Editing and Imputation for the NSDUH Roster Variables	353
11.1 Introduction.....	353
11.2 Editing the Respondent-Level Detailed Roster Variables	356
11.2.1 Roster Consistency Checks.....	357
11.2.2 Roster-Level Dataset.....	359
11.2.3 Roster Edits Involving the Self.....	359
11.2.4 Roster Edits for Other Household Members.....	360
11.2.5 Final Edited Respondent-Level Detailed Roster Variables	367
11.3 Editing and Imputation Procedures for Roster-Derived Household Composition Variables.....	368
11.3.1 Creation of Edited Roster-Derived Household Composition Variables	368
11.3.2 Imputation for Roster-Derived Household Composition Variables	369
11.4 Editing the Proxy Variables.....	371
11.4.1 Edited Indicator of Potential Proxies in Household (EDFAM18)	374
11.4.2 Editing the Proxy Variables when EDFAM18 = 1	374
11.4.3 Editing the Proxy Variables when EDFAM18 = 0	375
11.4.4 Additional Editing for PRXRELAT	375
11.4.5 Additional Editing for QP03A.....	375
12. Editing and Imputation for the NSDUH Income Variables.....	377
12.1 Introduction.....	377
12.2 Editing for Binary and Finer Category Income Variables	378
12.2.1 Binary Income Variables	378
12.2.2 Finer Category Income Variables	379
12.2.3 Additional Edits to Income Variables for Inconsistencies or Logical Inferences.....	380
12.3 Editing for Cell Phone Variables	381
12.4 Imputation for the Income Variables	381
12.4.1 Imputation for Binary Income Variables (Imputation Set 1).....	382
12.4.2 Imputation for Finer Category Income Variables (Imputation Set 2).....	387
13. Editing and Imputation for the NSDUH Health Insurance Variables.....	391
13.1 Introduction.....	391
13.2 Editing the Health Insurance Variables	392
13.2.1 Editing Because of Incorrect State Information.....	394
13.2.2 Determination of Any Health Insurance Coverage.....	394
13.2.3 Editing of Health Insurance Follow-Up Variables Based on Current Coverage	396
13.2.4 Miscellaneous Issues for Editing the Health Insurance Variables.....	396
13.2.5 Editing Types of Health Insurance and Any Other Health Insurance Variables, Imputation Sets 1 and 2.....	398

Table of Contents (continued)

Chapter	Page
13.3 Imputation for the Health Insurance Variables	400
13.3.1 Types of Health Insurance Variables, Imputation Set 1	400
13.3.2 Any Other Health Insurance Variable, Imputation Set 2	402
13.3.3 Health Insurance Recodes for Additional Analyses	403
14. Editing and Imputation for the NSDUH Pair Variables	405
14.1 Introduction.....	405
14.1.1 Pair Relationship Variables.....	406
14.1.2 Multiplicity Count Variables	406
14.1.3 Household-Level Person Count Variables.....	408
14.2 Stage One Editing: Pair Relationships	408
14.2.1 Editing the Household Roster of Each Pair Member.....	408
14.2.2 Creating the Pair Relationship Variable (PAIRREL)	409
14.3 Stage Two Editing: Multiplicity Counts	414
14.3.1 Determining the Multiplicity Count for Each Pair Member	416
14.3.2 Determining the Final Multiplicity Count	416
14.4 Stage Three Editing: Household-Level Person Counts	417
14.4.1 Determining the Household-Level Person Count for Each Respondent.....	419
14.4.2 Determining the Final Household-Level Person Count.....	423
14.5 Imputation for the Pair Variables.....	424
14.5.1 Creation of Covariates for Imputing Pair-Level Variables.....	424
14.5.2 Stage One Imputation: Pair Relationships	427
14.5.3 Stage Two Imputation: Multiplicity Counts	430
14.5.4 Stage Three Imputation: Household-Level Person Counts.....	432
References.....	437

Table of Contents (continued)

Appendix	Page
A Percentage of Cases Imputed or Logically Assigned, 2016 Survey	A-1
B Detailed Documentation of Edits for Variables from the Self-Administered Demographics, Substance Use, and Additional Self-Administered Sections in the 2016 NSDUH.....	B-1
C Race, Hispanic/Latino Group, and Drug Codes.....	C-1
D Model Summaries	D-1
E Hot-Deck/Collocated Stochastic Imputation Procedure Summaries	E-1
F Quality Control Measures Used in the Imputation Procedures	F-1
G Interviewer Explanations for Overrides to Consistency Checks in Household Roster	G-1
H Rules for Determining Pair Relationships	H-1
I Conditions Used to Reconcile Differing Multiplicity Counts between Pair Members	I-1
J Conditions Used to Reconcile Differing Household-Level Person Counts between Pair Members	J-1
K Priority Conditions Used to Create Household-Consistent Covariates	K-1
L Creation of Household-Level and Person-Level Files.....	L-1

List of Tables

Table	Page
1.1 Content of Self-Administered Substance Use Sections in the First Part of the 2016 NSDUH Interview	2
1.2 Content of Additional Self-Administered Sections in the 2016 NSDUH Interview	5
2.1 Effects on the Potential Usable Status of a Case Based on Responses to Follow-Up Questions for Refusals to Gate Questions in the Initial Substance Use Sections	24
3.1 Summary of Item Imputation Procedure Used, by Variable and NSDUH Year	51
3.2 Regression Models Used for Each Variable Imputed with Predictive Mean Neighborhood	59
3.3 Values of Delta for Various Predicted Probabilities	63
3.4 Examples of Variables Imputed Using Each of the Four Combinations of Univariate/Multivariate Matching and Assignment	65
3.5 PMN Types Applied to Each Variable Group and Imputation Set	67
3.6 Regression Models Used for Each Variable Imputed Using modPMN	74
3.7 Imputation Sets Processed Using modPMN	77
3.8 Example of Sequencing for Imputation Sets 4 and 8	78
3.9 Independent Variables Used in Response Propensity and Prediction Steps of modPMN	80
4.1 Hispanic/Latino Status of Item Nonrespondents for Race	113
4.2 Edited Race Variables and Their Imputation-Revised Counterparts	114
4.3 Mapping of EDHOGRP Levels to EDHOGRP2 Levels	116
4.4 Mapping of EDEDUC Levels to RESPEDUC Levels	117
5.1 Names in 2016 and 2015 for Current Education Variables	127
5.2 Mapping of Current Age with Expected Grades	130
5.3 Names in 2016 and 2015 for Military Family Variables	134
5.4 Categories of WRKSTATWK	137
6.1 Substances in the 2016 NSDUH with a Parent/Child Relationship	145
6.2 How the Flag and Impute Edit Procedures Handled Usual Inconsistencies Involving the CAI Recency Variables	155
6.3 How the Flag and Impute Edit Procedures Handled Inconsistencies between Parent and Child Recency Variables	158

List of Tables (continued)

Table	Page
6.4 "OTHER, Specify" Codes for Reasons for Misuse that Corresponded to Available Response Options.....	200
6.5 Inconsistent Patterns in the Prescription Drug Modules and their Resolution	206
6.6 Drugs and Drug Use Measures, PMN Imputation Sets	211
6.7 General Incomplete Recency Categories for Cigarettes	216
6.8 Elements of Full Predictive Mean Vector: Cigarettes	218
6.9 Elements of Full Predictive Mean Vector.....	232
6.10 Drugs and Drug Use Measures, modPMN Imputation Sets	237
6.11 Sequencing of Drugs within Imputation Sets, 2016	237
6.12 Summary of Response Propensity Models for Categorical modPMN Variables	240
6.13 Outcomes of Prediction Modeling.....	241
6.14 Final Imputation-Revised Categorical modPMN Variables	242
6.15 Summary of Response Propensity Models for Continuous modPMN Variables	245
6.16 Final Imputation-Revised Continuous modPMN Variables	247
6.17 Prevalence Recodes	251
6.18 Prevalence Recodes Incorporating More than One Recency Variable.....	252
7.1 Mapping of Cigarette Dependence Questions to Edited Variables	265
7.2 Quick Reference Guide to Creation of Edited Dependence Indicators	267
7.3 Value of Abuse as Derived from Dependence and the Abuse Criteria, Zero-Fill Drugs	272
7.4 Value of Edited Abuse as Derived from Edited Dependence and the Edited Abuse Criteria, modPMN Drugs.....	272
7.5 Summary of Response Patterns for 2016 NDSS Variables	274
7.6 Segregation of Dependence and Abuse Indicators for modPMN Drugs into Imputation Sets	275
7.7 Sequencing of Dependence and Abuse Indicators within Imputation Sets, 2016	275
7.8 Tables Reporting Covariate Lists for Response Propensity Models, Dependence and Abuse Indicators.....	276
7.9 Tables Reporting Covariate Lists for Prediction Models, Dependence and Abuse Indicators.....	277
8.1 States Where Marijuana Was Approved for Medical Use as of December 31, 2016, and Dates When the Laws Took Effect	286

List of Tables (continued)

Table	Page
10.1 Mapping of Edited Variables for Health Care Screening or Talks about Substance Use with a Health Care Professional	331
10.2 Depression Symptom and Score Variables	346
11.1 Roster Grid Example Where Number of Individuals in Household (QD54) Equals 4.....	354
11.2 Roster Relationship Codes Prior to 2016.....	354
11.3 Detailed Roster Relationship Codes Beginning in 2016.....	354
11.4 Roster-Derived Household Composition Variables.....	355
11.5 Proxy Variables.....	356
11.6 Levels of PRXRELAT	372
11.7 Levels of PRXRELATD	373
12.1 Comparison between 1999-2007, 2008-2014, and 2015-2016 Sets of Income Questions.....	378
12.2 Imputation Order for Binary Income Variables (Imputation Set 1)	382
12.3 Cross-Classification of FAMPMT and FAMSVC; Relevant Combinations for Binary Income Predictive Mean Vector when WELMOS Is Missing.....	386
13.1 Edits Pertaining to the Health Insurance Section.....	397
13.2 Mapping of Raw Health Insurance Variables to Edited Variables	398
14.1 Edited and Imputation-Revised Multiplicity Count Variables	407
14.2 Edited and Imputation-Revised Household-Level Person Count Variables.....	408
14.3 Levels of the Variable PAIRREL	410
14.4 Values of PAIRREL That Correspond to the Levels of the Variable RELMATCH.....	413
14.5 Frequencies of the Levels of the Variable RELMATCH: 2016	414
14.6 Multiplicity Counts for Each Pair Member	416
14.7 Amount of Imputation Required for Multiplicities in Various Pair Domains: 2016.....	417
14.8 Age Group Pairs with Associated Possible Pair Relationships	427
14.9 Modeled Pair Relationships within Age Group Pairs	430
14.10 Modeled Multiplicities within Domains	431

List of Figures

Figure	Page
3.1 Donor Selection Algorithm.....	62
3.2 PMN Type 1: Single Response Propensity/Single Prediction	67
3.3 PMN Type 2: Multiple Response Propensity/Multiple Prediction	69
3.4 PMN Type 3: Single Response Propensity/Multiple Prediction.....	70
3.5 Stochastic Imputation for a Three-Level Categorical Variable	75
3.6 Twelve-Step Block of modPMN Programs	79
3.7 Arranging the 12-Step Blocks to Complete Imputation for an Imputation Set	80
5.1 Follow-Up Question Logic for Current Education	126
5.2 Routing Logic for Respondents in the Employment Status Section.....	136
6.1 Routing Logic for Respondents in the Substance Use Sections Other than for Prescription Drugs	147
6.2 Routing Logic for Respondents in the Prescription Psychotherapeutic Drug Sections	175
7.1 Sequencing of Steps for Hallucinogens Dependence and Abuse	278
13.1 Routing Logic for Respondents in the Health Insurance Section	393

1. Introduction

Conducted annually, the National Survey on Drug Use and Health (NSDUH) is the primary source of information on alcohol use, tobacco use, illicit drug use (including misuse of prescription drugs), substance use disorders, and mental health issues for the U.S. civilian, noninstitutionalized population aged 12 years or older. In the 2016 NSDUH, this population included residents of noninstitutional group quarters (e.g., shelters, rooming houses, dormitories, and group homes) and civilians residing on military bases. The target population excluded people with no fixed household address (e.g., homeless or transient people not in shelters), residents of institutional group quarters (e.g., jails and nursing homes), children younger than 12 years of age, and active military personnel. As it has since 1999, the 2016 NSDUH utilized a 50-state, multistage cluster design that enables the Center for Behavioral Health Statistics and Quality within the Substance Abuse and Mental Health Services Administration (SAMHSA) to provide representative estimates for each state and the District of Columbia. Both direct and model-based state and substate estimates are produced on a variety of measures based on a combination of multiple years of data.

This report focuses on the editing and statistical imputation procedures that were applied to respondent data for the 2016 NSDUH. Logical editing uses data from elsewhere within the same respondent's record to reduce the occurrence of missing or ambiguous data or to resolve inconsistencies between related variables. Imputation is defined as the replacement of missing values with valid, nonmissing values. Statistical imputation usually involves some randomness to preserve the natural variability in the data.

1.1 Organization of the NSDUH Questionnaire and Overview of Content

The 1999 survey marked the transition from data collection based on paper-and-pencil interviewing (PAPI) to computer-assisted interviewing (CAI).¹ Since then, the NSDUH data have been collected using CAI, which allows a private mode of data collection for respondents to answer questions pertaining to drug use and other sensitive topics.

The CAI interview is organized by sections based on the mode of administration, content, and routing logic. For example, the alcohol section includes an initial question that asks whether respondents ever had a drink of an alcoholic beverage. Respondents who report that they have ever used alcohol are asked additional questions about their age when they first used alcohol, when they last used alcohol, and (if relevant) their use of alcohol in the 12 months or 30 days prior to the interview. Respondents who do not report lifetime alcohol use are routed to the next major section, which pertains to marijuana use.

The first section of the interview consists of questions about certain demographic characteristics, including birth date (which is used to determine a respondent's age), gender, Hispanic/Latino origin, racial group, marital status, and education level (highest grade completed). Computer-assisted personal interviewing (CAPI) is used for these questions, a process in which interviewers read the questions and respondents give their answers aloud to the

¹ CAI specifications for the 2016 NSDUH are available at <https://www.samhsa.gov/data/>.

interviewers, who then enter the responses into the computer. The logic for determining which questions the interviewers should ask is controlled by the computer program based on the responses to previous questions that interviewers enter into the computer. Consequently, interviewers can concentrate on asking questions and recording respondent answers, without having to concern themselves with following skip pattern instructions.

Following completion of these demographic questions, interviewers orient respondents about the use of specific features of the laptop computer and use of headphones for listening to questions. Respondents then complete a brief tutorial on answering self-administered questions, which are administered through use of audio computer-assisted self-interviewing (ACASI). ACASI is used for the majority of questions in the interview and is particularly useful for collecting data about sensitive topics such as (but not limited to) substance use, problems associated with substance use, risk and protective factors related to substance use, and mental health issues.

In the ACASI portion of the interview, respondents can read the questions on a computer screen and also are encouraged to listen to an audio recording of the questions on headphones. Respondents then enter their answers directly into a laptop computer. These features of ACASI prevent interviewers or others in the household from knowing what questions the respondents are being asked and how they are answering. The availability of audio recordings of the questions is especially useful for respondents with limited reading ability because they can listen to the questions instead of having to read them.

The first set of self-administered questions ([Table 1.1](#)) pertains to the use of tobacco, alcohol, marijuana, cocaine, crack cocaine, heroin, hallucinogens, inhalants, methamphetamine, and prescription psychotherapeutic drugs (i.e., pain relievers, tranquilizers, stimulants, and sedatives). Questions underwent a partial redesign in 2015 for smokeless tobacco, binge alcohol use (for females), hallucinogens, inhalants, methamphetamine, and prescription psychotherapeutic drugs. Except for prescription psychotherapeutic drugs (see Section 1.4.1), these questions did not undergo further changes for 2016. Questions did not change as part of the partial redesign in 2015 for cigarettes, cigars, pipe tobacco, alcohol (any use), binge alcohol use (for males), marijuana, cocaine, crack cocaine, and heroin to permit measurement of trends in use for 2002 to 2015. Chapter 6 presents documentation of the editing and imputation procedures for these questions.

Table 1.1 Content of Self-Administered Substance Use Sections in the First Part of the 2016 NSDUH Interview

Section	Content
Tobacco	• Lifetime use or nonuse of the following: – Cigarettes – Smokeless tobacco – Cigars – Pipe tobacco • Additional questions for lifetime users
Alcohol	• Lifetime use or nonuse • Additional questions for lifetime users

Table 1.1 Content of Self-Administered Substance Use Sections in the First Part of the 2016 NSDUH Interview (continued)

Section	Content
Marijuana	• Lifetime use or nonuse • Additional questions for lifetime users
Cocaine and Crack	• Lifetime use or nonuse of the following: – Any cocaine – Crack cocaine (if lifetime cocaine user) • Additional questions for lifetime users
Heroin	• Lifetime use or nonuse • Additional questions for lifetime users
Hallucinogens	• Lifetime use or nonuse of 10 types of hallucinogens, including any other hallucinogen besides the ones that had been listed • Additional questions if lifetime use was reported for any of the 10 types of hallucinogens
Inhalants	• Lifetime use or nonuse of 13 specific types of inhalants for kicks or to get high, including any other inhalant besides the ones that had been listed • Additional questions if lifetime use was reported for any of the 13 types of inhalants
Methamphetamine	• Lifetime use or nonuse • Additional questions for lifetime users
Pain Relievers Screener	• Any use in the past 12 months of 36 specific pain relievers and any other pain reliever • Any lifetime use of prescription pain relievers, if past year use was not reported
Tranquilizers Screener	• Any use in the past 12 months of 12 specific tranquilizers and any other tranquilizer • Any lifetime use of prescription tranquilizers, if past year use was not reported
Stimulants Screener	• Any use in the past 12 months of 26 specific stimulants and any other stimulant • Any lifetime use of prescription stimulants, if past year use was not reported
Sedatives Screener	• Any use in the past 12 months of 14 specific sedatives and any other sedative • Any lifetime use of prescription sedatives, if past year use was not reported
Pain Relievers Main Section	• Misuse in the past 12 months of any of the specific pain relievers that were used in the past 12 months, where misuse was defined as use "in any way not directed by a doctor" (see Section 6.3.4) • First misuse of specific pain relievers that were misused in the past 12 months • Lifetime misuse of pain relievers, if no past year misuse was reported • Initiation of misuse of all pain relievers in the past 12 months • Misuse of any pain reliever in the past 30 days, if misuse of specific pain relievers in the past 12 months was reported • Additional questions for misusers in the past 30 days • Additional questions for misusers in the past 12 months

Table 1.1 Content of Self-Administered Substance Use Sections in the First Part of the 2016 NSDUH Interview (continued)

Section	Content
Tranquilizers Main Section	• Misuse in the past 12 months of any of the specific tranquilizers that were used in the past 12 months, where misuse was defined as use "in any way not directed by a doctor" (see Section 6.3.4) • First misuse of specific tranquilizers that were misused in the past 12 months • Lifetime misuse of tranquilizers, if no past year misuse was reported • Initiation of misuse of all tranquilizers in the past 12 months • Misuse of any tranquilizer in the past 30 days, if misuse of specific tranquilizers in the past 12 months was reported • Additional questions for misusers in the past 30 days • Additional questions for misusers in the past 12 months
Stimulants Main Section	• Misuse in the past 12 months of any of the specific stimulants that were used in the past 12 months, where misuse was defined as use "in any way not directed by a doctor" (see Section 6.3.4) • First misuse of specific stimulants that were misused in the past 12 months • Lifetime misuse of stimulants, if no past year misuse was reported • Initiation of misuse of all stimulants in the past 12 months • Misuse of any stimulant in the past 30 days, if misuse of specific stimulants in the past 12 months was reported • Additional questions for misusers in the past 30 days • Additional questions for misusers in the past 12 months, including injection of stimulants
Sedatives Main Section	• Misuse in the past 12 months of any of the specific sedatives that were used in the past 12 months, where misuse was defined as use "in any way not directed by a doctor" (see Section 6.3.4) • First misuse of specific sedatives that were misused in the past 12 months • Lifetime misuse of sedatives, if no past year misuse was reported • Initiation of misuse of all sedatives in the past 12 months • Misuse of any sedative in the past 30 days, if misuse of specific sedatives in the past 12 months was reported • Additional questions for misusers in the past 30 days • Additional questions for misusers in the past 12 months

The second section of self-administered questions follows the substance use questions and asks about a variety of sensitive topics related to substance use and mental health issues (Table 1.2). These topics include (but are not limited to) injection drug use, perceived risks of substance use, substance use disorders (SUDs), arrests, treatment for substance use problems, pregnancy, mental illness, the utilization of mental health services, and further details for the consumption of alcohol. Documentation of the editing procedures for variables in these sections of the interview is included in Chapters 7 through 10. Documentation of the imputation procedures for the SUD variables also is discussed in Chapter 7.

Additional demographic questions were self-administered in 2016 (Table 1.2). These questions address topics such as immigration, current school enrollment, and employment and workplace issues. Questions also were added to this section of the interview beginning in 2015 for sexual attraction and sexual identity (administered only to adults), English language

proficiency, disabilities, and military families. Documentation of editing and imputation procedures for these self-administered demographic variables is described in Chapter 5.²

Once respondents complete the ACASI portion of the interview, they return the laptop computer to the interviewer. Remaining interview questions are administered through CAPI and address household characteristics, health insurance coverage, and income.

Table 1.2 Content of Additional Self-Administered Sections in the 2016 NSDUH Interview

Section	Content
Special Drugs	• Use of gamma hydroxybutyrate (GHB) or the use of cough and cold medications to get high • Routes of heroin administration (i.e., smoking, sniffing, injection) • Injection of cocaine, methamphetamine, or other drugs • General injection use behaviors
Risk/Availability	• Perceived risk of harm associated with use of cigarettes, alcohol, or specific illicit drugs • Perceived availability of illicit drugs • Miscellaneous risk behaviors (e.g., seatbelt use)
Blunts	• Use of cigars with marijuana in them (i.e., "blunts") • Use of marijuana that had been recommended by a doctor or other health care professional
Substance Dependence and Abuse	• Nicotine (i.e., cigarette) dependence • Symptoms of dependence and abuse for alcohol and illicit drugs, including methamphetamine and the misuse of prescription drugs
Special Topics	• Arrests in the lifetime and past 12 months • Probation and parole status, past 12 months • Driving under the influence of alcohol, marijuana, cocaine, heroin, hallucinogens, inhalants, or methamphetamine • Knowledge of state laws for marijuana possession
Prior Substance Use	• Use of marijuana, cigarettes, alcohol, or cocaine more than 12 months ago but within the past 24 months • Last use of substances included in the drug use section (see Table 1.1) if not used in the past 30 days
Substance Use Treatment	• Treatment for the use of alcohol or illicit drugs (i.e., not counting tobacco) in the lifetime and past 12 months • Perceived need for treatment for use of alcohol or illicit drugs • Barriers to receiving treatment for use of alcohol or illicit drugs
Health	• Pregnancy status of females aged 12 to 44 • Height and weight • Use of health services in the past 12 months, including emergency room visits, hospitalizations, and outpatient medical visits • Conversations about tobacco, alcohol, or illicit drug use with health care professionals in the past 12 months • Sexually transmitted diseases in the past 12 months • Lifetime history of specific medical conditions, age at first diagnosis, and miscellaneous issues

² For the 2016 NSDUH, marital status questions were moved back to the CAPI section at the beginning of the interview.

Table 1.2 Content of Additional Self-Administered Sections in the 2016 NSDUH Interview (continued)

Section	Content
Adult Mental Health Service Utilization	• Administered to respondents aged 18 or older • Use of inpatient or outpatient mental health services in the past 12 months • Payment for inpatient or outpatient mental health services • Use of prescription medication to treat a mental health condition • Unmet demand for mental health services • Use of alternative sources for mental health treatment
Social Environment	• Administered to respondents aged 18 or older • Specific illegal behaviors regardless of arrest (e.g., selling illegal drugs) • Attitudes about marijuana use • Religious involvement and beliefs
Parenting Experiences	• Administered to parents when two people were selected for an interview and the second selected person was an adolescent • Beliefs about whether their child has used tobacco, alcohol, or drugs • Talks with their child about substance use • Attitudes about drug education
Youth Experiences	• Administered to respondents aged 12 to 17 • Risk and protective factors for substance use • Fighting and delinquent behaviors • Attitudes and perceptions of others' attitudes about substance use • Exposure to drug prevention messages • Religious involvement and beliefs
Mental Health	• Administered to respondents aged 18 or older • Psychological distress in the past 30 days or past 12 months • Impairment in carrying out activities because of psychological distress • Suicidal thoughts and behavior
Adult Depression	• Administered to respondents aged 18 or older • Symptoms of depression in the lifetime and past 12 months • Impairment in carrying out activities because of depression symptoms • Treatment for depression
Youth Mental Health Service Utilization	• Administered to respondents aged 12 to 17 • Use of mental health services in the past 12 months • Reasons for receiving mental health services from specific sources
Adolescent Depression	• Administered to respondents aged 12 to 17 • Age-appropriate questions for symptoms of depression in the lifetime and past 12 months • Age-appropriate questions for impairment in carrying out activities because of depression symptoms • Treatment for depression
Consumption of Alcohol	• Number of drinks on the last occasion of alcohol use in the past 30 days • Underage alcohol use • Alcohol use in combination with illicit drugs • Initiation of binge alcohol use
Self-Administered Demographics	• Moves in the past 12 months and state residency 1 year ago (if moved in the past 12 months) • Immigrant status and length of time living in the United States • Sexual attraction and sexual identity (administered only to adults) • English language proficiency and disability

**Table 1.2 Content of Additional Self-Administered Sections in the 2016 NSDUH Interview
(continued)**

Section	Content
Education	• Current school enrollment • Military families
Employment	• Current employment status • Employment in the past 12 months (if not currently employed) • Absences from work in the past 30 days • Substance use education and testing policies in the workplace

1.2 Overall Data Quality Issues with CAI

Conversion of the NSDUH interview from PAPI (paper format) to CAI (computer format) in 1999 greatly reduced or, in some cases, eliminated the following data quality problems:

- illegible responses, multiple marks, or out-of-range values;
- item nonresponse (i.e., missing data);
- incorrectly executed skip patterns; and
- inconsistencies among related variables.

For example, when a question's instructions ask respondents to choose only one answer, PAPI respondents nevertheless are physically able to mark multiple answers. This cannot occur in the CAI because the computer program will permit entry of only one response to the item. In addition, when CAI respondents are allowed to choose one or more responses from a list (e.g., prescription drugs that they used in the past 12 months) or to enter a response of "95" to mean that none of the items in the list apply (e.g., if they had not used any of the prescription drugs in a given list in the past 12 months), then the computer program will not allow respondents to choose items from the list *and* enter "95." Similarly, the CAI has been programmed not to allow out-of-range values for certain items, such as frequency-of-use items, thereby reducing the amount and types of out-of-range values that would otherwise need to be addressed through editing. Further, the skip patterns that are embedded in the CAI were designed to reduce the occurrence of inconsistent data by not giving respondents the *opportunity* to provide inconsistent answers. The occurrence of inconsistent data was further reduced through the use of consistency checks, or "hard errors," built into the CAI program that prompted respondents to resolve inconsistencies that occurred between related items.³

Despite the potential for improvements in data quality through a CAI instrument, it was recognized that conversion to CAI would not completely eliminate data problems. For example, missing data were not completely eliminated because CAI respondents still had the option of

³ When respondents are presented with a hard error message because of inconsistent answers, they cannot proceed further with the interview until they revise their answers to make them consistent. For consistency checks, respondents are given an opportunity to revise their answers to make them consistent but can proceed with the interview regardless of whether they have resolved the inconsistency.

entering a response of "don't know" or "refused" when answering a given item. Similarly, even though consistency checks were designed to reduce inconsistent reporting, the CAI program was not equipped to address every possible inconsistent report that a respondent could make. Finally, in some situations, conversion to CAI could introduce new data quality issues. As discussed in Section 2.3.2 in Chapter 2, for example, the lack of direct interviewer monitoring could allow some respondents to use the computer keyboard to enter nonsensical patterns of answers for reasons such as accidental errors or disinterest.

1.3 Organization of this Report

Preliminary coding and processing of unedited data after interviewers transmitted the data from the field as well as the general principles of logical editing are discussed in Chapter 2 of this report. The predictive mean neighborhood (PMN) and modified predictive mean neighborhood (modPMN) imputation methodologies, which are used to impute NSDUH data, are described in detail in Chapter 3. It is recommended that readers first review Chapters 2 and 3 because these two chapters provide a foundation for Chapters 4 through 14. The information in both chapters helps to set an appropriate context for readers as they review the documentation for the specific editing and imputation procedures that were employed on a particular variable set.

Following these initial chapters, Chapters 4 and 5 address the editing and imputation procedures that were applied to the interviewer-administered and self-administered demographic variables, respectively. Editing and imputation procedures for the substance use variables that form the basis for published estimates of substance use and misuse in NSDUH reports and tables are discussed in Chapter 6. Chapter 7 addresses the editing procedures for the special drugs and substance dependence and abuse sections of the interview and the imputation procedures for relevant substance dependence and abuse variables. Chapter 7 also discusses the imputation procedures for cigarette dependence,⁴ which differ from the procedures used for other drug variables. Chapters 8 through 10 discuss editing procedures for additional self-administered sections of the interview; the variables that apply to these chapters did not undergo imputation in 2016. Chapter 11 describes the edits that were applied to the household roster, the creation of imputation-revised versions of the roster-derived household composition variables, and the creation of respondent-level variables with individual roster information. Chapter 12 summarizes the editing and imputation procedures that were applied to the income variables. Procedures for logical editing and imputation of missing values in the health insurance variables are described in Chapter 13. The editing and imputation processing of the pair relationship variable and related household count variables are detailed in Chapter 14.

This report also contains 12 appendices. Appendix A contains a number of tables that quantify the amount of imputation and logical assignment (i.e., editing) that selected analytic variables underwent during imputation processing in 2016. Appendix B provides a summary of data issues in different sections of the interview and the specific edits that were applied in response to these issues. Appendix C provides details on the handling of "OTHER, Specify" responses to the race and Hispanic/Latino group questions so that the data could be summarized in a meaningful way, as well as changes to the coding of "OTHER, Specify" data for prescription drugs to reflect the changes to these sections of the questionnaire beginning in 2015. The

⁴ This term is referred to in imputation reports prior to 2012 as "nicotine dependence."

covariates in each of the imputation models are listed in Appendix D. The tables in Appendix D also include (1) the starting list of covariates for each model and (2) descriptions of each level and identification of the reference level for categorical covariates. Appendix E provides details on each final hot-deck step, or each collocated stochastic imputation step for categorical variables imputed using modPMN, in the imputation procedures. The quality control measures used in the imputation procedures are summarized in Appendix F. Reasons that interviewers gave for overriding consistency checks in the household roster are presented in Appendix G, along with evaluations of their legitimacy and the resulting actions in editing the roster. The rules for determining pair relationships are defined in Appendix H. The conditions used for reconciling differing multiplicity counts between pair members are described in Appendix I, and the conditions used for reconciling differing household-level person counts between pair members are described in Appendix J. Appendix K details the priority conditions for creating household-consistent covariates. Appendix L contains detailed information about household-level and person-level eligibility and the completeness criteria used to construct the household-level and person-level files.

1.4 Changes from the 2015 Survey to the 2016 Survey

This section summarizes the changes in editing and imputation procedures implemented on the NSDUH since the 2015 survey year. It is intended for analysts who have had previous experience with the data. Those with little or no prior experience with NSDUH data are encouraged to review Chapters 2 and 3 before reviewing this section. For sections of the interview that underwent little or no change relative to the previous year, a careful assessment of the impact of the change is done to determine what, if any, effect the changes to the editing and imputation procedures would have on the estimates; this assessment is done before making formal changes to the procedures. The impact assessment involves applying the new process to the previous year's data and comparing the results. For variables whose creation has a stochastic component, the new process is run at least three times to assess whether the results are directly attributable to the new process or whether they could be due to random variation.

The changes that were implemented to the editing and imputation procedures between the 2015 and 2016 surveys are described in three sections below. In particular, Section 1.4.1 describes changes to the 2016 questionnaire and includes, where relevant, a brief discussion of the effects on editing and imputation procedures for 2016; starting with Chapter 4, the effects of the questionnaire changes outlined in Section 1.4.1 on editing and imputation procedures are discussed in greater detail as they relate to the relevant sections of the interview. Section 1.4.2 describes changes to the editing and imputation procedures that reduced the time required to process the data, with little to no impact on the final imputation-revised variables. Finally, Section 1.4.3 lists a few minor corrections and improvements to the editing and imputation procedures that also had little impact on the final edited or imputation-revised variables. The changes described in this chapter were implemented during the normal editing and imputation process.

1.4.1 Changes to Accommodate Questionnaire Changes

Changes to the 2016 questionnaire were not as extensive as the changes to the 2015 questionnaire. Nevertheless, some questions for 2016 were modified from the 2015 survey. For

example, marital status questions that had been moved to a self-administered section later in the interview for the 2015 survey were moved back to the interviewer-administered section at the beginning of the interview for the 2016 survey. Also, changes were made to the prescription drug questions for 2016 that affected selected variables for the use and misuse of prescription drugs (see Section 1.4.1.8).

The following sections describe changes to the 2016 NSDUH instrument relative to 2015. Where relevant, the sections also briefly discuss the effects on editing and imputation procedures.

1.4.1.1 Marital Status

In 2015, the questions for current marital status and the number of times that respondents had been married moved from the interviewer-administered demographic section at the beginning of the interview to the audio computer-assisted self-interviewing (ACASI) section of the interview following the education questions. This move to ACASI in 2015 affected the quality of the data, especially for respondents aged 15 to 25. In particular, respondents aged 15 to 25 chose category 2 (widowed) in the marital status question more often in 2015 than in 2014, when the question was interviewer administered. Being widowed is rare for individuals aged 15 to 25. Because category 2 in the self-administered questions often means "no," respondents may have chosen category 2 in the self-administered marital status question to mean that they were not married but without reading or listening to the entire question or response options. Missing data for marital status (i.e., don't know, refused, or blank because of breakoffs) also increased from fewer than 20 respondents aged 15 or older in 2014 to nearly 400 respondents in this age group in 2015.

Therefore, SAMHSA decided to move the marital status questions back to the interviewer-administered demographic section at the beginning of the interview for 2016. The editing and imputation processes for marital status remained unchanged in 2016, except to change the names for the final edited and imputed marital status variables back to their names from 2014. Specifically, the names for edited variables MARITSTAT and MARIEDNUM from 2015 changed to MARITAL and NOMARR in 2016. The final imputation-revised variable changed from IRMARITSTAT in 2015 to IRMARIT in 2016. Use of the names for edited and imputed variables in 2016 that corresponded to the names in 2014 indicated that these data were comparable with the data from 2014 but were not comparable with the data from 2015. Additional details on the editing and imputation procedures for marital status are discussed in Chapter 4.

1.4.1.2 Highest Level of Education and Current Education

Other than the change for marital status that was noted previously, the interviewer-administered demographic section at the beginning of the interview did not change for 2016. For the highest grade or level of education completed, the logic for the consistency check between interviewer-administered question QD11 (highest level of education completed) and self-administered question QD18 (current grade or year in school) was updated slightly so that the consistency check was *not* triggered in 2016 if respondents reported that their highest level of education was the 12th grade but they had not received a high school diploma (QD11 = 13) and

they reported in QD18 that they were currently in the 12th grade. However, this change for 2016 had a minimal effect on editing procedures for the highest level of education completed (EDUHIGHST) because the decision was made for 2015 that no editing should be done if this pattern occurred in the 2015 data. No changes were needed for the imputation procedures for 2016.

Self-administered education question QD17 for current school enrollment changed from "Are you now attending or are you currently enrolled in school?" to "Do you go to school?" for 2016. An instruction was added for how respondents should answer QD17 if they were on a holiday or a break from school and intended to return to school when their break was over. This change for QD17 did not affect how the self-administered education data were edited for 2016. However, review of the final data for 2016 indicated that this change affected the comparability of the self-administered education data between 2015 and 2016. Consequently, the edited self-administered education variables were renamed for 2016 to reflect this break in comparability. The edited self-administered education variables do not undergo imputation, so no changes were required for the imputation procedures. Additional details on the editing and imputation procedures for the highest level of education completed are discussed in Chapters 4 and 5.

1.4.1.3 Military Families

Questions about other immediate family members who were serving in the United States military that had been added to the 2015 questionnaire were modified for 2016. Specifically, respondents in 2016 were instructed to include only the specific family members who were listed in the introductory text for question QD10D. Respondents no longer had the option to report that another immediate family member was serving, and the follow-up question for respondents to specify the other members of their immediate family who were serving also was deleted. Consequently, "OTHER, Specify" data were no longer available for editing the military family variables in 2016. These military family variables were renamed for 2016 because the changes affected the comparability of the military family variables between 2015 and 2016.

These edited military family variables do not undergo imputation. Therefore, the changes to the edited military family variables for 2016 did not affect the imputation procedures. Additional details on the editing procedures for the military family variables are discussed in Chapter 5.

1.4.1.4 Health Insurance

Minor changes were made to the wording of questions for Medicare and Medicaid in 2016 to change "disabled persons" to "people with disabilities." These changes did not affect the editing or imputation procedures for health insurance or the comparability of health insurance data between 2015 and 2016.

The wording of question QHI06 was changed in 2016 to indicate that private health insurance included plans that respondents purchased through the Health Insurance Marketplace. Comparison of trend data in NSDUH and the National Health Interview Survey indicated that this change for 2016 did not affect the comparability of the data between 2016 and prior years for any health insurance coverage and private health insurance coverage. Therefore, this change

to QHI06 did not affect the editing and imputation procedures. Further details on the editing and imputation procedures related to health insurance variables are provided in Chapter 13.

1.4.1.5 Income

Minor changes were made to the questions in the income section for 2016. In particular, question QI107N was revised to move the instruction that benefits from the Supplemental Nutrition Assistance Program (SNAP) do not include the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) or free or reduced school lunches to be included with the other definitional text at the beginning of the question. In 2015, the instruction not to include WIC or free or reduced school lunches followed the question about the receipt of SNAP benefits. This change did not affect the editing or imputation procedures for income or the comparability of income data between 2015 and 2016. Further details on the editing and imputation procedures as applied to income data in 2016 are described in Chapter 12.

The order of cell phone questions CELL1 and CELL2, which appear in the income section of the questionnaire, also changed for 2016 to move CELL1 (any working telephone at the residence that is not a cell phone) to appear after CELL2 (whether anyone at the residence has a working cell phone number). This change did not affect the editing procedures or the comparability of the data between 2015 and 2016. The changes to the order of the cell phone questions also did not affect the imputation procedures because these variables do not undergo imputation.

1.4.1.6 Roster and Household Composition

In 2016, the questionnaire levels describing the relationships of household roster members to the respondent were updated and expanded to a more detailed list of relationships. The relationship codes new to the 2016 survey year include sibling-in-law, aunt or uncle, niece or nephew, cousin, and ex-spouse.⁵ The new detailed relationship levels were captured by the new detailed roster relationship variables ROSRLTD1-ROSRLTD25. To maintain comparability with previous survey years, ROSRLT1-ROSRLT25 were created by recoding the new relationship level to match the original 14 relationship levels captured in 2015 and prior survey years.

When using the detailed relationship levels to create ROSRLT1-ROSRLT25, the relationships sibling-in-law, aunt or uncle, niece or nephew, cousin, and ex-spouse were coded to "13 – Other relative." This decision was based on the proxy section of the questionnaire, which fills in response options based on reported relatives in the household, and all new categories were considered relatives. In addition, the CAI suppression checks relying on spouses and unmarried partners were updated to include ex-spouses. Additional details on the editing and imputation procedures for the roster and household composition variables are provided in Chapter 11.

⁵ For a detailed description of the roster relationship levels available prior to 2016, see [Table 11.2](#) in Chapter 11. For a detailed description of the roster relationship levels beginning in 2016, see [Table 11.3](#).

1.4.1.7 Proxy Variables

The household relationship of a selected proxy to the respondent (PRXRELAT) was another variable affected by the change in household relationship levels. To maintain consistency with previous years, PRXRELAT was created from the original 14 levels of household relationships available prior to 2016.⁶ Starting in 2016, the variable PRXRELATD was created to account for the newly available detailed relationship levels.⁷ Further details on the editing procedures for the proxy variables are discussed in Chapter 11.

1.4.1.8 Prescription Drug Variables

The following changes were made to the 2016 NSDUH questionnaire with regard to the prescription drug questions:

- Because respondents were instructed to key "95" in the questions for any past year use of specific prescription drugs if they had never used any of the prescription drugs in a given list, an example question was added to the ACASI tutorial to show respondents how to answer these types of questions.
- Consistency checks were added to the sections for any past year use of prescription drugs if respondents keyed three responses in a row of only "1" or only "2" to the first three items in a section. Respondents who triggered these consistency checks were asked to confirm that their answers were correct or to change their answers if necessary. In previous sections of the interview, responses of "1" or "2" typically meant "yes" or "no," respectively. Therefore, these consistency checks (along with the tutorial question mentioned above) were designed to reduce the occurrence of patterned responses in the prescription drug data and potential misreporting of past year use or misuse of specific prescription drugs.
- The response option for Tylenol[®] with codeine 3 or 4 in the section for any past year use of prescription pain relievers was modified to clarify that this drug was not the same as over-the-counter Tylenol[®]. This change was made to reduce potential confusion between these two similarly sounding drug names.
- The following specific prescription drugs from 2015 were removed as response options because they had been discontinued or were reported infrequently in the 2015 data:
 - Roxicet[®] (a pain reliever that contains oxycodone),
 - Actiq[®] (a pain reliever that contains fentanyl),
 - buspirone (a tranquilizer),
 - hydroxyzine (a tranquilizer),

⁶ See Table 11.6 in Chapter 11 for the levels of PRXRELAT.

⁷ See Table 11.7 in Chapter 11 for the levels of PRXRELATD.

- meprobamate (a tranquilizer), and
- Ritalin[®] SR (a stimulant).⁸
- Buprenorphine plus naloxone was added as a response option to the prescription pain relievers section.

The new tutorial question and consistency checks considerably reduced the amount of editing that was needed in 2016 to set variables to "bad data" because of patterned responses, especially for respondents who chose the prescription drugs that were the second response option in the items for any past year use.⁹ However, these consistency checks had relatively small effects on the data compared with the overall numbers of respondents who reported the past year use or misuse of any prescription drug in a given psychotherapeutic category (e.g., pain relievers) or subtypes of prescription drugs in a given psychotherapeutic category (e.g., hydrocodone products). Therefore, the additions of the new tutorial question and consistency checks were not considered to have affected the comparability of the 2016 prescription drug data.

Editing programs were updated for 2016 to remove variables that corresponded to prescription drugs that had been dropped from the questionnaire and to add new variables for buprenorphine plus naloxone. The edited variables for the past year use and misuse of Tylenol[®] with codeine 3 or 4 and Ritalin[®] LA were renamed in 2016 to reflect the questionnaire changes for those specific prescription drugs. Otherwise, the basic procedures for editing the prescription drug data did not change for 2016. In particular, the basic questions did not change in 2016 for creating the edited recency variables for any use and misuse and for creating the variables for the initiation of the misuse of prescription drugs in a given prescription drug category (e.g., pain relievers) in the past 12 months.

The determination was made based on further analysis of the 2016 data that estimates for the past year use and misuse of codeine products were not comparable between 2015 and 2016 because of the additional instruction for any past year use of Tylenol[®] with codeine 3 or 4 (see the 2016 NSDUH methodological summary and definitions report [Center for Behavioral Health Statistics and Quality, 2017d]). However, this change and the other changes for prescription pain relievers that were described previously did not affect the comparability of the estimates for the use or misuse of any prescription pain relievers or any prescription psychotherapeutic drugs. Similarly, the changes described previously for tranquilizers and stimulants did not affect the comparability of the prescription drug data between 2015 and 2016.

The changes to the prescription drug data for 2016 did not affect the prescription drug measures that were imputed. However, minor updates were made to the modPMN imputation procedures for the age at first misuse and the date of first misuse among past year initiates to account for the specific prescription drugs that were removed or added to the questionnaire in 2016. Although some variables for the use or misuse of subtypes of prescription drugs in the past 12 months were affected by these questionnaire changes (e.g., any use or misuse of codeine

⁸ Respondents in 2015 were asked to indicate whether they used "Ritalin SR or Ritalin LA" in the past 12 months. In 2016, respondents were asked to indicate whether they used "Ritalin LA" but not Ritalin SR.

⁹ The individual effects of the new tutorial question and consistency checks on the 2016 data cannot be determined.

products), these variables (with the exception of OxyContin[®]) did not undergo statistical imputation to replace missing data with valid data from donor respondents. Rather, missing values in variables pertaining to subtypes of prescription drugs were coded as "no use" or "no misuse" in the past 12 months as part of the procedures for creating recoded variables for use in the NSDUH detailed tables. A full description of the editing and imputation procedures for the prescription drug variables in the 2016 NSDUH is provided in Chapter 6.

1.4.2 Changes to Streamline Editing and Imputation Procedures

Between 2015 and 2016, steps were taken toward better integration of the editing and imputation procedures. Efforts focused mainly on prescription drugs because the prescription drug sections that were redesigned for the 2015 survey are structured very logically to reduce the opportunity for respondents to provide inconsistent answers. Resolution of the limited inconsistent response patterns that remain in the data and that have been determined by SAMHSA to be unallowable in the final analytic dataset can be handled in both the editing and imputation procedures through prespecified "edit rules." Details for implementing these edit rules and how these rules affect the editing and imputation procedures for prescription drugs are described in Chapter 6.

In addition, the imputation programs for the 30-day frequency for all seven drug categories for which modPMN is used (hallucinogens, inhalants, methamphetamine, and the four prescription drug categories) were modified so that several steps not involving any modeling or stochasticity are handled in a preliminary setup program. These details are also discussed in Chapter 6. The advantages of these small modifications are twofold. First, a smaller dataset can be carried through the modPMN programs. Second, by sharing common SAS[®] programs, statisticians can implement changes more efficiently and with less likelihood of making errors.

Another streamlining measure implemented in 2016 involved the imputation of the hallucinogen child recency variables. Beginning in 2016, the imputation process for these variables was updated to account for the edited age-at-first-use (AFU) and date-of-first-use (DFU) variables for overall hallucinogens. By accounting for these variables during the imputation of child recency variables, potential parent/child inconsistencies can be prevented earlier in the process. Prior to this modification, these types of issues would not have been identified until the imputation of hallucinogen child AFU/DFU and would have required reprocessing the child recency variables. Although no such inconsistencies were found in 2015 that required reprocessing of the child recency variables, these modifications should now eliminate this possibility for 2016 and future survey years.

1.4.3 Changes Involving Minor Corrections and Enhancements

Five changes were made to the 2016 editing and imputation procedures that involved minor corrections or enhancements. None of these changes had a significant impact on the estimates.

The first change involved updating the hot-deck procedures of PMN to accommodate new missingness patterns and to add attempts to find a donor by dropping likeness constraints when necessary. In 2016, a new missingness pattern was identified for cocaine and crack

recency/frequency where the recipient was a past year user of cocaine but use in the past month was unknown. This recipient was also known to be a past year but not past month user of crack, and the 12-month frequency for both cocaine and crack were missing. In order to find a donor for this particular case, the hot-deck step had to be updated, and the corresponding documentation was also updated to reflect this change. The hot-deck steps and documentation for health insurance imputations were also updated in 2016 to add attempts to find donors for particular missingness patterns.

A second change made between the 2015 and 2016 surveys involved the imputation indicators for the imputation-revised substance dependence and abuse variables.¹⁰ Beginning in 2016, respondents who were not past year users/misusers of modPMN drugs received values of 9 (legitimate skip) instead of 1 (questionnaire data) for the imputation indicators for substance dependence and abuse. This treatment is more consistent with other imputation indicators. One way these imputation indicators are atypical of most of the others, however, is in the handling of respondents who were imputed out of the domain (i.e., respondents whose domain indicator was missing but was later imputed to be negative). For most other imputation indicators, respondents like these receive values indicating statistical imputation (see Section 3.1) instead of values indicating a legitimate skip. This handling of imputation-revised skips simplifies the procedures by which imputation indicators are assigned for variables governed by skip logic, and data users can always trace back the domain variables if they are interested in whether skipped cases were imputed to be skips or were skipped based on actual questionnaire responses.

A third change was made to the imputation procedures for 2016 to account for a donorless case for the imputation-revised measure of the total number of children younger than 18 in a household, IRKID17. The previously imputed values for IRHH65 (imputation-revised adults in the household aged 65 or older) for the donorless case created a bounding inconsistency in the pre-hot-deck step for IRKID17. In 2015 and prior years, if the upper bound derived from the edited variable KID17 was missing, then the upper bound for IRKID17 was set equal to the imputation-revised household size, IRHHSIZE. This bounding did not consider the number of adults or adults aged 65 or older in the household. To remedy this situation and to prevent future bounding inconsistencies, the upper bound for IRKID17 was modified to account for the minimum number of adults that are known to be in the household in addition to the total household size.

A fourth change was made in 2016 to correct a minor documentation error. The skip code for past year any use of OxyContin®, IROXCNANYR, and past year misuse of OxyContin®, IROXCNNMYR, incorrectly indicated that the respondent never (mis)used OxyContin®. This level description was corrected for 2016 to indicate that the respondent did not (mis)use OxyContin® in the past year. The 2015 documentation was also corrected and will be reflected in any future redeliveries of these data files.

A fifth change involved a refinement to the edits for the nicotine (cigarette) dependence variables in 2016. Specifically, editing of the nicotine dependence variables prior to 2016

¹⁰ This change was applied retroactively to the 2015 data. The imputation indicators were redelivered to the 2015 master file in March 2017. Versions of the 2015 data file dated after March 2017 have imputation indicators with values of 9.

assigned a code of 93 (DID NOT USE CIGARETTES IN THE PAST 30 DAYS) to the skipped nicotine dependence variables for respondents who last smoked cigarettes at some point in the past 3 years (CIGREC = 39). Because cigarette use at some point in the past 3 years could include the past 30 days, the editing procedures for the nicotine dependence variables were updated for 2016 to leave these variables blank rather than to assign a code of 93. This change for 2016 had a minimal effect on the editing of the nicotine dependence variables because fewer than five respondents in most years were assigned a code of 39 in the editing of CIGREC. This change also had a minimal effect on the imputation of the nicotine dependence data because respondents whose nicotine dependence data are all missing are not imputed. The only difference between 2016 and prior years was that the imputed variables in 2016 would be coded as 98 (BLANK/DID NOT ANSWER) instead of being coded as 93 (DID NOT USE CIGARETTES PAST MONTH).

This page intentionally left blank

2. Procedures and General Principles for Editing the Computer-Assisted Data

2.1 Introduction

This chapter provides an overview of the procedures and general principles for editing the computer-assisted interviewing (CAI) data from the National Survey on Drug Use and Health (NSDUH). Logical editing typically uses data from elsewhere within the same respondent's record in a deterministic manner to reduce the occurrence of missing or ambiguous data or to resolve inconsistencies between related variables.¹¹ In contrast, statistical imputation procedures (see Chapter 3) apply probabilistic (stochastic) statistical methods to identify *another* respondent (i.e., a "donor") whose data are used to replace (1) missing values in the "recipient" respondent's record with nonmissing values from the donor or (2) ambiguous responses in the recipient respondent's record with more specific information from the donor.

As an example of ambiguous data, the CAI logic for the marijuana through methamphetamine sections in 2016 requires respondents to report that they have used a particular substance (e.g., marijuana) at least once in their lifetime in order to be asked when they last used it. However, a respondent can report lifetime use of a particular drug but not provide a definitive answer for when he or she last used the substance. In the subsequent imputation procedures for this recipient record, a donor is identified who specifically reported last using the substance in the past 30 days, more than 30 days ago but within the past 12 months, or more than 12 months ago. Here, the ambiguous answer in the recipient's record for most recent use is replaced with one of the more specific responses supplied by the donor record.

Section 2.2 describes the procedures for initially processing the transmitted NSDUH interview data to get it in a form for further data processing. Subsequent editing and coding steps are described in Section 2.3. The final section in this chapter, Section 2.4, discusses additional principles of data processing and editing that were applied once the transmitted interview data had been processed and cases with questionable data had been identified.

2.2 Initial Processing of Transmitted NSDUH Interview Data

The collected interview data were transmitted from the field as ASCII files, and daily SAS[®] datasets were created from these files. This daily processing included the following activities as part of the creation of initial unedited interview data files:

¹¹ There are a few situations where data from outside the respondent's record is used in logical editing. For example, some editing procedures involve data from the screener where an eligible member of the dwelling unit reports basic information about all members of the dwelling unit. Further, in situations where two members of the same dwelling unit are selected for the survey and complete the interview, data from the second respondent's record may be used in logical editing of some variables in the first respondent's record, or vice versa. This procedure allows use of information from both respondents to determine relationships among household members.

- assignment of standard NSDUH missing data codes (Section 2.2.1),
- remapping of responses to "enter all that apply" questions (Section 2.2.2), and
- identification of "usable" cases (Section 2.2.3).

Each day's processed SAS dataset was merged with the transmitted data to date until the end of the quarter (e.g., January through March = Quarter 1) when a cumulative data file (Section 2.2.4) was produced that contained all transmitted cases from that quarter.

2.2.1 Assignment of Standard NSDUH Missing Data Codes

A key activity in the initial processing of transmitted, unprocessed interview data involved the assignment of standard NSDUH missing data codes. The Blaise program for the CAI instrument uses a code of 8 (or 98 or 998, etc.) to denote responses of "refused" and a code of 9 (or 99 or 999, etc.) to denote responses of "don't know." However, in the NSDUH, a code of 98 is used to indicate when a variable is blank (i.e., not answered), and a code of 99 is used in editing procedures to indicate when the variable corresponding to a question was legitimately skipped. Therefore, the codes for missing data that were supplied by the Blaise program were replaced with the standard NSDUH codes for "don't know" (DK) and "refused" (REF). Assignment of codes as part of the editing procedures to indicate when a question was legitimately skipped is discussed in Section 2.4.2.

The following standard codes for missing data were relevant to the 2016 CAI data, depending on the number of digits for a given variable:

- 94 (or 994 or 9994, etc.) = DON'T KNOW (DK);
- 97 (or 997 or 9997, etc.) = REFUSED (REF); and
- 98 (or 998 or 9998, etc.) = BLANK (i.e., nonresponse [NR]).

Blanks were the default code for missing data if questions were not answered because of the CAI skip logic or if respondents broke off the interview before reaching a particular question. Section 2.4.3 describes situations in which variables retained codes of blank when respondents answered "don't know" to a question that governed CAI skip logic. Section 2.4.2 describes situations in which codes of blank in the unedited data could be replaced with codes to indicate that the questions had been legitimately skipped.

Codes for missing data in most unedited variables were two digits in length (i.e., 94, 97, or 98). For some variables, however, these values were part of the allowable range of responses. Questions that asked respondents to report the age when they first used a particular drug, for example, had an allowable range of up to 110 years.¹² For the variables corresponding to these age-related questions, the codes for missing data were three digits in length (i.e., 994, 997, or 998).

¹² Questions for prescription drugs referred to the age when respondents first misused specific prescription drugs that they misused in the past 12 months.

Finally, the CAI logic governed whether respondents were asked additional questions about a topic based on their age, gender, or answers to other preceding questions. For example, questions in the interview about pregnancy applied only to females aged 12 to 44 (Section 10.2.1). These pregnancy questions were skipped if interviewers reported that a respondent was male or if the CAI program recorded that a female respondent was aged 45 or older. When questions were skipped because the criteria in the CAI program were not satisfied for administering the questions, the unedited variables corresponding to the skipped questions retained a code of "blank" (i.e., 98, 998, etc.). In subsequent editing (described in Section 2.4.2), these variables were examined more closely to determine whether the questions had been legitimately skipped (i.e., they were not applicable) or whether they should retain codes of "blank."

2.2.2 Remapping of Responses to "Enter All That Apply" Questions

A second activity associated with the initial processing of transmitted interview data involved the remapping of responses to "enter all that apply" type questions, which allowed respondents to choose as many responses from a given list as applied. Respondents who wanted to report more than one answer from a list did so by typing the numeric codes that corresponded to the applicable responses and by separating each entry with a space.

The CAI program captured information from these "enter all that apply" questions as separate variables in the order that respondents keyed their answer choices. For example, the transmitted CAI data included five separate variables for the specific pain relievers listed in question PR01 to accommodate reports of any use in the past 12 months of these pain relievers, or to allow respondents to choose a response of "95" to indicate that they did not use any of the past relievers listed in PR01 in the past 12 months.¹³ Consequently, an "enter all that apply" variable in the transmitted data could have a different meaning across respondents, depending on how respondents answered the question. For example, if a respondent reported that he or she had not used any of the pain relievers listed in PR01 in the past 12 months, then a response of 95 would be captured in the *first* variable from the transmitted data, and the remaining "unused" variables from question PR01 would be blank. If a second respondent reported the use of generic hydrocodone and no other pain relievers from PR01 in the past 12 months, then a response of 5 (corresponding to the response option for hydrocodone) would be captured as the first variable in the transmitted data for PR01. If a third respondent chose Vicodin[®] and hydrocodone as his or her first and second responses from the list, then the first variable from PR01 would be occupied by the response for Vicodin[®], the second would be occupied by hydrocodone, and the remaining variables would be blank.

If these "enter all that apply" variables were allowed to remain in the form in which they came from the transmitted CAI data, then these variables would have retained the information about the order in which respondents chose their answers. However, this variable structure makes it difficult to analyze the data because the first variable could indicate use of a specific prescription drug for some respondents and nonuse of any of the prescription drugs in the list for other respondents. It would be more straightforward for subsequent editing if an individual

¹³ The CAI program does not allow respondents to report the use of specific prescription drugs in a list and also that they did not use any of the listed prescription drugs in the past 12 months.

"enter all that apply" variable had one, and only one, meaning for all respondents. To accomplish this purpose, part of the processing of transmitted data involved remapping responses to these "enter all that apply" questions according to the order of the response choices that respondents saw in the questionnaire. For example, six discrete variables were created as part of the remapping process to capture all reports of any use of Vicodin® in the past 12 months in the variable PR011 (i.e., the first variable in the series); reports of any use of Lortab® in the variable PR012; reports of any use of Norco® in the variable PR013; reports of any use of Zohydro® ER in the variable PR014; reports of any use of generic hydrocodone in the variable PR015; and reports of nonuse of any prescription drug in the list in the past 12 months (i.e., responses of 95) in the variable PR016. If respondents did not report the use of a specific drug from the list of pain relievers in PR01, then the remapped variable for that pain reliever was assigned a code of 98 (i.e., blank). Thus, if respondents chose an answer of 95, then PR016 was given a code of 95 and the remaining variables for PR011 through PR015 were assigned codes of 98. If respondents reported the use of at least one drug from the list in PR01, then PR016 received a code of 98.

In addition to choosing one or more applicable responses from a list, respondents could use function keys to answer "don't know" or "refused" as their first response to these "enter all that apply" types of questions. In situations where respondents answered "don't know," it would be reasonable to infer that the respondent did not know which particular item on the list applied to him or her. For example, if a respondent answered question PR01 as "don't know," this was inferred to mean that the respondent did not know whether he or she had used Vicodin®, Lortab®, Norco®, Zohydro® ER, or generic hydrocodone in the past 12 months. In this case, a code of "don't know" was propagated to each of the recoded "enter all that apply" variables as part of the daily processing of the transmitted data. Similarly, if a respondent *refused* to answer question PR01, a refusal code was propagated to all of the recoded variables on that list as part of the daily processing of the transmitted data.

2.2.3 Identification of Usable Cases

Once standard missing data codes had been assigned and the responses to "enter all that apply" questions had been remapped, the third key step in the preliminary processing of transmitted NSDUH data established the minimum item response requirements necessary for cases to be retained for weighting and further analysis (subsequently referred to as "usable" cases). These rules were designed to eliminate cases with unacceptable levels of item nonresponse (i.e., missing data), thereby retaining cases with lower levels of missing data and reducing the amount of statistical imputation needed for any given record. In addition, requiring lifetime use or nonuse to be fully defined for at least one substance (i.e., no missing data) can allow data for that substance to be used in statistical imputations for other substances with missing data for lifetime use or nonuse.

In order for a case to be considered usable in NSDUH, the requirements listed below must be met.¹⁴ The usable case requirements were updated beginning with the 2015 survey to reflect the changes to the questionnaire for smokeless tobacco, hallucinogens, inhalants,

¹⁴ The historical background and considerations for establishing usable case rules for the CAI data are discussed in a methodological chapter on editing the 1999 CAI data (Kroutil & Myers, 2002).

methamphetamine, and prescription drugs.¹⁵ The term "gate question" is used in connection with the usable case criteria because an affirmative response to these questions (e.g., "Have you ever, even once, used marijuana or hashish?") opens the "gate" to a series of other questions on use of the substance, whereas a negative response closes the "gate" and leads to the skipping of all other questions on use of that substance.

The requirements for a CAI record in 2016 to be considered usable were as follows:

1. The lifetime cigarette question (i.e., "Have you ever smoked part or all of a cigarette?") had to have been answered as "yes" or "no." This requirement was set so that lifetime use or nonuse would be fully defined for at least one substance. Consequently, data about lifetime use or nonuse of cigarettes could be used in subsequent statistical imputations for other substances where lifetime use or nonuse was undefined.
2. Responses to questions on at least nine (9) of the following additional substances had to contain information about use or nonuse: (a) smokeless tobacco, (b) cigars, (c) alcohol, (d) marijuana, (e) cocaine (in any form), (f) heroin, (g) hallucinogens, (h) inhalants, (i) methamphetamine, (j) pain relievers, (k) tranquilizers, (l) stimulants (i.e., independent of methamphetamine), and (m) sedatives. Crack cocaine was not included in the usable case criteria because the logic for asking about crack cocaine was dependent upon the respondent having answered the lifetime cocaine question as "yes." Although the CAI instrument also asked about pipe tobacco, this was not included in the usable case criteria because there was only one other question about pipe tobacco in addition to the lifetime pipe tobacco use question.

The usability criterion for smokeless tobacco through heroin and for methamphetamine (i.e., the drugs listed above as items a to f and item i) was that lifetime use or nonuse must be determined. For the "multiple gate" sections for hallucinogens and inhalants (items g and h above), at least one gate question in the series for that section was required to have an answer of "yes" or "no."¹⁶ Any of the following met the usability criteria for prescription drugs (items j to m above):

- use in the past 12 months of at least one specific prescription drug in a category (e.g., pain relievers) was reported;
- lifetime use or nonuse of any prescription drug in the category was reported; or
- nonuse of *all* specific prescription drugs in the past 12 months was reported, regardless of whether lifetime use or nonuse could be determined.

¹⁵ For more information about the development of the usable case requirements for the redesigned NSDUH questionnaire, see the 2012 Questionnaire Field Test report (Center for Behavioral Health Statistics and Quality, 2014a).

¹⁶ For hallucinogens and inhalants, the logic for asking more detailed questions about use of that category of drugs was based on respondents' answers to multiple "yes/no" questions about the lifetime use or nonuse of specific drugs within that category (e.g., lifetime use or nonuse of the following specific hallucinogens: lysergic acid diethylamide [LSD]; phencyclidine [PCP]; peyote; mescaline; psilocybin/mushrooms; Ecstasy or "Molly" [3,4-methylenedioxymethamphetamine (MDMA)]; ketamine; DMT, AMT, or "Foxy"; *Salvia divinorum*; or "any other" hallucinogen).

The interview also included follow-up probes for respondents who initially refused to answer a gate question or an entire series of gate questions. Follow-up probes were included in the following sections that were relevant to the usable case criteria: cigarettes, smokeless tobacco, cigars, alcohol, marijuana, cocaine, heroin, specific hallucinogens (i.e., LSD, PCP, Ecstasy/"Molly" [MDMA]), any use of inhalants, and methamphetamine. In particular, if respondents changed their initial refusal to a response of "yes" or "no," they were considered to have provided usable data to that drug's gate information.

Table 2.1 lists the follow-up questions in the initial substance use sections that were administered when respondents initially refused a gate question, including follow-up questions that were administered when there was a single gate question (i.e., for cigarettes through heroin and for methamphetamine) and follow-up questions that were administered in the sections for hallucinogens and inhalants, which had multiple gate questions. Table 2.1 also lists the implications for the usable case criteria according to how respondents answered these follow-up questions.

Table 2.1 Effects on the Potential Usable Status of a Case Based on Responses to Follow-Up Questions for Refusals to Gate Questions in the Initial Substance Use Sections

Section or Drug (if Applicable)	Question Number for Follow-Up Question	Consequence if Follow-Up Response Is "Yes" or "No"	Consequence if Follow-Up Response Is DK or REF
Cigarettes	CGREF1	Meets usable case criterion for cigarettes	Not a usable case
Smokeless Tobacco	CGREF3	Meets usable case criterion for snuff	Does not meet usable case criterion for snuff ¹
Cigars	CGREF4	Meets usable case criterion for cigars	Does not meet usable case criterion for cigars ¹
Alcohol	ALREF	Meets usable case criterion for alcohol	Does not meet usable case criterion for alcohol ¹
Marijuana	MJREF	Meets usable case criterion for marijuana	Does not meet usable case criterion for marijuana ¹
Cocaine	CCREF	Meets usable case criterion for cocaine	Does not meet usable case criterion for cocaine ^{1,2}
Heroin	HEREF	Meets usable case criterion for heroin	Does not meet usable case criterion for heroin ¹
Hallucinogens/LSD	LSREF1	Meets usable case criterion for hallucinogens	Does not affect usable case status for hallucinogens if another gate question (or follow-up question) is answered as "yes" or "no"
Hallucinogens/PCP	LSREF2	Meets usable case criterion for hallucinogens	Does not affect usable case status for hallucinogens if another gate question (or follow-up question) is answered as "yes" or "no"
Hallucinogens/Ecstasy	LSREF3	Meets usable case criterion for hallucinogens	Does not affect usable case status for hallucinogens if another gate question (or follow-up question) is answered as "yes" or "no"

Table 2.1 Effects on the Potential Usable Status of a Case Based on Responses to Follow-Up Questions for Refusals to Gate Questions in the Initial Substance Use Sections (continued)

Section or Drug (if Applicable)	Question Number for Follow-Up Question	Consequence if Follow-Up Response Is "Yes" or "No"	Consequence if Follow-Up Response Is DK or REF
Inhalants	INREF ³	Meets usable case criterion for inhalants	Does not meet usable case criterion for inhalants ¹
Methamphetamine	MEREF	Meets usable case criterion for methamphetamine	Does not meet usable case criterion for methamphetamine ¹

DK = don't know; REF = refused.

¹ Overall status as a usable case will still be met if the usable case criteria are met for a sufficient number of other substances or sections.

² The interview includes a follow-up probe if the gate question for crack cocaine is refused (CKREF), but crack cocaine is not included in the criteria for identifying usable cases.

³ Follow-up probe is asked if the respondent refused to answer all gate questions for that section.

In addition, the CAI program terminated interviews during the initial demographics questions if respondents were ineligible for the survey (i.e., younger than age 12 or on active duty in the U.S. military). These ineligible cases by definition did not meet the usable case criteria because the interviews were terminated before the respondents were asked the first cigarette use question.

Meeting the usable case criteria was a necessary, but not sufficient, requirement for a case to be considered a final respondent. Cases that had sufficient data to meet the usable case criteria could still be treated as nonrespondents if their interview data appeared to be of poor data quality, as evidenced by potential response pattern problems (Section 2.3.2).

2.2.4 Creation of Cumulative Quarterly Unedited Data Files

Following daily processing of the data, each day's SAS dataset was merged with the cumulative data that had been transmitted up to that point in the quarter. At the end of the quarter, a complete data file was produced that contained all cases that had been transmitted during the quarter. Each quarterly data file then underwent additional initial cleaning and processing (prior to the editing procedures) to modify or correct field errors, such as erroneous ID entries by the field staff. The cleaned-up (but otherwise unedited) SAS datasets from the first two quarters (also known as 6-month data) and from all four quarters (also known as 12-month data) served as the usual starting points for the subsequent logical editing procedures that are described in Chapters 4 through 14.¹⁷

¹⁷ Edits are run on preliminary data from the first two quarters of a given survey year to identify any updates that need to be made to the programs for use with the full data from all four quarters. Running the edits on data from the first two quarters is particularly useful for testing the programs for any sections of the CAI instrument that are new or have changed since the previous survey year.

2.3 Preliminary Editing and Coding of Processed Interview Data

In addition to procedures that were described in Section 2.2 following receipt of transmitted data from the field, preliminary coding and processing of unedited interview data encompassed the following activities:

- coding of "OTHER, Specify" data (Section 2.3.1);
- investigation of response patterns in records that otherwise met the usable case criteria (Section 2.3.2); and
- edits to "date-dependent" variables (if applicable) when the interview date was judged to be questionable (Section 2.3.3).

The first two of these activities could occur or did occur prior to creation of the cumulative quarterly unedited data files (Section 2.2.4). Edits to date-dependent variables were not performed until final interview dates had been created for respondents (Section 4.2.1).

Note that a code to denote "bad data" (i.e., to indicate an inconsistency or some other problem in the original data) could be assigned to variables during any of these activities. The following codes were assigned to denote "bad data": 85 (or 985, or 9985, etc.) = BAD DATA Logically assigned. Codes for bad data were treated as missing values. Any assignment of bad data codes was done in subsequent editing steps, not as part of the nightly processing of transmitted data.

2.3.1 Coding of "OTHER, Specify" Data

This activity took alphanumeric (text) answers that respondents or interviewers had typed (e.g., specific other drugs that were used or misused, specific other payment sources of treatment) and converted them into numeric codes. These alpha answers (and the numeric codes resulting from them) are referred to as "OTHER, Specify" data.

Coding of the "OTHER, Specify" variables was accomplished through computer-assisted procedures.¹⁸ "OTHER, Specify" responses were first converted to all capital letters because respondents could use different combinations of uppercase or lowercase characters to provide an otherwise identical response. If an exact match was found between what the respondent keyed and an existing entry in the data dictionary (e.g., "ALCOHOL"), the computer-assisted procedures assigned the appropriate numeric code (e.g., 807 for alcohol). The system could also accommodate commonly encountered misspellings (e.g., "ALCHOHOL").

Typed answers that the respondent provided that did not match an existing data dictionary entry were reviewed by analysts on a flow basis during the quarter to determine whether an existing code should be assigned to the response or a new code should be created. Based on these decisions, new entries were added to the relevant dictionaries on a quarterly

¹⁸ A system has been in place since 2002 for the daily coding and processing of the "OTHER, Specify" variables for drugs and tobacco brands. This system now encompasses the daily coding and processing of most "OTHER, Specify" variables from the CAI instrument that underwent assignment of numeric codes, except for the codes pertaining to race or ethnicity that are described in Section 2.3.1.3.

basis—including new dictionary entries corresponding to existing codes—for use in daily processing of data that were transmitted from the field in subsequent quarters. Analysts could also decide not to add a particular response to the data dictionary, in which case the response would be output for review and coding on a case-by-case basis if it was reported in a future quarter or survey year. In addition, analysts reviewed the codes that were assigned through the computer-assisted process to verify that these codes were being assigned correctly. Over time, these procedures have reduced the turnaround time and burden on analysts for producing the coded "OTHER, Specify" data.

Particularly for other drugs and other tobacco brands (Sections 2.3.1.1 and 2.3.1.2, respectively), which used the same data dictionaries for a number of "OTHER, Specify" variables, matching a written response to a numeric code in the data dictionary would result in that code being assigned *no matter where* respondents typed that answer. For example, a given entry in the data dictionary for drugs would be assigned the same code if it appeared in the hallucinogens section one year and appeared in the pain relievers section another year. Similarly, an entry in the data dictionary for tobacco brands would be assigned the same code regardless of the type of tobacco where the response appeared. In some situations, however, the same response could have different meanings depending on the context. For example, a given tobacco brand name with no other associated information could apply to a brand of cigarettes or to a brand of cigars, depending on whether it was specified as a cigarette brand or as a cigar brand.

In 2015, the data dictionary for illicit drugs was revised to reflect changes to the prescription drug sections. These changes are discussed further in Section 2.3.1.1 and in Section C.4 in Appendix C. [Table C.5](#) in Appendix C also shows changes for prescription drug codes between 2014 and 2015.

As with the "enter all that apply" data that were discussed previously, respondents could answer the "OTHER, Specify" questions as "don't know" or "refused," which were then reassigned to the respective codes of 9994 or 9997, as described in Section 2.2.1. Respondents could also type in an equivalent response to "don't know" (e.g., "no idea") or "refused" (e.g., "too personal"). These equivalent responses were assigned the relevant codes for missing data (Section 2.2.1) as part of the coding procedures. For typed responses that were nonsensical or otherwise nonresponsive to the request to specify additional information, codes for bad data (Section 2.3) were assigned.

The remainder of the discussion in this section focuses on issues related to coding of the "OTHER, Specify" data according to the type of other information that was requested from respondents:

- other drugs (Section 2.3.1.1),
- other tobacco brands that respondents used most commonly in the past 30 days (Section 2.3.1.2),
- other race or ethnicity (Section 2.3.1.3), and

- additional "OTHER, Specify" data in the remaining self-administered sections (Section 2.3.1.4).¹⁹

Except for mentions of other drugs (Section 2.3.1.1), "OTHER, Specify" data in the 2016 NSDUH typically were intended to capture a single "other" response (e.g., most important other reason for not receiving mental health treatment in the past 12 months). If respondents typed in responses for which multiple codes could apply (e.g., multiple reasons for not receiving mental health treatment instead of the single most important reason), the standard procedure was to assign a code to the first response that could be coded. Relevant sections elsewhere in this report indicate when exceptions were made to this more general approach (e.g., coding of "OTHER, Specify" data for youth mental health service utilization, described in Section 10.6).

2.3.1.1 Other Drugs

In the interview sections for hallucinogens and inhalants, respondents could report lifetime use of drugs other than those that were specifically asked about in the respective sections. For prescription drugs (i.e., pain relievers, tranquilizers, stimulants, and sedatives), respondents first were asked if they used other prescription drugs in that category in the past 12 months besides the specific prescription drugs that they were asked about in the respective screener sections (see [Table 1.1](#) and Section 1.4.1 in Chapter 1); respondents who reported any use of other prescription drugs in the past 12 months could subsequently report that they misused other prescription drugs in that period. Respondents also could report lifetime injection of drugs other than heroin, cocaine, methamphetamine, or other stimulants in the special drugs section. In the substance use treatment section, respondents could report current or most recent treatment or the perceived need for treatment in the past 12 months for drugs other than those for which they reported lifetime use in the corresponding substance use sections at the beginning of the interview. Respondents who reported the use or misuse of other drugs were asked to specify the names of these other drugs.

Codes for prescription drugs were revised in 2015 in response to changes to the prescription drug questions. Specifically, if prescription drug questions beginning in 2015 applied to one (and only one) prescription drug, then the "OTHER, Specify" codes were revised accordingly. Prior to 2015, for example, code 434 for pain relievers applied to more than one type of hydrocodone product. Beginning in 2015, however, for purposes of editing the 2015 data (see Section 2.4.5), code 434 was used for mentions of one specific pain reliever (i.e., Vicodin[®]) only. Consequently, new codes were created in 2015 for other hydrocodone products (e.g., Lortab[®], Norco[®]) to which code 434 had previously been assigned. In turn, the data dictionary for illicit drugs was updated for 2015 to reflect the changes to the codes for prescription drugs.

In all of the "OTHER, Specify" drug questions, respondents could enter up to five responses (in five data entry fields) for other drugs. For the most part, respondents specified a total of only one drug in these questions or they specified only one drug in each field. Neither of these situations required any special handling of the data.

¹⁹ See Section 1.1 and [Tables 1.1](#) and [1.2](#) in Chapter 1 for details about the contents of the self-administered sections of the 2015 NSDUH questionnaire.

However, additional procedures were required to handle the following situations in the "OTHER, Specify" drug data.

1. Respondents specified more than one drug within a single data entry field but they specified a total of five or fewer drugs across the five available fields.
2. Respondents specified a total of more than five substances across the five available data entry fields (i.e., by definition, they specified multiple drugs within one or more fields).

In the first situation, codes for the additional drugs were moved to the next unused "slot(s)" (i.e., variable[s]). For example, if a total of three drugs were specified, with two of them being specified in the first field and the third being specified in the second field, the additional code from the first field was moved to the third "OTHER, Specify" variable. Consequently, the responses in the individual "OTHER, Specify" drug variables did not always correspond to the order in which respondents reported use of these drugs.

When respondents specified more than five substances in the available fields, duplicate mentions of drugs were identified and removed. Duplicate mentions could include the same drug being mentioned more than once in the "OTHER, Specify" data or a drug being reported in the gate questions for a given section and also being reported in the "OTHER, Specify" data. For example, if question LS01a for LSD was answered as "yes" and LSD also was reported in the "OTHER, Specify" data for hallucinogens, then the duplicate code for LSD was not kept in the "OTHER, Specify" data for hallucinogens to allow codes to be used for other drugs. Similarly, if a respondent answered question TX11 as "yes" for perceiving the need for treatment for alcohol use in the past 12 months and specified alcohol and more than four other substances in the data for the perceived need for treatment, then the duplicate code for alcohol was not kept in the corresponding "OTHER, Specify" data. If more than five unique mentions of drugs remained after any redundant mentions had been eliminated, further priority in retaining responses in the final drug codes was given to (1) mentions of illegal or prescription drugs, as opposed to "over-the-counter" (OTC) drugs that are legally available without a prescription and (2) mentions that were relevant to the category of interest (e.g., in the hallucinogens section, giving priority to mentions of hallucinogens over mentions of drugs that were not classified as hallucinogens). Because the "OTHER, Specify" questions in the special drugs section for injection of other drugs applied to substances in any category, this second criterion was not used in deciding which codes to keep for special drugs.

In rare instances, the procedures described previously still did not yield five or fewer unique drugs in the "OTHER, Specify" data for a given section. In these situations, the highest priority was given to retaining codes for drugs that had no approved medical use in the United States or were deemed to have a greater potential for dependence or abuse, and the lowest priority was given to retaining codes for drugs that were not classified as controlled substances.²⁰ In the extremely rare situations where five or more unique drugs remained after all of these measures had been applied, the first five remaining "OTHER, Specify" codes were retained and

²⁰ The drug scheduling classifications, Schedules I through V, under the Controlled Substances Act were used in making these determinations. See <https://www.deadiversion.usdoj.gov/21cfr/21usc/index.html> and <https://www.deadiversion.usdoj.gov/schedules/index.html> for details.

mentions of any additional drugs were dropped from the coded data that were available for further editing or analysis.

2.3.1.2 Other Tobacco Brands

The CAI instrument included questions to identify the specific brands of tobacco that were used most commonly by respondents who reported use in the past 30 days of cigarettes or cigars. For both of these types of tobacco products, respondents who reported use in the past 30 days could choose from a list of brands shown on the computer screen²¹ or they could indicate use of "a brand not on this list." Respondents who gave the latter answer were asked to type in the name of the specific other brand that they used.

The basic coding scheme for the "OTHER, Specify" tobacco brand categories in 2015 was as follows.

- Codes of 101-199 and 1001-1999 were reserved for cigarette brands.
- Codes of 401-499 and 4001-4999 were reserved for cigar brands.
- Codes of 501-599 and 5001-5999 were reserved for pipe tobacco brands.
- Codes in the 601-699, 701-799, and 801-899 series were reserved for miscellaneous tobacco and nontobacco responses.

Codes were assigned to the "OTHER, Specify" tobacco brand data according to these categories, regardless of whether the response came from the section for cigarettes or cigars.

In 2015, the questions for snuff and chewing tobacco were replaced with questions about the use of any smokeless tobacco product, and questions were not included for the use of specific smokeless tobacco brands in the past 30 days. Nevertheless, the following tobacco brand codes continued to be available for 2015:

- Codes of 201-299 and 2001-2999 were reserved for chewing tobacco brands.
- Codes of 301-399 and 3001-3999 were reserved for snuff brands.

Although rare (i.e., occurring for fewer than five respondents in 2015 and for none in 2016), situations sometimes occur in which a smokeless tobacco brand was reported for the brand of cigarettes that respondents smoked most often in the past 30 days.

This coding scheme was particularly relevant for smokers of cigarillos or little cigars, which have a similar appearance to cigarettes but are considered to be types of cigars. Thus, if a respondent specified that the brand of other cigarettes that he or she smoked most often in the past 30 days was actually a brand of little cigars, the "OTHER, Specify" response for cigarette brands was assigned a code in the 400 or 4000 series for cigars.

Note that the coding for a particular tobacco brand did not capture further details, because the main aim in the coding was to capture information about any use within a particular brand

²¹ For cigarettes, the listing of brands was split between two different computer screens.

label. For example, the code for a particular brand did not capture details such as cigarette length, the size or shape of cigars, or flavor varieties.

2.3.1.3 Other Race or Ethnicity

In the interviewer-administered demographics section (see Chapter 4), NSDUH respondents were asked about their Hispanic/Latino origin and race information in QD04 and QD05, respectively. If respondents reported in QD05 that they were Asian, they were asked in QD05ASIA to report which Asian group best described them. Respondents who considered that the categories presented to them for Hispanic/Latino origin, race, or Asian ethnicity did not apply to them reported their "other" Hispanicity, race, or Asian ethnicity to the interviewers, who then typed in the respondents' answers.

The computer-assisted coding procedures and use of data dictionaries that were described in Section 2.3.1 applied to these "OTHER, Specify" data for race and ethnicity as well. In most cases, new responses were just new misspellings of an already established category, such as a response of "Porto Rican" instead of "Puerto Rican."

Regardless of the source (QD04, QD05, QD05ASIA) for these "OTHER, Specify" data, the write-in responses were used in subsequent editing of Hispanicity and race to determine the respondents' final Hispanicity and race (see Sections 4.2.6, 4.3.1, 4.3.2, and 4.3.3). Thus, in coding the "OTHER, Specify" data, each write-in was assigned two codes, one for race and the other for Hispanicity. If an interviewer entered both a geographical entity and a group within a particular race in the "OTHER, Specify" response, such as "Black Cape Verdean," the geographical entity was ignored in the race code and the respondent was coded as "Black/African American" for the race code. The geographic information Cape Verde was captured in the Hispanic/Latino code for "Cape Verde."

2.3.1.4 Additional Self-Administered "OTHER, Specify" Data

There were three types of "OTHER, Specify" questions in the self-administered sections:

- those where the "OTHER, Specify" item was a follow-up to a lead question that typically was answered as "yes" or "no";²²
- those where the "OTHER, Specify" item was a follow-up to a response category for "other" in an "enter all that apply" question; and
- those where respondents did not get the opportunity to choose the "other" response (and specify something) if they already chose another category from the list.

Coding of these "OTHER, Specify" variables was performed according to the general principles described in Section 2.3.1. Otherwise, minimal additional decision making was involved in assigning codes to the responses for these variables. Descriptions of these variables are included here for completeness.

²² Depending on the nature of the lead question, either an affirmative or a negative response to the lead question could govern whether respondents were asked to specify something.

An example of the first type of "OTHER, Specify" question is question TX42JSP in the substance treatment section (i.e., specify other source that paid/will pay for the last or current substance treatment). Only those respondents who reported in question TX42J that some other source paid for their last substance treatment or counseling or that some other source would pay for their current treatment (TX42J = 1) were routed to TX42JSP and asked to report the other payment source. Respondents could report specific sources of payment for their treatment in questions TX42A to TX42I and also report some other payment source in TX42J.

Question ADMT15 in the adult mental health service utilization section (i.e., specify the other location where outpatient mental health treatment was received in the past 12 months) is an example of the second type of "OTHER, Specify" question. Adult respondents aged 18 or older could report in question ADMT14 that they received mental health treatment or counseling in the past 12 months in up to six different types of outpatient settings, including "some other place." Adults were routed to ADMT15 if they reported receiving outpatient treatment in some other place in ADMT14. Because ADMT14 was an "enter all that apply" question, respondents could choose any of the specific locations that were listed in the question (e.g., an outpatient mental health clinic or center) and also report receipt of mental health treatment in some other place.

The third type of "OTHER, Specify" question is represented by the following three questions:

- SD21SP, which was associated with question SD21 (how respondents obtained their last needle for injecting drugs) in the special drugs section;
- TX25SP, which was associated with question TX25 (main place where the respondent received or was currently receiving substance use treatment in the past 12 months) in the substance treatment section;²³ and
- TX38SP, which was associated with question TX38 (outcome of the last substance use treatment).

Respondents first were presented with a list of options in the "lead" question (i.e., SD21, TX25, or TX38), including an option for "other" (e.g., other way in SD21 in which respondents got the last needle they used to inject drugs). If respondents chose any response from the list of options in the lead question except for "other," they were *not* routed to the "OTHER, Specify" question. Rather, data from the lead question and the specify question were combined into a single final variable.²⁴ When respondents chose the other category in the lead question, but they specified something that was coded as a missing value (i.e., don't know, refused, bad data, or blank), then the final edited variable (e.g., GNNDGET) retained a code corresponding to other, as opposed to being assigned a missing value.

²³ Although question TX04ISP (specify other location where the respondent received treatment for alcohol or other drug use in the past 12 months) was preceded by a "yes/no" question, TX04ISP used the same codes as TX25SP to allow the same data dictionary to be used for processing both of these "OTHER, Specify" variables (see Section 2.3.1).

²⁴ The final edited variables in 2016 were GNNDGET (based on data from SD21 and SD21SP) and TXLTVMNPL (based on data from TX25 and TX25SP).

The edits applied to GNNDGET, TXLTYNPL, and TXLTYOOCOM (including similar types of questions that could be added in the future) were designed to provide analysts with a *standardized* way to readily identify when it could be logically inferred that respondents should have chosen a given response option from the preceding question (i.e., rather than "other"). For GNNDGET as an example, codes of 1 to 4 applied to answers that respondents gave directly from question SD21 (e.g., 1 = Bought the needle from a pharmacy). Category 5 in SD21 was "Got the needle some other way." Although the coding sequence for "OTHER, Specify" responses in question SD21SP could have resumed at number 6 for responses corresponding to "Bought the needle from a pharmacy," assigning a code of 6 for "OTHER, Specify" responses that corresponded to category 1 in question SD21 would not enable analysts to readily see the logical connection between the "OTHER, Specify" response and the available response choice in the question that preceded it.

Responses in the "OTHER, Specify" variables for GNNDGET, TXLTYNPL, and TXLTYOOCOM were coded in a manner consistent with procedures from prior years. Specifically, responses that corresponded to existing response categories in their respective lead questions were coded starting with the number 21, with the coding proceeding in the order of the existing response categories. This procedure allowed for the possibility that lead questions could have 10 or more available choices and prevented overlap of "OTHER, Specify" codes with available responses. This procedure also offered analysts a standardized way to interpret these values. For example, if analysts wanted to treat codes of 21 to 24 in GNNDGET that came from "OTHER, Specify" data as being equivalent in an analysis to the corresponding codes of 1 to 4, the analyst could simply subtract 20 from any codes with values of 21 to 24 to recode these values to the corresponding codes of 1 to 4.

For similar reasons, the coding sequence for responses in these "OTHER, Specify" variables that did not correspond to responses from the list in the lead question resumed at number 41 (e.g., for GNNDGET, 41 = Given by/stolen from friend/acquaintance of friend/nonrelative). Although codes for responses that did not correspond to available choices from a lead question could have resumed at a lower number—especially for SD21SP—this approach standardized the assignment of codes across these types of "OTHER, Specify" variables (i.e., new codes always started at 41), minimized the risk of overlap between codes for these two types of responses, and allowed flexibility if new response choices were added to a lead question in a future year.

2.3.2 Investigation of Response Patterns in Usable Records

Although conversion to CAI reduced or eliminated some data quality problems that could occur in a PAPI format, it also was recognized that the audio computer-assisted self-interviewing (ACASI) environment could encourage some respondents to use the computer keyboard to enter nonsensical patterns of answers if they were not paying attention to questions or were not taking the interview seriously for other reasons. Thus, even if a respondent had sufficient data to meet the usable case criteria described in Section 2.2.3, certain patterns of answers could call into question the validity of the respondent's data, either for the entire interview or for specific sections of the interview.

2.3.2.1 Investigation of Response Patterns in Initial Substance Use Sections

A data diagnostics program was developed for substance use questions other than prescription drugs to screen for the following patterns of responses that might raise questions about the validity of the interview as a whole:

- high numbers of "yes" responses to lifetime use of specific hallucinogens or inhalants, which might indicate that respondents were indiscriminately keying data without paying attention to what they were entering;
- alternating "yes/no" responses to questions about lifetime use of specific hallucinogens or inhalants, which might indicate some type of pattern-making;
- high numbers of illegal drugs that respondents reported using every day or just about every day in the past year or past month (where applicable), in which case one might question either the validity of the answers or the respondent's competence to complete the interview;
- high numbers of substances that respondents reported first using at age 1 or 2, which might indicate indiscriminate keying of 1s or 2s, especially given that the age-at-first-use questions followed gate questions where a response of 1 denoted "yes" and a response of 2 denoted "no"; and
- consistent keying of the same code (either 1 or 2) throughout one or more sections (i.e., after lifetime use has been reported), which would suggest a pattern of indiscriminate answering.²⁵

These patterns of responses to substance use questions other than prescription drugs were examined on a case-by-case basis to determine whether a case should be retained as a final respondent or dropped.

Because of the changes to the prescription drug questions for 2015, a separate data diagnostics program was developed to screen for the following patterned responses in the prescription drug data:

- high numbers of prescription drugs that were reported in the screener sections for any use in the past 12 months or in the main sections for misuse in the past 12 months;
- high numbers of responses of 1 or 2 in the screener sections, which might indicate that respondents were confusing responses of 1 to mean "yes" and responses of 2 to mean "no," based on the meanings of these responses in earlier sections (e.g., a response of 2 in question ME01 for methamphetamine meant "no");
- reported use in the past 12 months of all prescription drugs in a list for screener items that included three or more prescription drugs;

²⁵ An important change to the CAI instrument since 2001 is that response categories for certain consistency checks use 4 = yes/6 = no instead of the format of 1 = yes/2 = no that was used prior to 2001. For example, if a respondent reported first using marijuana at age 1 or 2, the respondent could not use a response of "1" or "2" in question MJCC05 to verify that this age at first use was correct. This change to the CAI instrument was designed to stop respondents if they had been engaged in a pattern of keying responses of "1" or "2."

- miscellaneous patterns of keying in the screener sections (e.g., keying only a response of 1 in PR01, keying only 2 in PR02, keying only 3 in PR03, keying only 4 in PR04);
- keying only of responses of 1 for questions in the main prescription drug sections (e.g., pain relievers), after respondents had reported misuse in the past 12 months of at least one prescription drug in that category;
- high numbers of prescription drugs for which respondents reported first misuse at age 1 or 2, which might indicate indiscriminate keying of 1s or 2s;
- consistent keying of the same age at first misuse (AFM) for prescription drugs that respondents misused in the past 12 months;
- reports of misuse on every day or just about every day in the past month; and
- all possible responses chosen for the ways in which respondents misused prescription drugs (e.g., question PRYWAYS for pain relievers) or reasons for misusing prescription drugs (e.g., question PRYMOTIV for pain relievers) in the past 12 months.

Fewer than 10 cases in 2016 met the usable case criteria but were treated as nonrespondents because their responses were of questionable validity, based on one or more of the patterns described above.

In addition, fewer than 10 cases in 2016 were retained as respondents, but their original responses to questions in one or more of the sections for alcohol through methamphetamine were replaced with bad data codes. This process included setting responses to bad data that indicated they were lifetime users of a given drug. For example, data for some respondents were set to bad data because the respondents keyed values of 1 or 2 to every question that was asked in a section, beginning with the age-at-first-use question. Data for the lifetime variables for these cases were set to bad data as part of the edits for the lifetime drug use variables (Section 6.2.1). For example, if a case was identified that had "bad" data for inhalants, the data for the lifetime use of inhalants corresponding to responses in questions IN01A through IN01L were set to bad data as part of the lifetime drug use edits, and a flag was set to indicate that data subsequently needed to be set to bad data for related variables in the inhalants section of the interview and for variables in subsequent sections that were contingent on respondents being users of inhalants.

The changes to the prescription drug questions for 2015 resulted in a greater number of respondents providing patterned responses to prescription drug data in 2015 than in 2014. In particular, when respondents answered gate questions in previous sections of the interview, they were instructed to key responses of 1 to mean "yes" and to key responses of 2 to mean "no." When respondents reached the prescription drug screeners, however, responses of 1 or 2 corresponded to specific prescription drugs. In question PR01 in the screener for pain relievers, for example, respondents were instructed to key 1 if they used Vicodin® in the past 12 months and to key 2 if they used Lortab® in that period. Respondents were instructed to key an answer of 95 if they never used any of the pain relievers in the list for PR01.

Two changes were made to the 2016 questionnaire because of large numbers of respondents in 2015 (i.e., more than 100) who keyed responses of only 1 or only 2 wherever possible in one or more of the prescription drug screeners. First, a question was added to the tutorial section (i.e., practice questions) before the ACASI section of the interview to show

respondents how to key a response of "95" if none of the items in a list applied to him or her. Second, consistency checks were added to the prescription drug screener sections to ask respondents to verify their answers if they entered three responses in a row of only 1 or three responses in a row of only 2 (e.g., if respondents chose only responses of 2 in questions PR01, PR02, and PR03).

These changes beginning in 2016 considerably reduced, but did not completely eliminate, the occurrence of patterned responses in the prescription drug data. The prescription drug variables were set to bad data if definite patterned responses still occurred in the data. As an example of improvements because of these changes in 2016, the edited variable LTABANYR (any past year use of Lortab[®], corresponding to the second response choice in question PR01) was set to bad data for fewer than 10 respondents in 2016 compared with more than 60 in 2015.

2.3.2.2 Editing of Other Self-Administered Data Because of Patterned Responses

In the self-administered sections of the interview for special drugs through consumption of alcohol, two types of edits were made because of patterned responses in the interview.²⁶ The first type involved edits because of patterned responses in related variables from corresponding substance use sections (see Section 2.3.2.1). When a case was retained as a final respondent but with variables from one or more initial substance use sections being assigned the appropriate codes for bad data (Section 2.4.3), codes for bad data also were assigned to the corresponding variables in the subsequent sections. This edit was implemented regardless of whether the corresponding variables in these subsequent sections had patterned responses. For example, if a respondent's variables in the inhalants section were assigned codes for bad data because of patterned responses in that section and the respondent was routed to questions pertaining to inhalants in the substance dependence and abuse section, then codes for bad data were assigned to the edited variables for inhalant dependence or abuse.

The second type of edit applied to other occurrences of patterned responses specifically in the self-administered data. These edits were implemented for fewer than 30 cases in 2016. Responses were replaced with bad data codes when these patterns occurred. In particular, answers were replaced with bad data codes if respondents keyed answers of "1" (or multiple-digit responses of "1," where applicable) to every question they were asked within a section. In the youth experiences section, adolescents who keyed answers of "2" wherever possible also had their answers replaced with bad data codes. For respondents to key all possible items as "2" in the youth experiences section would mean that they keyed this response to 42 items in a row, including that they had engaged in all "problem" activities (such as getting into fights, carrying a handgun, or selling illegal drugs) exactly "1 or 2 times" in the past year. These respondents who

²⁶ Edits for patterned responses were implemented in the self-administered demographic questions beginning in 2016. However, these new edits failed to identify patterned responses in the self-administered demographic data for about 15 additional respondents in 2016 because marital status variables (which were moved back to computer-assisted personal interviewing in 2016 but were in the self-administered section in 2015) were included in the checks for patterned responses. Consequently, only respondents aged 12 to 14 (who were not asked the marital status questions) were flagged for patterned responses in the self-administered demographic data. Further investigation indicated that not editing the data for the additional respondents who were identified instead of setting the responses to "bad data" would have minimal effects on estimates. Corrected checks for identifying patterned responses in this section of the interview have been implemented for 2017.

keyed "2" wherever possible also would have participated in exactly two school-based activities, two community-based activities, two church- or faith-based activities, and two other kinds of activities in the past year.

2.3.3 Edits to "Date-Dependent" Variables

Beginning in 2015, respondents could view electronic calendars on the laptop computer screens that showed the start dates for reference periods pertaining to the "past 30 days" and "past 12 months." These electronic calendars replaced the paper calendars that interviewers filled out with respondents at the beginning of the interview to mark the start dates for these reference periods.

With the introduction of the electronic calendars in 2015, the CAI program used the interview date information that was stored by the computer to indicate the start dates in the calendars for these reference periods. Specifically, the starting date for the past 30-day period was calculated and displayed to respondents in the calendar as exactly 30 days prior to the stored interview date.²⁷ Similarly, the starting date for the 12-month reference period was displayed to respondents in the calendar as the same month and day as the interview date but in 2014. The CAI program also used the interview date information to create "date fills" during the interview that indicated the starting dates for questions pertaining to the "past 30 days" and "past 12 months." For example, in the recency-of-use questions that asked respondents when they last used the drug of interest, the response category "within the past 30 days" included a date fill to remind respondents when the past 30-day reference period began for them. Similarly, introductions to specific questions about the frequency of use of a particular drug in the past 12 months and past 30 days included date fills to remind respondents of the period they should be thinking about when answering these questions.

Data that could be affected by questionable interview dates were edited as needed once the edited interview date variable INTDATE had been created for all final respondents (Section 4.2.1). As part of the procedures for creating INTDATE, an indicator variable (EIIDATE) was created that specified how the final interview date was assigned. For example, EIIDATE could indicate that the interview date that had been stored by the CAI program during the interview had an incorrect year or was outside of the quarter in which the interview had been completed.

The occurrence of problematic interview dates is rare. With the exception of 2008, no cases from 2004 to 2015 had an originally entered interview date that was sufficiently problematic to call into question the respondent's answers to date-dependent questions in self-administered sections of the interview. In 2016, however, a small number of respondents (fewer

²⁷ If a respondent broke off and then resumed the interview at a later date, the calendars and "fills" in the interview that defined the past 30-day and past 12-month reference periods were updated for remaining questions when the interview resumed. This was done because 30 days and 12 months from the date when an interview resumed might be later than the 30-day and 12-month periods when a respondent had previously been asked questions. Thus, these reference periods during the interview were dynamic rather than static. If a respondent resumed the interview and went back to a section that he or she had completed prior to the breakoff, the date displayed in the calendar and the date "fill" would be reset based on the new date when the interview resumed. In practice, however, the tendency is to proceed with the parts of the interview that have not been completed rather than to go back to earlier sections.

than 10) had questionable interview dates. Data in the self-administered sections that were dependent on the interview date were therefore considered problematic. For example, if the CAI program calculated a 30-day reference period based on an incorrect interview date, answers that the respondent gave on the number of days that he or she used different drugs in the past 30 days could reflect use of these drugs in a period other than the intended 30-day reference period.

This interview date issue did not present a problem if respondents had never engaged in a particular behavior (e.g., never used marijuana). Moreover, if the respondent reported never having engaged in a particular behavior, the CAI program skipped that person out of questions where the interview date would be important for establishing reference periods. For these reasons, some cases where there was some question about the interview date could still be retained as final respondents.

Instead of cases being dropped, the edits that were in place for 2016 set the following types of self-administered questions to bad data if respondents were routed to them:

- questions pertaining to a specific time period that was less recent than the respondent's lifetime (e.g., the past week, past 30 days, past 12 months);
- questions pertaining to the most recent time that an event occurred (e.g., when a respondent last used a drug of interest); and
- questions pertaining to the respondent's age when an event occurred (e.g., the age when the respondent first used a drug of interest).

Self-administered questions about age at first use and ages when other events occurred were not related directly to the interview date but were related indirectly via respondents' ages; that is, respondents' ages were calculated by comparing the date of birth with the interview date. In turn, age-at-first-use and other age data in the self-administered sections were compared for consistency with the respondents' ages.

For the types of questions that were described previously, respondents' answers were set to bad data before any further editing was done. Setting the responses to bad data prior to any subsequent editing allowed analysts during the editing stage of data processing to distinguish between situations where the data were deemed to be bad prior to any further editing being done and situations where a variable might be set to bad data in subsequent editing steps because of inconsistencies with other data in the respondent's record (Section 2.4.3).

This edit did not apply to gate questions that asked whether a respondent had ever engaged in a behavior (e.g., "Have you ever smoked part or all of a cigarette?"). As noted previously, whether the respondent had ever engaged in a particular behavior prior to being interviewed is not dependent on the value stored for the interview date. This edit also did not apply to questions within a section that asked whether a more detailed behavior of interest was *ever* true for this respondent (e.g., "Have you smoked at least 100 cigarettes in your entire life?"). The decision also was made that this edit would not apply to sections of the interview that were administered directly by the field interviewers (FIs).

2.4 General Issues and Principles for Editing NSDUH Data

Previous sections discussed broader data processing and coding procedures in the data as a whole or in large sections, such as the initial substance use sections of the interview. In contrast, this section focuses on the general principles of data processing and editing across and within interview sections, once the initial processing of transmitted interview data had been completed (i.e., the procedures described in Section 2.2) and after cases with patterned responses or questionable interview data information had been identified (Sections 2.3.2 and 2.3.3).

The following specific issues are addressed in this section:

- editing across sections (Section 2.4.1);
- assignment of relevant "not applicable" codes (Section 2.4.2);
- additional assignment of NSDUH missing data codes (Section 2.4.3);
- assignment of codes to edited "enter all that apply" variables (Section 2.4.4);
- "back-editing" based on "OTHER, Specify" data (Section 2.4.5);
- the "lockout" procedure that prevented respondents and interviewers from returning to self-administered sections of the interview after respondents had completed their last self-administered section (Section 2.4.6); and
- general principles for the "flag and impute" procedures for substance use variables other than for prescription drugs (Section 2.4.7).

2.4.1 Editing across Sections

An important principle that was followed in editing the data was that responses from one section (e.g., hallucinogens) generally were not used to edit variables in another section (e.g., inhalants). In particular, the self-administered data following the initial substance use sections (special drugs through employment) were *not* used to edit related variables in the initial substance use sections (tobacco through sedatives). For example, if a respondent reported in the heroin section that he or she last used heroin more than 12 months ago, but then reported in the later special drugs section that he or she last used heroin with a needle more than 30 days ago but within the past 12 months, then the heroin recency variable HERREC from the heroin section was *not* edited to take into account this more recent indication of heroin use from the later section. Rather, HERREC retained the respondent's initial response that he or she last used heroin more than 12 months ago. Consequently, the documentation in NSDUH codebooks for the variables in sections of the interview that follow the initial substance use sections includes a footnote to indicate that these variables may be inconsistent with data in other sections of the interview.

This principle of not using substance use data in later interview sections to edit data in the initial substance use sections has been important for maintaining consistent data to assess trends in substance use (i.e., not including situations in which new baselines have been established because of questionnaire changes in 2015). If variables in the initial substance use sections were allowed to

be edited based on respondents' answers in the later sections, then key substance use estimates could change across years as later questions or entire interview sections were added or deleted.

One exception to this principle involved situations in which responses in one or more sections governed whether respondents were asked questions in a later section. For example, the substance use treatment section was relevant only for respondents who reported some lifetime use of alcohol or other drugs, excluding tobacco products. Respondents who reported in the initial substance use sections that they had never used alcohol, marijuana, cocaine, heroin, hallucinogens, inhalants, methamphetamine, or misused prescription psychotherapeutics (i.e., pain relievers, tranquilizers, stimulants, or sedatives) were not asked corresponding follow-up questions in the substance use treatment section. In such cases, during the editing process, blank values in the substance use treatment variables were replaced with codes to indicate that respondents were not asked the follow-up questions in this section because they reported never having used or misused any of the relevant substances.

2.4.2 Assignment of Relevant "Not Applicable" Codes

Because the CAI logic controlled whether respondents were skipped out of some questions based on their answers to previous questions, an important aspect of editing the NSDUH data involved replacing missing data codes in the unedited data with appropriate codes to indicate that the questions had been skipped because they did not apply. The following codes were assigned when respondents were skipped out of a given question *and it could be determined unambiguously* that the respondent had legitimately been skipped out of the question, based on the answer(s) to one or more previous questions:

- 91 (or 991, or 9991, etc.) = NEVER USED [DRUG(S) OF INTEREST];
- 93 (or 993, or 9993, etc.) = USED [DRUG] BUT NOT IN THE PERIOD OF INTEREST; and
- 99 (or 999, or 9999, etc.) = LEGITIMATE SKIP.

Strictly speaking, codes of 91 and 93 in the CAI data could be considered variants of the more generic legitimate skip code. Their use was designed to provide analysts with more information about the reason that respondents were skipped out of a particular question.

Codes of 91 and 93 were used most often in the initial substance use sections of the interview. For example, a code of 91 (or 991, etc.) in the marijuana section denotes the pattern where respondents were skipped out of all remaining marijuana questions because they answered "no" to the lifetime marijuana question MJ01. Similarly, a code of 93 in the marijuana section denotes situations where respondents were lifetime marijuana users but were definitely not users in the past 30 days or past 12 months or both.

Codes of 91 and 93 also were used to a limited extent in sections of the interview following the initial substance use sections because the CAI logic took into account respondents' prior answers to the initial substance use questions to determine whether particular questions applied in later sections of the interview. For example, questions about cocaine in the substance dependence and abuse section were relevant to respondents who used cocaine in the 12 months

prior to the interview. Thus, if a respondent last used cocaine more than 12 months prior to the interview, a code of 93 in the substance dependence and abuse variables pertaining to cocaine would signify to an analyst why the CAI program skipped the respondent out of these questions. Similarly, a code of 91 in the substance dependence and abuse variables for cocaine signified to an analyst that the respondent was skipped out of these questions because he or she had never used cocaine.

A legitimate skip code of 99 was used most often in sections of the interview following the initial substance use sections. For example, the youth experiences section was intended to be administered only to respondents aged 12 to 17. If a respondent was 18 or older, a code of 99 was assigned in the editing process to the skipped youth experiences variables. Similarly, if a respondent had used alcohol or some other drug at least once in his or her lifetime, but answered the lifetime substance treatment question TX01 as "no," then the CAI program skipped the respondent out of all remaining substance treatment questions. A code of 99 was assigned to the skipped substance treatment variables in this situation to signify that the respondent had used alcohol or drugs at least once but had never received substance abuse treatment.

The following codes also were assigned through editing:

- 81 (or 981, or 9981, etc.) = NEVER USED [DRUG(S)] Logically assigned;
- 83 (or 983, or 9983, etc.) = USED [DRUG(S)] BUT NOT IN THE PERIOD OF INTEREST Logically assigned; and
- 89 (or 989, or 9989, etc.) = LEGITIMATE SKIP Logically assigned.

These codes were given values in the 80s to signify that existing values were to be overwritten during the editing process. For example, if a respondent was somehow routed into the youth experiences section, but that respondent was subsequently classified as being 18 or older (Section 4.2.2), then any answers that the respondent gave in the youth experiences section were overwritten with a code of 89 (or 989, etc.). These codes signify that this adult respondent logically was not eligible to be asked the youth experiences questions. Chapter 6 presents examples of edits where codes of 81 or 83 were assigned.

However, the codes described in this section were assigned *only* in situations where there was total certainty that a respondent should have skipped a question. For example, if a respondent reported in question MJ01 that he or she had never used marijuana, it was absolutely clear that subsequent questions about marijuana use (e.g., age at first use of marijuana) did not apply. See Section 3.1 for a description of imputation indicators, including a brief discussion of the codes assigned to respondents whose imputed values meant that they legitimately skipped out of the corresponding questions.

The CAI skip logic often treated responses of "don't know" or "refused" to gate questions as equivalent to a negative response. For example, if a respondent was uncertain whether he or she had ever used marijuana (and answered question MJ01 as "don't know"), the CAI program skipped the respondent out of all remaining questions about marijuana use, as though the respondent had never used it. From the standpoint of respondent burden, this makes sense. There is little value in asking a respondent who did not know whether he or she had ever used

marijuana, "How old were you the first time you used marijuana or hashish?" Implicit in this question is that respondents have used marijuana at least once in their lives.

Although the CAI program skipped respondents out of questions if they answered a gate question as "don't know" or "refused" (or gave similar answers on follow-up if they initially refused to answer a gate question), these types of responses to a gate question are ambiguous and do not provide an analyst with conclusive information one way or the other about the behavior or event of interest. Consequently, such responses could be thought of as *potentially* affirmative responses, as opposed to inferring that they are negative responses. In particular, as noted previously, respondents who initially refused to answer a question about their lifetime use or nonuse of a drug were given a second opportunity to answer the question as "yes" or "no." Similarly, if a respondent who initially did not know whether he or she had ever used a drug had thought about the issue further, the respondent may have recalled a time when he or she in fact had used it—and more detailed questions about use of the drug would have been relevant for this respondent. Alternatively, if the respondent gave more thought to the issue and decided that he or she really should answer the lifetime drug use question as "no," an analyst would have a solid basis for determining that subsequent questions did not apply.

Further, the procedures for statistically imputing missing data did not automatically infer lifetime nonuse when respondents provided ambiguous information about whether they had ever used a given drug. Rather, such respondents were eligible to be statistically imputed to be lifetime users or nonusers. For these reasons, variables retained missing values in the editing procedures when questions were skipped due to respondents answering a lead question as "don't know" or "refused" (or answering in a similar manner in response to a follow-up probe).

2.4.3 Additional Assignment of NSDUH Missing Data Codes as Part of Editing within a Section

Previous sections described the procedures for assigning missing data codes in the interview data as a whole or in large sections, such as in self-administered sections, when the interview date was questionable (Section 2.3.3). This section discusses additional principles for assigning (or retaining) missing data codes as part of the editing procedures for a given section.

In particular, if respondents refused a single or multiple gate question or questions that governed a skip pattern in a section, refusal codes were often assigned to all of the subsequently skipped items in the section as part of the editing procedures (i.e., the lead refusal was propagated); that is, it was logically inferred that the respondents were globally refusing to answer *any* questions on that topic.²⁸

This propagation did not occur when respondents answered a gate question or questions as "don't know." Rather, values of "blank" (no answer) were retained in the skipped questions. Unlike the situation for responses of "refused," it does not follow logically that a response of "don't know" to a gate question would imply that the respondent would answer "don't know" to all subsequent questions on that topic. For example, if a respondent answered the lifetime

²⁸ Because of the focus of the questions for specific prescription drugs on past year use or misuse rather than on lifetime misuse (i.e., and no gate questions for lifetime use or misuse), propagation of refusals was not done for the prescription drug data beginning in 2015.

marijuana question as "don't know," assigning a "don't know" code to the age-at-first-use variable (corresponding to question MJ02) would imply that the respondent was a lifetime user but did not recall the age when he or she first used.

In addition, data sometimes were identified that were inconsistent with other data in a respondent's record. For the sections of the interview for tobacco through methamphetamine, for example, if a respondent reported first using a drug at an age greater than his or her current age, the CAI program indicated to the respondent that this age at first use was inconsistent. The respondent was prompted to revise the age at first use, his or her current age, or both, to make the data consistent.²⁹ As noted in Chapter 1, however, respondents did not always resolve these types of inconsistencies. If the age at first use was still inconsistent with the respondent's age, despite the opportunity that the respondent had to resolve the inconsistency, then a bad data code was assigned to the age-at-first-use variable to indicate that the data were inconsistent with other data. As was the case with prior survey rounds, the following codes were assigned to denote "bad data" (i.e., usually inconsistent with other data): 85 (or 985, or 9985, etc.) = BAD DATA Logically assigned. Other situations where bad data codes were assigned are discussed below in connection with specific steps in the editing process.

2.4.4 Assignment of Codes to Edited "Enter All That Apply" Variables

As noted in Section 2.2.2, the initial creation of separate variables for "enter all that apply" questions involved assignment of the relevant code that was shown to respondents during the interview for that question, or a code of 98 (blank) if respondents did not choose that item from the list; these procedures were implemented for "enter all that apply" questions anywhere in the entire interview. The additional edits described in this section were implemented within sections that had "enter all that apply" variables. Most variables corresponding to "enter all that apply" questions that existed prior to 2013 were coded as 1 if respondents chose that item from the list, and values of 98 were recoded with a value of 6. Documentation for these edited values for "enter all that apply" variables was as follows:

- 1 = Response entered, and
- 6 = Response not entered.

For example, adolescent respondents who reported the receipt of mental health services in the past year from a private therapist, psychologist, psychiatrist, social worker, or counselor in question YSU16 could choose reasons from questions YSU18 or YSU18A for receiving services from these providers. The code of 4 that was assigned in the initial processing to the variable from YSU18 for receiving services because of breaking rules or "acting out" was reassigned a code of 1 as part of the editing procedures. If the respondent did not choose this reason from the list in question YSU18, but reported another reason from the lists in YSU18 or YSU18A, then the code of 98 in the variable for breaking rules or "acting out" (which had been assigned during

²⁹ For prescription drugs beginning in 2015, respondents who reported misuse of a specific prescription drug at an age that was later than their current age could revise their age at first misuse but could not change their current age.

the remapping of "enter all that apply" variables, as described in Section 2.2.2) was replaced with a code of 6 during the editing process.

Use of the code of 6 was intended to indicate to analysts that not choosing an available response from the list was not exactly the same as an answer of "no" in questions that required respondents explicitly to answer "yes" or "no" about a behavior of interest. In other words, a response of "no" in a "yes/no" type of question can be thought of as an *active* indication that the behavior or characteristic of interest did not apply, whereas not choosing a response in an "enter all that apply" list can be thought of as a *passive* indication that a particular behavior did not apply. In practice, however, not choosing a response from a list often was treated as being *equivalent* to a response of "no" in subsequent editing steps.

However, this procedure was modified for "enter all that apply" questions that were added to the interview beginning in 2013 (including "enter all that apply" questions that were added for prescription drugs in 2015). These new variables were coded as 1 if respondents chose that item from the list, and values of 98 were recoded with a value of 2. Documentation for these values was as follows:

- 1 = Yes, and
- 2 = No (not entered).

For example, questions QD10A, QD10B, and QD10C were added in 2013 for respondents who had ever served on active duty in the United States Armed Forces or Reserve components. Respondents who reported in question QD10A that they had ever served on active duty were asked in question QD10B to indicate the period when they served on active duty, with question QD10B being an "enter all that apply" question. If a respondent reported serving on active duty between May 1975 and July 1990, the code of 3 that was assigned to the variable in the initial processing was reassigned a code of 1 as part of the editing procedures. If the respondent did not report being on active duty during this period but reported being on active duty during another period, then the code of 98 in the variable for this period of active duty military service was replaced with a code of 2 during the editing process. As for the code of 6 that was described previously, the documentation of "No (not entered)" for the code of 2 was intended to indicate to analysts that not choosing an available response from the list was not exactly the same as an answer of "no" in questions that required respondents explicitly to answer "yes" or "no."

As noted previously, assignment of codes of 1 or 2 also was implemented for the variables corresponding to the "enter all that apply" questions for any use of specific prescription drugs that were added in 2015. For example, if a respondent reported using generic hydrocodone in the past 12 months in question PR01, then the code of 5 that was assigned to the variable for hydrocodone use as part of the initial processing was reassigned a code of 1. If the respondent did not choose generic hydrocodone from the list of drugs in PR01 but reported using another hydrocodone product from the list, then the code of 98 for generic hydrocodone was replaced with a code of 2. Editing of the variables for any use of specific prescription drugs in the past 12 months when respondents chose the response of 95 (i.e., meaning that they had not used any of the drugs listed in a given question) are discussed in Section 6.3.

2.4.5 "Back-Editing" Based on "OTHER, Specify" Data

"Back-editing" refers to situations in which answers to a given question can be used to make inferences about how a preceding question *within the same section* should have been answered. Specifically, the principles and procedures that are discussed in this section refer to use of "OTHER, Specify" data to go "back" and edit an earlier variable according to what respondents specified in a later series of questions for that section.

In particular, a type of inconsistency that could commonly occur in the NSDUH data involved situations in which respondents did not answer a question affirmatively (e.g., in question LS01A, whether they ever used LSD). In the same section, however, they reported something in "OTHER, Specify" items that indicated that the other question should have been answered as "yes" (e.g., specifying use of LSD as some "other" hallucinogen that the respondent had ever used). When respondents specified something that corresponded to an item they had been asked about previously, but they had not answered that previous item as "yes," then the editing procedures assigned a value of "yes" to the relevant question. The following code typically was used when a response of "yes" was logically inferred: 3 = Yes LOGICALLY ASSIGNED.

If there was a lead to the "OTHER, Specify" question that was in the form of a "yes/no" question (e.g., "Have you ever, even once, used any other hallucinogens besides the ones that have been listed?"), the affirmative answer was retained in the lead to the "OTHER, Specify" question (i.e., having ever used any other hallucinogens), in addition to the inference being made that the answer to another question logically was "yes" (e.g., "Have you ever, even once, used LSD, also called 'acid'?"). The redundant specify code also was retained in the "OTHER, Specify" variable(s) to indicate to analysts the source of the logically inferred "yes" value.

This principle also applied to the editing of variables for the use and misuse of specific prescription drugs in the past 12 months based on answers in "OTHER, Specify" data for the misuse of other prescription drugs in that period. Editing of the prescription drug data based on the "OTHER, Specify" data varied according to whether respondents had reported any use of a given prescription drug in the past 12 months. For example, suppose a respondent reported using the pain reliever Norco[®] in the past 12 months. In this situation, the edited variable NRCOANYR had already been assigned a code of 1 according to the procedures described in Section 2.4.4. If this respondent did not report misusing Norco[®] in the past 12 months in question PRY03, but specified Norco[®] as some "other" pain reliever that the respondent misused in that period, then the edited variable for the past year misuse of Norco (NRCONMYR) was assigned a code of 3 to indicate that question PRY03 logically should have been answered as yes.

However, suppose a respondent did not choose Norco[®] from the list of pain relievers in question PR01. In this situation, question PRY03 would have been skipped. If this respondent specified Norco[®] as some other pain reliever that he or she misused in that period, then it could logically be inferred that the respondent misused Norco[®] in the past 12 months and also that the respondent used this prescription drug in the past 12 months for any reason. A code of 5 was assigned to the variables for any use and misuse (e.g., NRCOANYR and NRCONMYR for Norco[®]), where 5 = Yes LOGICALLY ASSIGNED. This code of 5 distinguished the

respondents who specified the misuse of a specific prescription drug in the past 12 months but did not report the use of that drug in the past 12 months from other respondents who reported use but not misuse in that period. The individual edited variable corresponding to lifetime nonmedical use of codeine was assigned a code of 3.

For "enter all that apply" variables other than the screener variables for prescription drugs, a code of 3 typically was used when a response was logically inferred based on "OTHER, Specify" data. If this variable existed prior to 2013, then this code of 3 had the following meaning: Response entered LOGICALLY ASSIGNED. For example, if an adult respondent who received mental health services as an outpatient in the past 12 months did not report the receipt of services at an outpatient mental health clinic or center in question ADMT14, but specified the receipt of services in this type of location as another place where he or she received outpatient mental health services, then it was inferred that this location logically should have been chosen from the list.

For "enter all that apply" variables that have been added to the survey since 2013, a code of 3 that indicated logical inference of a response from "OTHER, Specify" data had the following meaning: Yes LOGICALLY ASSIGNED. For example, respondents in 2016 who reported in question HLTH26 that they had some other type of cancer in their lifetime were asked to specify the other type of cancer that they had. Thus, for example, if a respondent did not choose "Melanoma" from the list in HLTH26, but specified that melanoma was the "other" type of cancer that he or she had, then a code of 3 was assigned to the variable CAMELANOM from HLTH26.

2.4.6 "Lockout" Following Completion of Self-Administered Sections

Since 2003, once the self-administered portion of the interview had been completed, interviewers and respondents are locked out of this section of the interview to provide additional protection of the privacy of respondents' answers. However, based on answers that respondents gave in the interviewer-administered questions following the self-administered section, interviewers could go back and change information in demographics questions that were in the interviewer-administered section that preceded the self-administered questions. For example, interviewers could discover when they were completing the household roster section of the interview that a female respondent's gender had been incorrectly keyed as male, or vice versa (Section 4.2.4). The interviewer was allowed to change the information about the respondent's gender at the beginning of the interview.

If interviewers went back and changed answers to the interviewer-administered demographic questions, however, the CAI program did not reroute respondents back into the self-administered section to answer questions that previously had been skipped. The CAI program also did not update the answers in the self-administered section to remove data from questions that logically should have been skipped based on any new information that the interviewers entered. For example, this "lockout" feature could result in some respondents being asked pregnancy questions because their gender was originally reported as female but their final gender was recorded as male. In this situation, the pregnancy data for respondents whose final gender was reported to be male were overwritten with codes of 89 (LEGITIMATE SKIP Logically assigned; see Section 2.4.2) to indicate that these data did not apply. Similarly, the

pregnancy questions would have been skipped for respondents aged 12 to 44 who were originally reported to be male. If the final gender for these respondents was recorded as female, then the edited variables would have missing data (i.e., blank) because female respondents aged 12 to 44 were expected to be asked these questions.

2.4.7 "Flag and Impute" Principles for Substance Use Data Other than Prescription Drugs

The editing procedures for establishing when respondents last used a substance are critical for creating final published estimates from NSDUH of the prevalence of substance use in the United States. In addition, data from the initial substance use sections on the most recent use are important for establishing whether skipped questions in later sections truly were not applicable or if there might be some question about whether these skipped questions might have applied to the respondent. For example, respondents who reported that they used cocaine in their lifetime but that they last used it more than 12 months ago were not asked questions about cocaine dependence or abuse. However, if any of these respondents also reported that they first used cocaine at their current age, these reports of recent initiation would suggest that they may have used cocaine in the past 12 months, in which case they should have answered the dependence or abuse questions for cocaine.

Under the deterministic edits for the old PAPI format, as a general rule, if a respondent indicated in one question on a substance's answer sheet that he or she had never used a substance and indicated use of that substance in another question on the answer sheet, logical editing coded the person as a user of that substance. If a respondent reported two (or more) different answers on the same answer sheet with respect to how recently he or she had used a substance, the editing procedures typically assigned the category indicating the more recent use. Relatively little statistical imputation was done to the PAPI recency variables following the editing step. A drawback of this approach was that decisions to infer more recent use could have an appreciable impact on estimates of use in the past 30 days for less commonly used substances, such as cocaine and heroin.

Conversion of the instrument to a CAI format in 1999 provided an excellent opportunity to reexamine the procedures and underlying assumptions for editing the recency variables. Further, the logic in the CAI instrument, in which respondents were skipped past questions that did not apply to them, precluded the same kinds of edits for the CAI recency variables as were done for PAPI.

Four possible ways of editing the data for most recent use of a substance were examined as part of the methodological research for data processing and estimation procedures using the CAI data (Kroutil & Myers, 2002). The flag and impute rule that was adopted for editing the CAI recency-of-use variables in 2015 for substances other than prescription drugs flags inconsistencies between a recency variable and related variables but does not make a decision about the final recency category. Rather, this rule leaves these inconsistencies for recency-of-use and related data to be resolved through statistical imputation.

For example, if a respondent in 2016 originally reported in the sections of the interview for tobacco through methamphetamine that he or she last used a substance more than 12 months

ago but also reported first using it at his or her current age, this procedure inferred that the respondent was *at least* a lifetime user. In the imputation procedures, this case's data for most recent use was imputed to indicate most recent use in any period (i.e., in the past 30 days, more than 30 days ago but within the past 12 months, or more than 12 months ago). Also, the data on past year initiation for substances other than prescription drugs that were inconsistent with most recent use of a given substance more than 12 months ago were set to bad data (Section 2.4.3), which treated the initiation data as missing. Consequently, if the respondent was imputed to have last used the substance at some point in the past 12 months (including use in the past 30 days), then the respondent could be imputed to have initiated use of the substance within the past 12 months or more than 12 months ago based on data from a donor whose reports of first use and most recent use were consistent. However, if the respondent was imputed to have last used the substance more than 12 months ago, then initiation data from the donor respondent also would be consistent with initiation more than 12 months prior to the interview date.

The beauty of this edit rule lies in its simplicity: If a respondent gives an answer within a substance's section that conflicts with the original answer to the recency-of-use question, then the recency variable is statistically imputed using data from a suitable donor record without these inconsistent data (see Chapter 3). A second attractive feature of this rule is that if the respondent provides conflicting information, it is not necessary to try to deduce from the data when the respondent last used the substance. Moreover, this rule does not automatically discount indications of more recent use than the respondent originally reported in a recency question, nor does it automatically infer that the respondent last used a substance more recently than he or she originally reported (as was the case with editing procedures for the former PAPI data). However, if final assignment occurred to indicate use in a more recent period than the respondent originally reported, this decision typically was made through statistical imputation rather than deterministic editing.³⁰

The changes to the prescription drug questions beginning in 2015 (see Section 1.4.1) greatly reduced the opportunity for respondents to give inconsistent answers to these questions. For example, respondents were asked questions about their first misuse of prescription drugs only for the specific prescription drugs that they misused in the past 12 months. Consequently, beginning in 2015 respondents could not report that they initiated the misuse of prescription drugs in a given psychotherapeutic category (e.g., pain relievers) without also having reported misuse of some prescription drugs in that period. Because the changes to the prescription drug questions reduced the opportunity for respondents to answer the questions inconsistently, these "flag and impute" procedures did not need to be applied to the editing of the prescription drug data in 2016.

³⁰ Limited exceptions that involved deterministic editing of recency variables are discussed in Chapter 6.

3. Imputation Procedures Used in the 2016 NSDUH

3.1 Introduction

As with most large-scale sample surveys, the respondent datasets for the 2016 National Survey on Drug Use and Health (NSDUH) contained missing responses for some items, inconsistent or invalid responses, and violations of skip patterns. Although the survey instrument was designed to enforce skip patterns and to perform some consistency checks as data were collected, invalid and inconsistent responses still occur. These response errors are a source of bias in the analysis of NSDUH data (Cox & Cohen, 1985).

Deterministic editing to correct erroneous and inconsistent responses and to replace missing values is appropriate when a unique association exists between predictor variables and the variable to be predicted (Cox & Cohen, 1985). For instance, gender often can be inferred from the respondent's relationship to the head of a household (e.g., son, daughter). However, even when good predictor variables are present, an unambiguous prediction may not be possible for every record having missing or faulty data (e.g., "cousin" does not clarify the gender of a respondent). In such cases, the remaining faulty or missing data often are replaced with statistically imputed data.

"Imputation" is the term used to describe the replacement of missing data with plausible values. Most commonly, imputation is used when a respondent answers some questions on a survey but not others. This is a condition known as "item nonresponse." By contrast, when a selected individual does not respond to any question on the survey at all, or does not respond to enough key questions for the case to be useful for research purposes, this is a condition referred to as "unit nonresponse." In such cases, weighting adjustments are normally employed to account for these missing data. As an initial step, prior to any processing of the data, unit nonrespondents were discarded, and only unit respondents (i.e., item respondents and item nonrespondents for any given questionnaire item) were included in the subsequent editing, imputation, and analysis of NSDUH data. Throughout the remainder of this report, the terms "nonresponse" and "nonrespondent" are used to denote item nonresponse and item nonrespondent, respectively.

Once processed, imputed values cannot be distinguished from nonmissing values for a given variable in the final dataset. Therefore, observations with imputed data must be identified with a concomitant indicator variable. The vast majority of imputation-revised variables for the 2016 NSDUH have the prefix "IR" attached to their names. Although no missing data were possible for gender because a response to this item was required before the interview could proceed, the "IR" prefix for IRSEX was maintained for continuity with past years. Each imputed variable has an associated indicator variable, identified by the prefix "II" that can be used to identify which values were imputed and which were not. For some imputation-revised variables,

additional imputation indicators were created with the prefix "II2."³¹ These indicators gave more details about the source of the imputed or logically assigned value. The levels of a typical "II" imputation indicator are as follows:

- 1 = From questionnaire
- 2 = Logically assigned
- 3 = Statistically imputed
- 9 = Legitimate skip

Assignment of a code of 9 to an imputation indicator reflected the prior assignment of a legitimate skip code as part of the editing process (Section 2.4.2). Also, if a question had missing or ambiguous data and this question governed the skip logic for subsequent questions, respondents typically were skipped out of the subsequent questions. Respondents requiring imputation for the variables that governed a skip pattern typically receive a value of 3 (statistically imputed) for the imputation indicators associated with both the governing variables and the imputed variables that were nested within the skip pattern. For example, if a respondent had missing data for whether he or she had ever used a particular substance and was imputed to have never used it, the imputation indicators for recency of use of the substance, age at first use, frequency of use in the past 12 months (if applicable), and frequency of use in the past 30 days (if applicable) all were assigned a code of 3.

Various methods of imputation have been used since the NSDUH was first administered in the early 1970s.³² With the expansion of the NSDUH sample size in 1999, the predictive mean neighborhood (PMN) method for imputation was implemented and is currently used for most variables. Beginning in 2015, the modified predictive mean neighborhood (modPMN) imputation methodology was implemented for a subset of the drug variables including inhalants, hallucinogens, methamphetamine, and prescription drugs (i.e., pain relievers, stimulants, sedatives, and tranquilizers). Throughout the remainder of this report, the term "modPMN drugs" will be used to refer to this group of substances (i.e., inhalants, hallucinogens, methamphetamine, pain relievers, stimulants, sedatives, and tranquilizers). Similarly, the term "PMN drugs" will be utilized throughout the report, as appropriate, to refer to the set of substances that includes alcohol, tobacco products (i.e., cigarettes, cigars, pipes, and smokeless tobacco), marijuana, cocaine and crack cocaine, and heroin. Both PMN and modPMN are designed to incorporate the complex interrelationships among items in the current NSDUH, thus maintaining data consistency within individual respondent records. [Table 3.1](#) provides a summary of the types of imputation procedures used for each of the variables imputed in the NSDUH samples from 1999 through 2016.

³¹ Beginning in 2015, additional details about the source of the imputed or logically assigned value are stored in the indicator variable with the prefix "II" rather than the indicator variable with the prefix "II2" for imputed variables associated with drugs that were no longer considered comparable with data from survey years prior to 2015 (i.e., smokeless tobacco, inhalants, hallucinogens, methamphetamine, prescription pain relievers, prescription stimulants, prescription tranquilizers, and prescription sedatives).

³² Prior to the 2002 survey year, when it was renamed, the National Survey on Drug Use and Health (NSDUH) was originally known as the National Household Survey on Drug Abuse (NHSDA).

Table 3.1 Summary of Item Imputation Procedure Used, by Variable and NSDUH Year

Variable	1999¹	2000	2001	2002-2003	2004-2014	2015-2016
Interview Date	Random ²	Random	None	None	None	None
Age	None ³	None	None	None	None	None
Birth Date	None	Random	Random	Random	Random	Random
Gender	None	None	None	None	None	None
Race	USHD ⁴	PMN	PMN	PMN	PMN	PMN
Hispanic or Latino Origin Indicator	USHD	PMN	PMN	PMN	PMN	PMN
Marital Status	USHD	PMN	PMN	PMN	PMN	PMN
Hispanic or Latino Origin Group	USHD	PMN	PMN	PMN	PMN	PMN
Education	USHD	USHD	PMN	PMN	PMN	PMN
Employment Status	USHD	USHD	PMN	PMN	PMN	PMN
Immigrant	Not imputed	Not imputed	Not imputed	WSHD ⁵	PMN	PMN
Health Insurance	PMN	PMN	PMN	PMN ⁶	PMN	PMN
Lifetime Drug Usage: Alcohol, Tobacco Products, Marijuana, Cocaine/Crack, Heroin, Hallucinogen Child Drugs	PMN	PMN	PMN	PMN	PMN	PMN
Lifetime Drug Usage: Inhalants, Hallucinogens, Methamphetamine	PMN	PMN	PMN	PMN	PMN	modPMN
Recency and Frequency of Use ⁷ : Alcohol, Tobacco Products, Marijuana, Cocaine/Crack, Heroin	PMN	PMN	PMN	PMN	PMN	PMN
Recency and Frequency of (Mis)Use ⁸ : Inhalants, Hallucinogens, Methamphetamine, Pain Relievers, Stimulants, Sedatives, Tranquilizers	PMN	PMN	PMN	PMN	PMN	modPMN
Age at First Use: Alcohol, Tobacco Products, Marijuana, Cocaine/Crack, Heroin	PMN	PMN	PMN	PMN	PMN	PMN
Age at First Use: Inhalants, Hallucinogens, Methamphetamine	PMN	PMN	PMN	PMN	PMN	modPMN
Age at First Misuse: Pain Relievers, Stimulants, Sedatives, Tranquilizers	PMN	PMN	PMN	PMN	PMN	Random
Age at First Daily Cigarette Use	PMN	PMN	PMN	PMN	PMN	PMN
Past Year Initiation of Drug Misuse ⁹ : Pain Relievers, Stimulants, Sedatives, Tranquilizers	N/A	N/A	N/A	N/A	N/A	modPMN

Table 3.1 Summary of Item Imputation Procedure Used, by Variable and NSDUH Year (continued)

Variable	1999 ¹	2000	2001	2002-2003	2004-2014	2015-2016
Past Year Substance Dependence and Abuse: Alcohol, Marijuana, Cocaine, Heroin	Zero-fill	Zero-fill	Zero-fill	Zero-fill	Zero-fill	Zero-fill
Past Year Substance Dependence and Abuse: Inhalants, Hallucinogens, Methamphetamine, Pain Relievers, Stimulants, Sedatives, Tranquilizers	Zero-fill	Zero-fill	Zero-fill	Zero-fill	Zero-fill	modPMN
Personal and Family Income (Binary)	PMN	PMN	PMN	PMN	PMN	PMN
Personal and Family Income (Finer Categories)	PMN	PMN	PMN	PMN	PMN	PMN
Nicotine Dependence	Not imputed	Not imputed	Regression	Regression	Regression	Regression
Household Size (Roster-Derived)	PMN	PMN	PMN	PMN	PMN	PMN
Other Household Composition (Roster-Derived)	PMN	PMN	PMN	PMN	PMN	PMN

modPMN = modified predictive mean neighborhood; N/A = not applicable.

¹ The 1999 survey year also included a paper-and-pencil interviewing sample. The procedures listed here are from the computer-assisted interviewing sample.

² "Random" refers to a random assignment within a quarter for the interview date and a random assignment using age and interview date for the birth date. For age at first misuse, "Random" refers to the selection of age at first misuse for the prescription drugs. For this subset of drugs (i.e., pain relievers, stimulants, tranquilizers, and sedatives), imputation of an age at first misuse is only conducted for past year initiates of the drug in question. A random assignment was done using birth date, interview date, and information on month and year of first misuse for specific prescription drugs.

³ "None" means that no missing values were encountered after editing, and thus no imputation was necessary. For gender (from the 2002 survey onward) and age, missing values were precluded by design (see Chapter 4).

⁴ "USHD" refers to the unweighted sequential hot-deck method of item imputation described in this report (see Section 3.1.1.1).

⁵ "WSHD" refers to the weighted sequential hot-deck method of item imputation described in this report (see Section 3.1.1.2).

⁶ Although predictive mean neighborhood (PMN) imputation was the method used for health insurance in all years since the 1999 survey, imputation also was applied to more detailed health insurance variables in the surveys from 2002 onward.

⁷ "Recency and Frequency of Use" included variables measuring recency of use, 12-month frequency of use, 30-day frequency of use, and binge drinking frequency in past 30 days. "Binge drinking" was defined as having five or more drinks on the same occasion on a given day for males and four or more drinks on the same occasion on a given day for females.

⁸ "Recency and Frequency of (Mis)use" included variables measuring recency of any use (prescription drugs only; not measured/imputed prior to 2015), recency of (mis)use, 12-month frequency of use, and 30-day frequency of (mis)use. For prescription drugs, the recency of any use and recency of misuse measures included lifetime any use and lifetime misuse, respectively. For inhalants, hallucinogens, and methamphetamine, lifetime use is imputed separately from recency of use.

⁹ Since 2015, respondents who indicate misuse of prescription drugs in the past year are asked whether or not they ever misused the drug of interest more than one year prior to the interview date. The response to this question is used to determine whether or not the respondent is a past year initiate of prescription drug misuse. Prior to 2015 for prescription drugs, and for nonprescription drugs in all years, past year initiation status is obtained indirectly from the final imputed age and date of first use variables.

3.1.1 Previously Used Hot-Deck Imputation Methods

With any method of imputation, missing responses for a particular variable (hereafter, termed "base" variable) are replaced by values from similar respondents with respect to a number of characteristics (hereafter, "auxiliary variables"). If "similarity" is defined in terms of a single predicted value from a model, these auxiliary variables can be represented by that value. The respondent with the missing value for the base variable is called the "recipient," and the respondent from whom values are borrowed to replace the recipient's missing value is called the "donor." Donors and recipients are distinguished by the completeness of their records with regard to the variable(s) of interest (i.e., the donor has complete data for that variable, and the recipient does not). The term "hot deck" is used to refer to imputations made on recipient base variables using donor values from the same dataset. The PMN methodology, and modPMN methodology for continuous variables, utilized on the NSDUH are specialized hot-deck methods and are described in greater detail later in this chapter. For more information on the general hot-deck method of item imputation, see Little and Rubin (1987, pp. 62-67).

For the 2016 NSDUH, the only imputations that did not incorporate either the PMN method or the modPMN method were those used for the birth date, date of first (mis)use, age at first misuse for prescription drugs, and nicotine dependence variables. Imputation procedures for the birth date are described in Chapter 4. Methods used to assign the date of first (mis)use and age at first misuse for prescription drugs are described in Chapter 6, and imputation procedures for the nicotine dependence variables are described in Chapter 7. Two other hot-deck methods—unweighted sequential hot deck (USHD) and weighted sequential hot deck (WSHD) (Cox, 1980, pp. 721-725; Iannacchione, 1982)—were used in past surveys.³³ In the sections that follow, the features and limitations of USHD, WSHD, and the random nearest neighbor hot deck (NNHD) are discussed as background for the development of the PMN and modPMN methodologies.

3.1.1.1 Unweighted Sequential Hot Deck

In a sequential hot-deck procedure, data are first ordered using specific criteria, and the last reported value in the sequence is substituted for each missing value as the data are processed. In USHD, the selection of a response for imputation purposes is independent of the sampling weight associated with the data record from which the response is taken and the data record to which a response is being imputed. USHD imputation is, therefore, based upon the tacit assumption that nonrespondents would answer in a manner similar to that of respondents immediately adjacent to them in an appropriately sorted data file and hence that the data associated with the nearest neighbor are appropriate for the imputation of missing values (Cox, 1980, p.721).

³³ The USHD method was used exclusively for the 1991-1998 surveys, for the paper-and-pencil interviewing sample from the 1999 survey, and for all demographic variables in the computer-assisted interviewing sample from the 1999 survey. In the 2002-2003 surveys, missing values in the immigrant variables required WSHD imputation. Note, however, that the USHD and WSHD methods have not been used on the NSDUH since the 2000 and 2003 survey years, respectively.

Implementation of the USHD method (and of hot-deck methods, in general) involves three basic steps:

1. **Construct imputation classes.** When there is a strong logical association between the base variable and certain auxiliary variables, the dataset is partitioned by these auxiliary variables, and imputation procedures are implemented independently within the resulting imputation classes defined by the cross of these auxiliary variables.
2. **Sort the analytic file.** Within each imputation class, the file is sorted by auxiliary variables relevant to the item being imputed. The sort order of the auxiliary variables is chosen to reflect the degree of importance of the auxiliary variables in their relation to the base variable being imputed (i.e., those auxiliary variables that are better predictors for the item being imputed are used as the first sorting variables). In general, two types of sorting procedures—a straight sort and a serpentine sort³⁴—were used in previous surveys to sort the files prior to imputation.
3. **Replace missing values with imputed values.** The sorted file is read sequentially. Each time an item respondent is encountered (i.e., the base variable is nonmissing), the base variable response is stored, updating the donor response. Any subsequent nonrespondent in the file receives the stored donor response, which in turn results in a statistically imputed response. Because the file is sorted by relevant auxiliary variables, the preceding item respondent (donor) closely matches the neighboring item nonrespondent (recipient) with respect to the auxiliary variables.

For any particular item being imputed under USHD, there is the risk of several nonrespondents appearing next to one another on the sorted file; in this situation, each would receive imputed values from the same donor. To detect this problem on the NSDUH in the survey years prior to 2001, the imputation donor was identified for every item being imputed, and frequencies by donor were examined. If several nonrespondents were aligned next to one another after sorting, sort variables were added or eliminated, or the ordering of the sort variables was modified, to ensure that multiple nonrespondents did not comprise adjacent records on the resulting file.

3.1.1.2 Weighted Sequential Hot Deck

WSHD improves upon USHD by incorporating the sampling weights when replacing missing values among recipient records in the final hot-deck assignment step. The earlier steps taken to impute for missing values under the WSHD method are the same as those for the USHD method; as in USHD, WSHD requires the formation of imputation classes and appropriate sorting (straight or serpentine) of the analytical file.

³⁴ Under a straight sort, a set of variables is sorted in ascending order by the first variable specified. Then, within each level of the first variable, the file is sorted in ascending order by the second variable specified, and so forth. In a serpentine sort, a set of variables is sorted so that the direction of the sort (ascending or descending) for subsequent variables changes each time the value of the preceding variable changes. The serpentine sort has the advantage of minimizing the change in the entire set of auxiliary variables every time any one of the variables changes its value. For an example of each, see Appendix A of the 2009 imputation report (Ault et al., 2011).

The WSHD procedure used in surveys prior to 2004 followed directly from Cox (1980). Specifically, once the imputation classes were formed, the data were divided into two datasets: one for respondents and one for nonrespondents. Scaled weights v_j were then derived for all nonrespondents using the following formula:

$$v_j = w_j s_+ / w_+; j = 1, 2, \dots, n,$$

where n is the number of nonrespondents, w_j is the sample weight for the j^{th} nonrespondent, w_+ is the sum of the sample weights for all the nonrespondents, and s_+ is the sum of the sample weights for all the respondents. The respondent data file was partitioned into zones of width v_j , where the imputed value for the j^{th} nonrespondent was selected from a respondent in the corresponding zone of the respondent data file. This selection algorithm is an adaptation of Chromy's (1979) sequential sample selection method.

WSHD controls the number of times a donor can be selected and allows each respondent the chance to be a donor because a respondent is selected within each v_j . Consequently, the most important benefit of the weighted sequential hot-deck method is the elimination of bias in the estimates of means and totals, particularly when the response rate is low or when the covariates explain only a small amount of variation in the specified variable. In addition, many surveys sample subpopulations at different rates, and using the sample weights allows the imputed data for the nonrespondents to have the same mean (for the specified variables) as the respondents. In other words, the weighted hot deck preserves the respondent's weighted distribution in the imputed data (Cox, 1980).

3.1.1.3 Unweighted Random Nearest Neighbor Hot Deck

Another commonly used imputation method—one not directly used on the NSDUH, but related to the PMN and modPMN methods—is random nearest neighbor hot deck (NNHD) (Little & Rubin, 1987, p. 65). With this method, a donor set or neighborhood deemed "close to" the recipient, with respect to a number of covariates, is used to select a donor at random. The distance between the values of the recipient and potential donors for each of the auxiliary variables is calculated, and then the donors for the neighborhood are chosen such that the maximum of these distances is less than a certain threshold value, referred to as "delta." This neighborhood is restricted, using imputation classes described previously, so that the potential donors' values of the base variable are consistent with the recipient's preexisting nonmissing values of related variables.

Because a distance function is used to define "closeness" between the recipient and a donor under NNHD, there is less of a problem of sparseness of the donor class when imputing for continuous variables. It should be noted, however, that the distance function involving categorical or nominal variables is typically ad hoc and often hard to justify.

3.2 Development of the Predictive Mean Neighborhood Methodology

The PMN methodology developed for and implemented on the 1999 NSDUH was an attempt to address the shortcomings, while retaining the positive characteristics, of the hot-deck

imputation methods discussed above. It is a combination of two commonly used imputation methods: non-model-based NNHD (Little & Rubin, 1987, p. 65) and a modification of Rubin's model-assisted predictive mean matching (PMM) method (Rubin, 1986), which matches a missing value to the observed value with the closest predicted mean. The PMN method enhances Rubin's PMM method, in that PMN can be applied to both discrete and continuous variables, either individually or jointly. PMN also enhances the NNHD method for discrete variables so that the distance function used to find neighbors is no longer in terms of the original predictor variables and therefore does not require arbitrary scaling.

In addition, the PMN methodology offers the following advantages over the imputation methods employed on earlier NSDUHs:

- **A greater number of auxiliary variables may be used to determine donors.** Using a model-based hot-deck technique like PMN allows auxiliary variables to be incorporated in two ways: first, as covariates in models, and second, in likeness constraints³⁵ applied to potential donors. Under USHD and WSHD, the number of auxiliary variables is limited in part by the problem of sparse neighborhoods; donors must match recipients for all variables used to form imputation classes. If too many variables are used to form imputation classes, some classes may be very small and contain few or no item respondents to serve as donors. By contrast, under PMN, the donors need only be "close" to the recipients with respect to the predicted values determined by the models, even when the models include numerous covariates. Moreover, PMN ensures that a sufficient number of potential donors comprise the donor neighborhood, so that likeness constraints may be applied on the donor set as needed.
- **Relative importance of auxiliary variables is determined by standard estimating equation techniques.** Under USHD and WSHD, as implemented, the selection of classing and sorting variables was sometimes ad hoc, and in the former instance, weights were not utilized. In PMN, by contrast, objective criteria based on a more rigorous methodology (i.e., regression) quantify the relationship between a given covariate and the response variable in the presence of other covariates, so that the response variable itself is indirectly used to determine donors. Further, the sampling weights can be incorporated in PMN regression models without difficulty.
- **Internal consistency of the post-imputation record is guaranteed.** In PMN, the donor pool can be restricted to those making the post-imputation record logically consistent. For example, if a recipient must receive a cocaine past year frequency of use between 30 and 50, the donors can be restricted to ensure that the recipient receives such a value. In USHD and WSHD, the classing and sorting variables cannot easily be used to guarantee this; there may not be a donor in the imputation class who will create a consistent record.

³⁵ Likeness constraints are flexible constraints that govern the similarity between donors and recipients. See Section 3.3.3.1 for details.

- **Correlations across response variables are accounted for by making the imputation multivariate.** In comparison with other model-based methods, discrete and continuous variables can be handled jointly and relatively easily in PMN by using the idea of sequential univariate modeling. Further, differential weighting factors can be objectively assigned to different elements of the predictive mean vector depending on the variability of predicted means in the dataset.

3.3 Implementation for the Predictive Mean Neighborhood Methodology

The implementation of PMN on the NSDUH involves three basic steps: response propensity (RP) adjustment, prediction (PRD) modeling, and hot-deck imputation. At the most basic level, the RP adjustment reallocates the weights of the item nonrespondents to item respondents; the PRD model calculates predicted means for both; and the hot-deck step assigns final values to the item nonrespondents based on a distance function derived from these predicted means. These steps are described in more detail in the following sections and are combined in three different ways, called PMN "types" (Section 3.4), to complete imputation procedures.

3.3.1 Step 1: Response Propensity Adjustment

Response propensity is defined as the probability of response, whether at the unit level or item level. The purpose of response propensity is to adjust the sampling weights for item nonresponse so that the item respondent weights that are used only during the imputation process are representative of the entire domain of interest. In the RP step of PMN, the item response propensity is modeled as a function of a predetermined set of covariates. The model can be thought of as a special case of the generalized exponential model (GEM)³⁶ developed for weighting procedures, in that imputations that are done at the item level are similar in nature to the weight adjustments made for entire units.

There are four key inputs to the item RP modeling step:

1. **Analysis weights.** For all imputation procedures, the best available analysis weight is used as an input to the NSDUH imputation procedures. Because of the timing of 12-month processing and, in particular, the coordination between the weighting and imputation tasks in each NSDUH year, most variables that undergo imputation utilize the preliminary analysis weight, PANALWT. PANALWT was developed by the NSDUH weighting team for imputation purposes. The dwelling unit-level design weights were adjusted for nonresponse using a weighting class method by state. The nonresponse-adjusted dwelling unit weights and the person-level design weights were further adjusted for nonresponse at the person level using a weighting class method by state and age groups (12-17, 18-25, 26-34, 35-49, and 50+). PANALWT is the product of all design weights and the two nonresponse adjustment factors. For those variables that are processed later in the annual cycle, the final analysis weight ANALWT may be used instead, if it is available at the time of imputation processing

³⁶ The GEM macro, which was written in SAS/IML[®] software, was developed at RTI International for weighting procedures and is described in detail in Appendix A of the person-level sampling weight calibration report in the 2016 NSDUH methodological resource book (Center for Behavioral Health Statistics and Quality, 2018a).

for that variable. The pair variables described in Chapter 14 utilize yet another weight, PRANALWT. See the person-level sampling weight calibration report in the 2016 NSDUH methodological resource book (MRB; Center for Behavioral Health Statistics and Quality, 2018a) for full descriptions of the final analysis weights. See the questionnaire dwelling unit-level and person pair-level sampling weight calibration report in the 2016 NSDUH MRB (Center for Behavioral Health Statistics and Quality, 2018b) for full descriptions of the pair weights.

2. **Domain indicator.** In this report, a "domain" is defined as the set of respondents who are included in models and for whom predicted means are calculated. For many NSDUH variables that undergo imputation, the domain includes all unit respondents. For others, the domain is a subset of unit respondents. For example, the marital status question is only asked of respondents aged 15 or older. In this case, the domain indicator is set to zero for respondents aged 12 to 14 and one for respondents aged 15 or older. The domain indicator is an important input to the tables in Appendix A, where item response rates are reported for each variable that undergoes imputation on the NSDUH.
3. **Item response indicator.** The item response indicator is set to zero for item nonrespondents and one for item respondents. GEM uses this indicator to identify the item nonrespondents and item respondents when reallocating the weights appropriately. The item response indicator is an important input to the tables in Appendix A, where item response rates are reported for each variable that undergoes imputation on the NSDUH.
4. **Covariates.** GEM uses a predetermined list of covariates to allocate the weights from item nonrespondents to item respondents. The covariates tend to be variables that are correlated with (1) the propensity to respond, (2) the variable that is undergoing imputation, or (3) both. The goal is to avoid bias in the PRD models by allocating the weights of the item nonrespondents to similar item respondents, without too greatly inflating the variance of the estimates that utilize these weights (Center for Behavioral Health Statistics and Quality, 2018a). Appendix D lists the starting and final covariate lists for each RP model fit in NSDUH imputation procedures.

3.3.2 Step 2: Prediction Modeling

Utilizing the RP-adjusted weights that were derived in the previous step, the PRD model calculates predicted means, which are used in the hot-deck step(s) to create neighborhoods and select donors. The dependent variable in the model is usually the variable, or some transformation of that variable, that is undergoing imputation. Each model is built using only those cases within the domain with complete responses for that item. Predicted means are then calculated for all of the domain members, whether or not they were item respondents, using the values for the covariates and the estimates for the regression coefficients.

For categorical outcome variables, logistic regression models are used for the PRD models. For continuous variables, linear regression models are fit. For count variables, Poisson regression models are used. For response variables that are proportions (e.g., months on welfare; see Chapter 12), a logit transformation is applied to the proportion, and a linear regression model is utilized. The variable sets in which some transformations of the response variables were

implemented include the self-administered demographics (Chapter 5), drugs (Chapter 6), and income (Chapter 12).

The goal of the PRD model is to protect against bias for as many analysis domains as possible, so these models tend to start with long lists of covariates (Allison, 2008).³⁷ Appendix D lists the starting and final covariate lists for each PRD model fit in NSDUH imputation procedures. In contrast to explanatory (association) models where model parsimony is a relevant metric of a model's appropriateness, the focus in a prediction setting is on the predicted values only.

The SUDAAN[®] software package³⁸ is used to fit nearly all the PRD models used in the NSDUH.³⁹ Under PMN, all covariates from the applicable starter list are utilized unless SUDAAN produces warning messages, which indicate nonconvergence or model instability. Generally, in these cases, covariates whose coefficients have high p-values are dropped until SUDAAN no longer produces these warning messages. The primary advantage of using SUDAAN to fit PRD models is that the standard errors associated with the regression coefficients properly account for the complex survey design. The predicted means are the same using SUDAAN as they are using, for example, the analogous base SAS[®] procedure (given the same set of covariates),⁴⁰ but the decision on which covariates to drop in the event of model instability or nonconvergence is more informed under SUDAAN because the standard errors, and the corresponding test statistics and p-values for model coefficients, account for stratification and clustering.

In the particular case of some of the logistic regression models, the warning messages produced by SUDAAN may be triggered when a cross-classification of the outcome variable and a covariate has empty or nearly empty cells. Covariates of this type are highly correlated with the outcome variable but cannot be used in the PRD model. However, they are often used in the hot-deck step to identify suitable donors. For the types of regression-based PRD models used for each variable that underwent imputation using PMN, see [Table 3.2](#).

Table 3.2 Regression Models Used for Each Variable Imputed with Predictive Mean Neighborhood

Variable	Domain	Type of Regression Model	SAS/SUDAAN Procedure ¹
Demographics			
Marital Status	15 years or older	Multinomial Logistic	MULTILOG
Race	All respondents	Multinomial Logistic	MULTILOG
Hispanic or Latino Indicator	All respondents	Binomial Logistic	RLOGIST
Hispanic or Latino Group	Hispanics	Multinomial Logistic	MULTILOG
Education Level	All respondents	Multinomial Logistic	MULTILOG
Employment Status	15 years or older	Multinomial Logistic	MULTILOG

³⁷ Allison (2008) writes that "... imputation models should be relatively 'rich' so that they may be congenial with lots of different models that could be of interest."

³⁸ SUDAAN[®] is a registered trademark of Research Triangle Institute.

³⁹ The only exception is the finer income categories, described in detail in Chapter 12.

⁴⁰ SAS[®] software is a registered trademark of SAS Institute Inc.

Table 3.2 Regression Models Used for Each Variable Imputed with Predictive Mean Neighborhood (continued)

Variable	Domain	Type of Regression Model	SAS/SUDAAN Procedure¹
Immigrant Status: Born-in-U.S. Indicator	All respondents	Binomial Logistic	RLOGIST
Immigrant Status: Age of Entry	Not born in U.S.	Simple Linear	REGRESS
Drugs²			
Lifetime Drug Use	All respondents ³	Binomial Logistic	RLOGIST
Recency of Drug Use, "Hierarchical" Drugs	All lifetime users for past year vs. not past year; all past year users for past month vs. not past month	Binomial Logistic	RLOGIST
Recency of Drug Use, Pipes	All lifetime users	Binomial Logistic	RLOGIST
Recency of Drug Use, Other Drugs	All lifetime users	Multinomial Logistic	MULTILOG
12-Month Frequency of Drug Use	All past year users	Simple Linear	REGRESS
Daily Drug Use over Past 30 Days, Cigarettes	All past month users	Binomial Logistic	RLOGIST
30-Day Frequency of Drug Use, Cigarettes	All past month users except those who used daily over the past 30 days	Simple Linear	REGRESS
30-Day Frequency of Drug Use, Other Drugs	All past month users	Simple Linear	REGRESS
Age at First Drug Use	All lifetime users	Simple Linear	REGRESS
Household Composition			
Total Number of Rostered People	All respondents	Poisson	LOGLINK
Total Number of Children Younger than 18	All respondents	Poisson	LOGLINK
Total Number of People Aged 65 or Older	All respondents	Poisson	LOGLINK
Indicator of Whether the Respondent Has Family Members in Household	All respondents	Binomial Logistic	RLOGIST
Total Number of Respondent's Family Members in the Household (Includes Foster Relationships)	All respondents	Poisson	LOGLINK
Total Number of Respondent's Family Members in the Household Younger than 18 (Includes Foster Relationships)	All respondents	Poisson	LOGLINK

Table 3.2 Regression Models Used for Each Variable Imputed with Predictive Mean Neighborhood (continued)

Variable	Domain	Type of Regression Model	SAS/SUDAAN Procedure ¹
Income			
Source of Income	All respondents	Binomial Logistic	RLOGIST
Months on Welfare	All respondents who received welfare payments or welfare services in the past year	Simple Linear	REGRESS
Total Income (Binary)	All respondents	Binomial Logistic	RLOGIST
Finer Category Income	All respondents	Time-to-Event (Survival)	LIFEREG
Health Insurance			
Specific Types of Health Insurance Coverage	All respondents	Binomial Logistic	RLOGIST
Any Other Health Insurance	All respondents with none of the specific types of health insurance	Binomial Logistic	RLOGIST

¹ See RTI International (2013) for more information on all SAS®-callable SUDAAN® procedures in this table except LIFEREG and LOGISTIC. See SAS Institute Inc. (2014) for more information on the LIFEREG and LOGISTIC procedures. SAS software is a registered trademark of SAS Institute Inc. SUDAAN is a registered trademark of Research Triangle Institute.

² The drugs imputed using the predictive mean neighborhood methodology include alcohol, cigarettes, cigars, pipes, smokeless tobacco, marijuana, cocaine/crack, heroin, and lifetime use of hallucinogen "child" drugs (i.e., LSD, PCP, Ecstasy, ketamine, DMT/AMT/Foxy, *Salvia divinorum*, and "any other" hallucinogen).

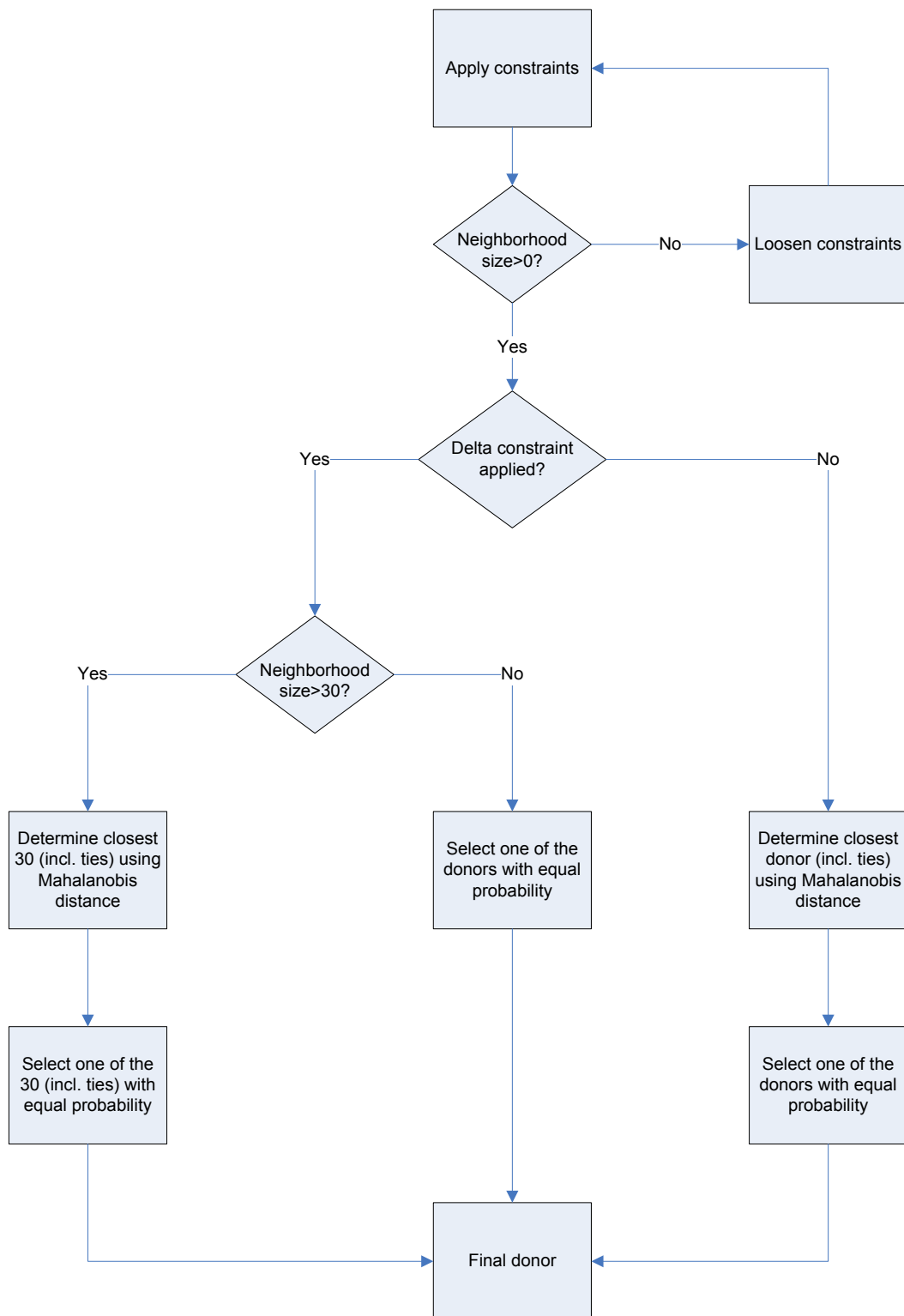
³ The domain for lifetime use of hallucinogen child drugs is defined by the imputation-revised variable for lifetime use of overall hallucinogens.

3.3.3 Step 3: Hot-Deck Imputation

After sampling weights have been appropriately adjusted in the RP step and predicted means have been calculated in the PRD step, the hot-deck step⁴¹ of PMN is applied to select a donor for each item nonrespondent. The algorithm used to select donors is graphically displayed in the flowchart in [Figure 3.1](#). Briefly, likeness constraints are loosened in an iterative fashion until PMN yields a nonempty donor neighborhood. Mahalanobis distance is then used to rank donors by closeness to the item nonrespondent, and a final donor is selected at random for a given recipient. Many of the hot-deck components used in PMN are described below and appear explicitly in the tables of Appendix E.

⁴¹ Under PMN, there are two situations on the NSDUH in which the imputation is not a hot-deck step, but is a stochastic imputation based solely on the predicted mean(s) of the recipient, of the type described in Section 5.1 of the evaluation of imputation methods for the NSDUH report (Center for Behavioral Health Statistics and Quality, 2017e). These ideas have their origin in Singh, Grau, and Folsom (2004), where "Centered PMN" is discussed as an alternative to PMN. These exceptions include the provisional hot-deck step for imputation set 2 for the health insurance variables (Chapter 13) and the provisional hot-deck step for hallucinogen child drugs (Chapter 6).

Figure 3.1 Donor Selection Algorithm



3.3.3.1 Logical and Likeness Constraints

Logical constraints and likeness constraints are restrictions placed on the set of donors to make imputed values consistent with preexisting, nonmissing values of the item nonrespondents (recipients) and to make candidate donors as much like the recipients as possible. Logical constraints are fixed constraints that prevent logical inconsistencies between variables, and likeness constraints are flexible constraints that govern the similarity between donors and recipients.

The logical constraints are never removed, because to do so would risk the selection of a donor that produces an inconsistent post-imputation record. For example, for the employment status variable, if the item nonrespondent is known to be employed, but full-time vs. part-time status is unknown, the imputed value must come from a donor who is employed as well.

Likeness constraints are placed on the pool of donors to make the attributes of the neighborhood as close as possible to those of the recipient. For example, age and employment status are correlated. A likeness constraint exploits this correlation by requiring the donor's age to be within 5 years of the item nonrespondent's age, but likeness constraints may be loosened if they happen to force the donor pool to be empty.

One likeness constraint that is used in all hot-deck steps, regardless of the variable being imputed, is the delta constraint. This particular likeness constraint requires the donor's predicted mean to be within 5 percent (delta) of the item nonrespondent's predicted mean for each element of the predictive mean vector. If the predicted means are probabilities, the values of delta vary depending upon the value of the predicted mean.

Each delta is defined as 5 percent of the predicted probability if the probability were less than 0.5 and is defined as 5 percent of 1 minus the predicted probability if the probability were greater than 0.5. This allows for a looser delta for predicted probabilities close to 0.5 and a tighter delta for predicted probabilities close to 0 or 1. The range of values for delta across various predicted probabilities is shown in [Table 3.3](#). Logical constraints and likeness constraints, including the order in which likeness constraints are loosened for some variables, are presented in the tables in Appendix E.

Table 3.3 Values of Delta for Various Predicted Probabilities

Predicted Probability (p)	Delta
$p \leq 0.5$	$0.05p$
$p > 0.5$	$0.05(1 - p)$

NOTE: In very rare cases, due to transformations of the response variable for the prediction model, p was slightly less than 0 or slightly greater than 1. This occurred only for 30-day frequency and 12-month frequency. If p was less than 0, delta was set to $0.05|p|$. If p was greater than 1, delta was set to $0.05(p - 1)$. See Sections 6.5.1.3.6 and 6.5.1.6.2 for a detailed description of these transformations and their associated back-transformations.

3.3.3.2 Predictive Mean Vector

The predicted means from the PRD step play a central role in the donor selection algorithm depicted in [Figure 3.1](#), through the construct of the predictive mean vector. The predictive mean vector is essentially a list of predicted means from the PRD modeling step. In simple cases, the predictive mean vector contains only one element, such as the predicted age at which a respondent began using a drug. In complex cases, the predictive mean vector includes several elements from several different PRD models, such as the predicted recency and predicted frequency of use for a given drug.

When the PRD model is a logistic regression model, predicted means are calculated for each level of the outcome variable. For example, the employment status variable that undergoes imputation has four levels: employed full time, employed part time, unemployed, or other. Therefore, a single PRD model is fit using a four-level outcome variable, yielding predicted probabilities for each level, as follows:

- E1: P(respondent is employed full time)
- E2: P(respondent is employed part time)
- E3: P(respondent is unemployed)
- E4: P(other)

Note, however, that the predictive mean vector for the employment status variable contains only three elements. It does not include the predicted probability for the reference cell, which in this case is the "other" level, because that level is implicitly defined by the presence of the other three predicted means.

Occasionally, the predicted means are adjusted so that they are made conditional on what is known for a given respondent. Continuing the example above, some respondents report that they have a job but are unclear about the number of hours they usually work in a week. Because the NSDUH definition of part-time versus full-time employment status was based on working a minimum of 35 hours in a usual week, the predictive mean vector is made conditional on employment of any sort for these respondents. Therefore, the single predicted mean used for these respondents is equal to $E1 \div (E1 + E2)$, P(respondent is employed full time | respondent is employed). Conditional probabilities are also used in the binary income hot-deck step and the drug recency/frequency hot-deck steps. Predictive mean vectors are presented in the tables in Appendix E.

3.3.3.3 Univariate vs. Multivariate Matching and Assignment

If the predictive mean vector consists of only one element, *univariate* matching is used to select a donor. If the predictive mean vector consists of more than one element, *multivariate* matching is used to select a donor. The donor may also give values to the item nonrespondent for more than one variable, a situation known as multivariate assignment. Similarly, if the donor provides values for only one variable, the hot-deck step uses univariate assignment. [Table 3.4](#) shows examples of NSDUH variables that were imputed using each of the four combinations of univariate/multivariate matching and assignment.

Table 3.4 Examples of Variables Imputed Using Each of the Four Combinations of Univariate/Multivariate Matching and Assignment

	Variables Imputed One at a Time (Univariate Assignment)	Variables Imputed in a Set (Multivariate Assignment)
Predictive Mean Vector Has One Element (Univariate Matching)	Hispanic/Latino Origin (Section 4.3.3)	Finer Category Income (Section 12.3.2)
Predictive Mean Vector Has More Than One Element (Multivariate Matching)	Marital Status (Section 4.3.1)	Lifetime Drug Use (Section 6.5.1.1)

Whether the hot-deck step employs univariate or multivariate matching, Mahalanobis distance is used to rank the donors by closeness to the item nonrespondent. The Mahalanobis distance is used instead of Euclidean distance in order to standardize the distance in terms of the population variances and covariances of vector components.⁴² It is given by

$$\sqrt{(\boldsymbol{\mu}_R - \boldsymbol{\mu}_{NR})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu}_R - \boldsymbol{\mu}_{NR})},$$

where $\boldsymbol{\mu}_R$ refers to the predictive mean vector for a given item respondent, and $\boldsymbol{\mu}_{NR}$ is the predictive mean vector for a given item nonrespondent. The matrix $\boldsymbol{\Sigma}$ is the variance-covariance matrix of the predictive mean vector, using the set of item respondents that comprise that domain. Because the square of the Mahalanobis distance is a monotone function of the distance itself, and only the ranking of the donors (instead of the absolute distance measure) is used in the algorithm, the additional step of taking the square root of the squared distance is not performed in practice.

3.3.3.4 Missingness Patterns

For many variables imputed on the NSDUH, item nonrespondents were segregated into patterns of nonresponse called missingness patterns. Missingness patterns arise in two ways. First, for sets of variables that underwent multivariate assignment, item nonrespondents were segregated into missingness patterns based on which variables were missing. Second, a new missingness pattern could emerge when logical editing restricted an item nonrespondent to only a subset of the variable's possible values. The example for employment status discussed above applies here as well: respondents whose employment status was completely unknown had a different missingness pattern than did those who were known to be employed. Often, different predictive mean vectors were used, and different constraints were applied, for different missingness patterns. Many of the tables in Appendix E are segregated by missingness pattern for this reason.

⁴² When univariate matching is employed, the ranking of donors by Mahalanobis distance is equivalent to the ranking of donors by Euclidean distance. This is not necessarily true when multivariate matching is employed.

3.3.3.5 Final Assignment of Donor Values

Logical and likeness constraints are used to form a neighborhood of potential donors from the pool of item respondents within each missingness pattern. Logical constraints are always imposed to maintain internal consistency, whereas likeness constraints are removed or relaxed in a predetermined order until this donor neighborhood is nonempty. Once a nonempty neighborhood is found, the rest of the PMN donor selection algorithm depends on whether or not the delta constraint was applied.

If the delta constraint was applied, all the members of the neighborhood are similar to the recipient with respect to the predictive mean vector. The final donor is then randomly selected with equal probability from among the "closest" (in terms of Mahalanobis distance) 30 members of the neighborhood; potential donors whose Mahalanobis distance from the recipient are equal ("ties") are accounted for in the donor selection algorithm depicted in [Figure 3.1](#). If, on the other hand, the delta constraint was not applied, to ensure that the final donor is as close to the item nonrespondent as possible with respect to the predicted means, the donor with the smallest Mahalanobis distance is selected as the final donor. If there is more than one "closest" donor (i.e., there are ties), the final donor is randomly selected with equal probability from among the closest donors. At the conclusion of the hot-deck step, the item nonrespondent receives values from the selected donor for a single variable (in the univariate assignment case) or for a set of variables (multivariate assignment).

3.4 Predictive Mean Neighborhood "Types"

There are three types of PMN as applied on the NSDUH: Type 1, *single RP/single PRD* (Section 3.4.1); Type 2, *multiple RP/multiple PRD* (Section 3.4.2); and Type 3, *single RP/multiple PRD* (Section 3.4.3). Each of the three PMN types is a coordinated application of the three basic steps of PMN discussed in Section 3.3.

In PMN, an imputation "set" is a set of variables for which a single donor is used in the final hot-deck step.⁴³ Sets are formed based on the extent of correlation among variables and the level of missingness in the data. Variables with few missing values and no strong relationships with other variables tend to be processed in an imputation set by themselves. Closely related variables tend to be processed together in the same set to preserve, as much as possible, correlations between variables in the data. However, the more variables that are included in a multivariate set, the less likely it is that a nonempty neighborhood can be found using the delta constraint. Even though there are many advantages to using a multivariate imputation set, one disadvantage in several instances is not being able to apply the delta constraint. [Table 3.5](#) lists the imputation sets for each variable group imputed using the PMN methodology discussed in this report and the PMN type used to process each set.

⁴³ Section 3.4.2 defines and discusses the differences between *provisional* and *final* hot-deck steps in the context of PMN.

Table 3.5 PMN Types Applied to Each Variable Group and Imputation Set

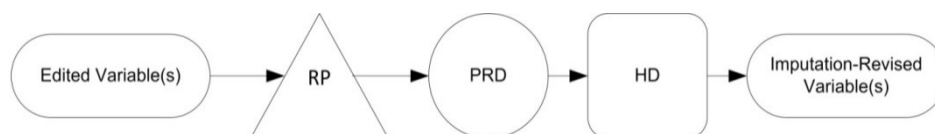
Variable Group	Imputation Set	PMN Type
Interviewer-Administered Demographics	All (5 sets)	1 (Single RP/Single PRD)
Self-Administered Demographics	All (3)	1
Drugs ¹	Lifetime (2)	3 (Single RP/Multiple PRD)
	Recency of Pipe Use	1
	Recency/Frequency, Other Drugs (7)	2 (Multiple RP/Multiple PRD)
	Cigarette Ever Daily Used	1
	Age at First Use (8)	1
Roster	All (8)	1
Income	Binary	3
	Finer Categories	1
Health Insurance	Types of Health Insurance	3
	Any Other Health Insurance	1
Roster Pair	Pair Relationship	1
	Multiplicities (6)	1
	Household Counts, Sibling-Sibling and Spouse-Spouse (4)	1
	Household Counts, Parent-Child	2

PRD = prediction; PMN = predictive mean neighborhood; RP = response propensity.

¹The drugs imputed using the PMN methodology include alcohol, cigarettes, cigars, pipes, smokeless tobacco, marijuana, cocaine/crack, heroin, and lifetime use of hallucinogen "child" drugs (i.e., LSD, PCP, Ecstasy, ketamine, DMT/AMT/Foxy, *Salvia divinorum*, and "any other" hallucinogen).

3.4.1 Type 1: Single Response Propensity/Single Prediction

PMN Type 1, the single RP/single PRD type, involves a single iteration of the three basic steps described in Section 3.3: RP, PRD, and hot-deck imputation. Many variables that undergo imputation using the PMN methodology in the standard processing cycle use this type, including all the demographics and roster variables and the age-at-first-use drug variables. [Figure 3.2](#) illustrates the single RP/single PRD type of PMN imputation.

Figure 3.2 PMN Type 1: Single Response Propensity/Single Prediction

Usually the single RP/single PRD type involves univariate assignment in the hot-deck step,⁴⁴ but it may involve univariate or multivariate matching, depending on the PRD model. If

⁴⁴ Finer income categories is an example of an imputation set that uses the single RP/single PRD type, but its hot-deck step utilizes multivariate assignment. If the item nonrespondent is missing the finer income category at both the personal and family level, the donor will provide values for both variables in a single hot-deck step. The PRD model is fit using the family-level finer income category.

the PRD model is a dichotomous logistic regression, linear regression, or Poisson regression model, univariate matching is used because the model produces only one predicted mean. If, on the other hand, the PRD model is a polytomous logistic regression model, multivariate matching is used because the model produces more than one predicted mean (i.e., the predicted probability associated with each level of the response variable). In either implementation, there is only one PRD model.

In the single RP/single PRD type, for the univariate assignment case, the item response indicator is based on the single variable that is being assigned in the hot-deck step. If the single variable is missing, the case is an item nonrespondent; otherwise, the case is an item respondent. In the multivariate assignment case, the case is an item respondent if all variables that are assigned in the hot-deck step are nonmissing.

3.4.2 Type 2: Multiple Response Propensity/Multiple Prediction

PMN Type 2, multiple RP/multiple PRD, involves multiple iterations of the single RP/single PRD type. However, for all iterations except the last, the hot-deck step is provisional instead of final and involves univariate matching and univariate assignment.⁴⁵ These provisional hot-deck steps tend to be straightforward with respect to constraints and predictive mean vectors, because their only purpose is to fill in missing values so that variables earlier in the sequence can be used as covariates in the RP and PRD models for variables later in the sequence.⁴⁶ In the last iteration, a final hot-deck step is completed, where final imputed values are assigned for all variables involved in the models. This final hot-deck step always involves multivariate matching and multivariate assignment. The predicted means from all PRD models are used in this final hot-deck step, and a single record is used to fill all the missing values, thus preserving the relationships among the variables of interest. This is the most refined type of PMN. The recency and frequency variables (within each drug family) that are imputed using the PMN methodology follow this type. [Figure 3.3](#) illustrates the multiple RP/multiple PRD type of PMN imputation.

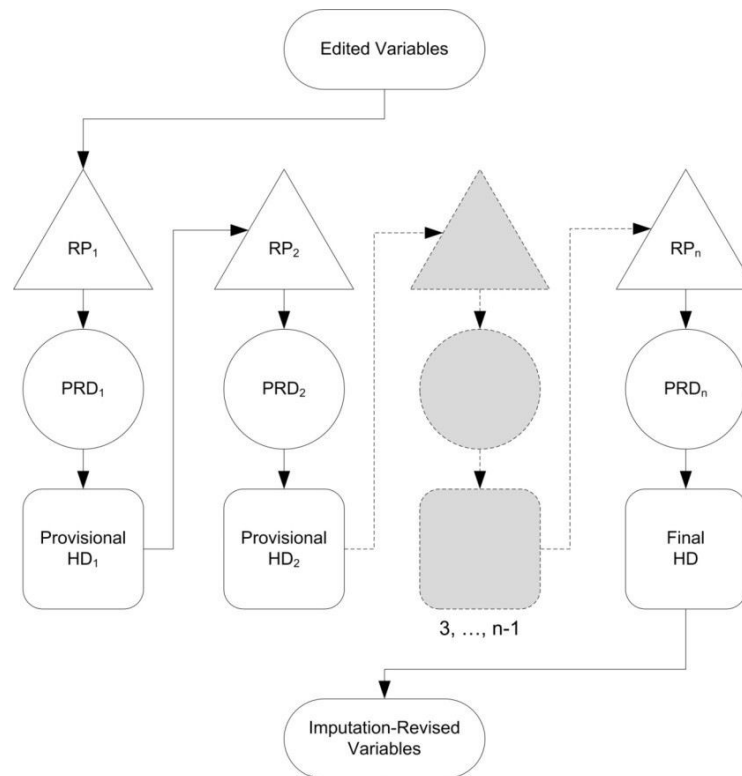
In the multiple RP/multiple PRD type, multiple univariate PRD models are used. The standard approach to multivariate modeling, with a given set of outcome variables (including both discrete and continuous), is likely to be computationally intensive due to the volume of model parameters and the difficulty in specifying a suitable covariance structure. Following Little and Rubin's (1987) proposal of a joint model for discrete and continuous variables, and its implementation by Schafer (1997), it is possible to fit a pure multivariate model for multivariate imputation, but it would require making distributional assumptions. Moreover, because of the obvious problem of specifying an accurate probability distribution underlying survey data, none of the existing solutions take the survey design into account. In the multiple RP/multiple PRD

⁴⁵ There is one exception to the rule that provisional hot-deck steps involve univariate matching and univariate assignment. The provisional hot-deck step for cocaine and crack lifetime use utilizes multivariate assignment because both variables are used in the subsequent PRD model for heroin. The delta constraint refers to both predicted means, but in the calculation of Mahalanobis distance, only the cocaine predicted mean is used. Therefore, with respect to matching, this is not strictly univariate or multivariate; it is a little of both. See Section 6.5.1.1.4.

⁴⁶ There are exceptions. In a few imputation sets that use PMN Type 2 or PMN Type 3 (single RP/multiple PRD), provisional hot-deck steps are not completed because the variables earlier in the sequence are not used as covariates for variables later in the sequence. This occurs for some of the imputation sets for health insurance (Chapter 13) and roster pairs (Chapter 14).

type, a multivariate model is fitted by a series of univariate parametric models (including the polytomous case), such that variables modeled earlier in the sequence have a chance to be included in the covariate set for subsequent models in the sequence.

Figure 3.3 PMN Type 2: Multiple Response Propensity/Multiple Prediction



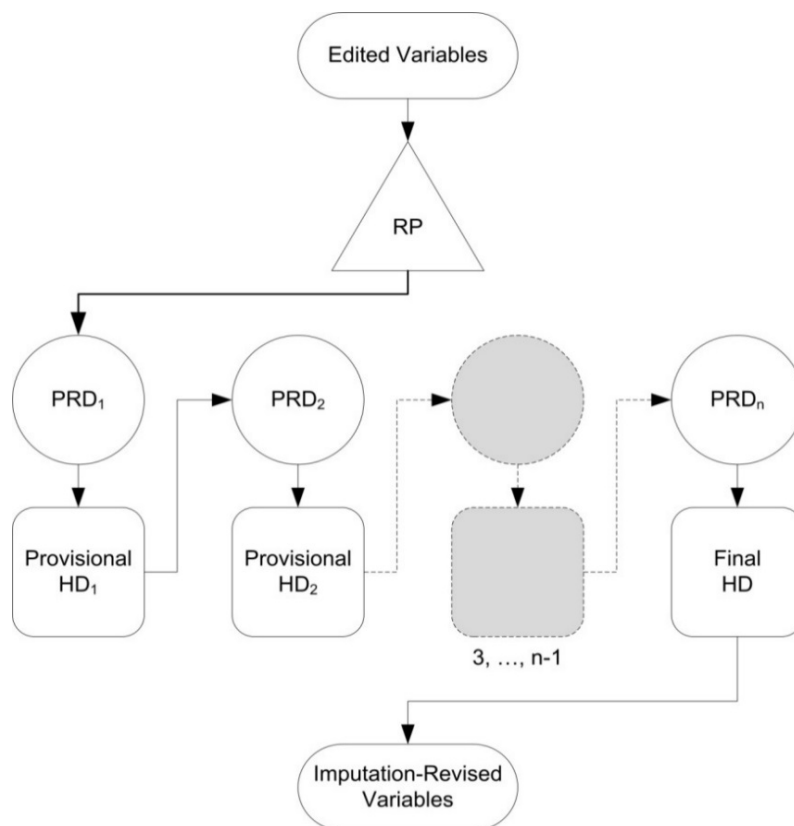
For variables imputed by PMN Type 2 and PMN Type 3 (single RP/multiple PRD), the order in which variables were modeled is of some importance because variables early in the sequence have the potential to be part of the set of covariates for variables later in the sequence, but variables late in the sequence cannot be used for modeling for the earlier variables because of missing values. Note that usually not all variables in the sequence were missing for a particular incomplete record. Nevertheless, models were developed for all the variables in a univariate fashion for reasons mentioned earlier. For the drugs imputed using the PMN methodology, the sequence of imputation was determined by considering such factors as the level of prejudice and discrimination associated with the drugs, the level of "missingness" in the data (Appendix A), and the degree to which one set of drugs could be used as predictors for other drugs. The decisions on sequencing for other imputation sets were made using similar criteria. For some respondents, some but not all of the variables in the imputation set are missing. This gives rise to missingness patterns (Section 3.3.3.4). Typically, in the final hot-deck step, only the predictive mean vector elements corresponding to missing variables are used to match donors with item nonrespondents. However, likeness constraints (and sometimes logical constraints) are often used to preserve relationships between the missing and nonmissing variables. Although the nonmissing values would not be replaced by the corresponding values from the donor, some degree of correlation between missing and nonmissing variables is expected to be preserved using these constraints.

The multiple RP/multiple PRD type works well for closely related variables that have different domains and different nonresponse patterns, because the separate RP steps account for these. The recency and frequency variables provide a good example: the domain of the recency models for drugs imputed using the PMN methodology consists of all lifetime users; the domain of the 12-month frequency model (if applicable) consists of all past year users; and the domain of the 30-day frequency model (if applicable) consists of all past month users. The provisional imputation-revised values may be used as covariates in later models, or even may be used to define the domains of later RP models.

3.4.3 Type 3: Single Response Propensity/Multiple Prediction

In PMN Type 3, the single RP/multiple PRD type, a single RP model is applied to all the variables modeled in the PRD steps. This is a less refined version of the preceding type, because it involves the fitting of only one RP model and is not as sensitive to different domains and response patterns among the outcome variables. The same weights are used for all PRD models. Among variables imputed using the PMN methodology, the lifetime drug use variables and source-of-income variables are examples of imputation sets that follow this type. [Figure 3.4](#) illustrates the single RP/multiple PRD type of PMN imputation.

Figure 3.4 PMN Type 3: Single Response Propensity/Multiple Prediction



3.5 Development of the Modified Predictive Mean Neighborhood Methodology

Just as the PMN methodology was an attempt to improve the hot-deck imputation methods discussed in Section 3.1.1, the modPMN methodology is an attempt to improve PMN. It retains the three basic steps of PMN—RP adjustment, PRD modeling, and imputation⁴⁷—but introduces some improvements into each of the three steps and includes a second cycle where each cycle looks much like PMN Type 2 (Section 3.4.2). The modPMN methodology addresses the following perceived shortcomings of PMN.

- Under modPMN, categorical variables are imputed stochastically using only the item nonrespondent's predicted values from the model. This collocated stochastic imputation step for categorical variables replaces the hot-deck step of PMN, which does not "center" the imputed value. This decision was made in part due to the results described in Singh, Grau, and Folsom (2004), where the notion of "Centered PMN" was introduced. Stochastic imputation is a much simpler way to center the imputed value of a categorical variable than hot-deck imputation, and it is implemented in a way that reduces the probability of extreme results (e.g., where either all item nonrespondents are imputed to "Yes" or all item nonrespondents are imputed to "No" for a dichotomous variable).
- For closely related variables, instead of using multivariate assignment in the imputation step (Section 3.3.3.3), closely related variables are imputed one at a time under modPMN, using a second cycle in a manner similar to IVEware (Raghuathan, Solenberger, & Van Hoewyk, 2002). The second cycle ensures that variables later in the sequence can be used to assist with imputation for variables earlier in the sequence. The one-at-a-time imputation approach greatly simplifies the predictive mean vectors and constraints that are used by PMN for hot-deck steps utilizing multivariate assignment.
- Instead of imputations being done separately within age groups (as under PMN), the domain members under modPMN are segregated into five bins based on their estimated propensity to respond, and the PRD modeling and imputation steps are implemented separately for each bin. This promotes "double protection," as described in Kott and Folsom (2010), without creating extreme weights, which can occur if item nonresponse is low. In contrast, the weight used in the hot-deck step of PMN does not promote double protection.
- Under modPMN, the segregation of drug variables into imputation sets is done within measures and across drugs instead of within drugs and across measures (as under PMN). For example, instead of including hallucinogens recency, 12-month frequency, and 30-day frequency in a single imputation set, hallucinogens recency is in the same set as the recencies of other modPMN drugs, hallucinogens 12-month frequency is in the same set as the 12-month frequencies of other modPMN drugs,

⁴⁷ As opposed to PMN where the third step typically involves a hot-deck imputation, the imputation step under modPMN depends on the type of variable being imputed. Under modPMN, continuous variables are imputed using a hot-deck step, whereas categorical variables are imputed using a stochastic process. For additional details, see Section 3.6.3.

and hallucinogens 30-day frequency is in the same set as the 30-day frequencies of other modPMN drugs. This simplifies the constraints and streamlines the logic.

- Under modPMN, child drugs (see Section 6.1.2) are handled after parent drugs have already undergone imputation. This also simplifies the constraints and logic and promotes a more customized approach to imputation within the small domains that are involved when child drugs are treated.
- Several warning messages produced by the modeling procedures in SUDAAN are disregarded when the PRD models are fit. This reduces the time required to fit the models with little impact on the predicted values.
- The fitting of PRD models was partially automated. The model-fitting process can sometimes be time-consuming. However, under modPMN, the code is set up so that the statistician fitting the PRD model can easily implement automated methods like stepwise selection to get the model to converge quickly.

The RP, PRD, and imputation steps are described in Section 3.6, and the arrangement of those steps into a procedure (which includes "cycling" as described above) is described in Section 3.7.

3.6 Procedural Steps for the Modified Predictive Mean Neighborhood Methodology

The implementation of modPMN on the NSDUH involves the same three basic steps that are used in PMN, as described in Section 3.3: RP adjustment, PRD modeling, and imputation. These steps are described in more detail in the following sections and are organized into blocks of 12 SAS programs to complete the modPMN imputation procedures as described in Section 3.7. The descriptions below assume that the reader is already familiar with Section 3.3, and the focus is mainly on the differences between these steps in modPMN and the analogous steps in PMN.

3.6.1 Step 1: Response Propensity Adjustment

The RP adjustment step under modPMN is very similar to the RP adjustment step under PMN as described in Section 3.3.1. The differences are as follows:

- Under PMN, typically the RP step is run independently within age groups (see Section 3.8.1). Under modPMN, all age groups are included in the same RP model.
- Under PMN, each item respondent is assigned a weight adjustment factor: namely, the reciprocal of the estimated response propensity. This factor is multiplied by the analysis weight, typically PANALWT, to create the weight used in the PRD model. Under modPMN, all domain members are divided into five "bins" of equal size based on their estimated response propensity, and the PRD modeling step and the imputation step are almost always completed within these bins.

These improvements are based on the "double protection" idea described in Kott and Folsom (2010). In PMN, the weights w_k used in the PRD models are

$$w_k = \frac{d_k}{p_k},$$

where w_k is the weight used in the PRD model for item respondent k , d_k is the analysis weight for item respondent k , and p_k is the estimated response propensity for item respondent k . When modPMN was first tested, as described in Chapter 5 of the evaluation of imputation methods for the NSDUH report (Center for Behavioral Health Statistics and Quality, 2017e), the weights described in Kott and Folsom (2010) were used in the PRD models:

$$w'_k = d_k \frac{1 - p_k}{p_k}.$$

However, this often resulted in extreme weights when the item response rate was high. This problem was addressed for modPMN by forming five RP bins instead of using the weights w'_k directly.

3.6.2 Step 2: Prediction Modeling

The PRD modeling step under modPMN is very similar to the PRD step under PMN as described in Section 3.3.2. The differences are as follows:

- Under PMN, typically the PRD step is run independently within age groups (see Section 3.8.1). Under modPMN, the PRD step is run independently within the five RP bins. Section 3.6.1 explains how the five RP bins are formed.
- Under PMN, independent variables are dropped until the modeling procedure in SUDAAN produces no warning messages. Under modPMN, most of these warning messages are disregarded. As long as the algorithm converges, reasonable estimates of regression coefficients are produced, and reasonable test statistics associated with the regression coefficients are produced, then the PRD model is considered final.
- Under PMN, when a PRD model fails to converge or SUDAAN produces warning messages, independent variables are dropped until the model converges and the warnings do not appear. Under modPMN, statisticians use a more automated method to fit the model. These automated methods are used mainly to save processing time, because there are more models to fit under modPMN than there are under PMN.⁴⁸ If the model does not converge with the full set of independent variables, then the statistician runs a procedure that automatically drops all independent variables that are causing the matrix of independent variables to be singular, and then re-fits the model. If the model still does not converge, then the statistician runs a procedure that uses stepwise selection in SAS to reduce the list of independent variables, drops any remaining independent variables that cause a singularity problem, and re-fits the

⁴⁸ There are two cycles instead of one, and there are five RP bins instead of three or four age groups.

model. Finally, if the model still does not converge, the statistician follows the same procedure that he or she would in PMN, which requires more manual intervention.

Table 3.6 is the modPMN analogue of Table 3.2 for PMN. It shows which SUDAAN procedures are used to fit the PRD models for modPMN. Note that modPMN drugs fall into two categories: prescription drugs (prescription pain relievers, prescription tranquilizers, prescription stimulants, and prescription sedatives), and other modPMN drugs (hallucinogens, inhalants, and methamphetamine).

Table 3.6 Regression Models Used for Each Variable Imputed Using modPMN

Variable¹	Domain	Type of Regression Model	SUDAAN Procedure²
Lifetime Drug Use, Other modPMN Drugs	All respondents	Binomial Logistic	RLOGIST
Recency of Any Drug Use, Prescription Drugs	All respondents	Multinomial Logistic	MULTILOG
Recency of Drug Misuse, All modPMN Drugs	All lifetime users	Multinomial Logistic	MULTILOG
12-Month Frequency of Drug Use, Other modPMN Drugs	All past year users	Simple Linear	REGRESS
30-Day Frequency of Drug Use, All modPMN Drugs	All past month users/misusers	Simple Linear	REGRESS
Age at First Drug Use, Other modPMN Drugs	All lifetime users	Simple Linear	REGRESS
Past Year Initiation, Prescription Drugs	All past year misusers	Binomial Logistic	RLOGIST

modPMN = modified predictive mean neighborhood.

¹ The drug variables imputed using the modPMN methodology include those associated with the four prescription drug categories (prescription pain relievers, prescription tranquilizers, prescription stimulants, and prescription sedatives) and the three "other" modPMN drugs (hallucinogens, inhalants, and methamphetamine).

² See RTI International (2013) for more information on all SAS®-callable SUDAAN® procedures in this table except LIFEREG and LOGISTIC. See SAS Institute Inc. (2014) for more information on the LIFEREG and LOGISTIC procedures. SAS software is a registered trademark of SAS Institute Inc. SUDAAN is a registered trademark of Research Triangle Institute.

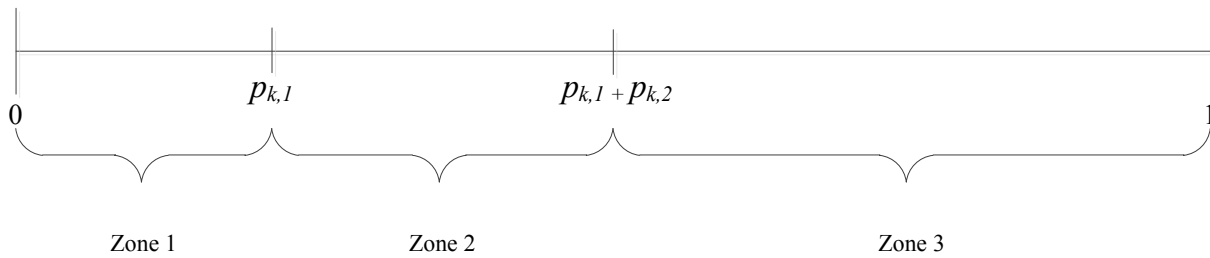
3.6.3 Step 3: Imputation

For continuous variables, the imputation step of modPMN is identical to the imputation step of PMN (Section 3.3.3). The hot-deck step for modPMN always employs univariate matching and univariate assignment (Table 3.4). However, for categorical variables, the imputation step of modPMN is not a hot-deck step. It is a random imputation based solely on the predicted means of the item nonrespondents. It also includes a feature that reduces the probability of extreme imputation results. The imputation step for categorical variables in modPMN is called "collocated stochastic imputation." As for the PRD step, imputations are normally performed within the RP bin. For continuous variables, a likeness constraint is included that requires the donor to be in the same bin as the item nonrespondent. This section explains stochastic imputation first and then describes the collocation process in detail.

To implement stochastic imputation, a random number between 0 and 1 is drawn for each item nonrespondent, and the imputed value is equal to the predicted mean "zone," which contains the random number. In the simplest case, the variable to be imputed is dichotomous (0 = No and 1 = Yes), and if the predicted mean from the PRD step for a given item nonrespondent k is p_k and the random number drawn for item nonrespondent k is s_k , then the imputed value is 1 if, and only if, $s_k < p_k$. If the categorical variable to be imputed has G levels, then item nonrespondent k will have G predicted means from the PRD model such that $\sum_{g=1}^G p_{k,g} = 1$, where $p_{k,g}$ is the predicted probability that item nonrespondent k has a value of g .

Figure 3.5 shows stochastic imputation for a three-level variable that can take a value of 1, 2, or 3. If the random number s_k is in Zone 1 (i.e., $0 < s_k < p_{k,1}$), then the imputed value for item nonrespondent k is 1. If the random number s_k is in Zone 2 (i.e., $p_{k,1} < s_k < p_{k,1} + p_{k,2}$), then the imputed value for item nonrespondent k is 2. Finally, if the random number s_k is in Zone 3 (i.e., $p_{k,1} + p_{k,2} < s_k < 1$), then the imputed value for item nonrespondent k is 3.

Figure 3.5 Stochastic Imputation for a Three-Level Categorical Variable



Note that if the s_k values are drawn independently, then an extreme imputation can result by chance if all the s_k values happen to be close to 0 or close to 1, or if just by chance each s_k happens to fall into a narrow zone. The collocation process embedded in collocated stochastic imputation is intended to reduce the probability of an extreme imputation by taking an approach more akin to systematic sampling than Poisson sampling, where each unit in the population is drawn into the sample resulting from independent decisions based on selection probability. In that case, the randomness of the process could lead to a very large sample or a very small sample.

A simple way to implement the systematic sampling-style approach for a dichotomous variable would be to randomly sort the R item nonrespondents (perhaps using s_k), choose f randomly from the uniform distribution on the interval $[0,1)$, and then impute k with a 1 if, and only if,

$$c_k = f / R + (k - 1) / R < p_k ,$$

where the value of k indicates the random sort order of the nonrespondents from 1 to R .

Observe that if all the p_k values were exactly equal to p , this approach would come very close to imputing a value of $1(p \times 100)$ percent of the time. This suggests a slightly better methodology would replace c_k with

$$c'_k = \left(\sum_{j=1}^k w_j + [f-1]w_k \right) / \sum_{j=1}^R w_j < p_k ,$$

where w_j is the weight of element j (and the elements were again randomly sorted and labeled in sort order). Here, if all the p_k values were exactly equal to p , this revised approach would produce a weighted average among the R elements needing imputation very close to p . Note that if all the w_j values were equal, c'_k would be the same as c_k .

If there are more than two categories in the variable needing imputation, the p_k value in the above method can be replaced by $p_{k,g}$, where $p_{k,g}$ is the probability that item nonrespondent k is in one of the first g categories (if there is no logical order for the categories, any order will do). Item nonrespondent k would be imputed into category g if, and only if,

$$p_{k,g-1} \leq c_k < p_{k,g} \text{ or } p_{k,g-1} \leq c'_k < p_{k,g} ,$$

where $p_{k,0}$ is defined to be 0. In fact, this latter approach involving the weights has been implemented on the NSDUH since 2015.

3.7 Implementation of the Modified Predictive Mean Neighborhood Methodology

Unlike PMN, there are not three types of modPMN as described in Section 3.4. All imputation sets except two in the 2016 NSDUH included more than one variable,⁴⁹ and all were implemented using the same general arrangement of steps, somewhat akin to Type 2 of PMN (Section 3.4.2), but with a second cycle. Section 3.7.1 describes how variables are ordered in modPMN, and although the ordering is still important, the second cycle makes it less important for modPMN than it is for PMN. Section 3.7.2 describes the building blocks for modPMN implementation, a set of 12 SAS programs that implement the RP, PRD, and imputation steps for a single variable. This is similar to Type 1 of PMN (Section 3.4.1). Finally, Section 3.7.3 describes how the building blocks described in Section 3.7.2 are arranged to complete the imputation procedures for an imputation set.

Note that this section describes the implementation of modPMN in general terms. Readers interested in the implementation of modPMN for a specific imputation set are referred to the sections of the report covering that imputation set. The imputation sets that were processed with modPMN and the corresponding sections of the report are presented in [Table 3.7](#). See the footnotes in [Table 3.1](#) for details about these imputation sets.

⁴⁹ The exceptions to this are the imputation sets for OxyContin, which include only one variable in each set (OxyContin recency of any use; OxyContin recency of misuse). Because only one variable is included in each of the sets for OxyContin, no cycling is performed.

Table 3.7 Imputation Sets Processed Using modPMN

Number	Variables in Set	Section
1	Lifetime Drug Usage: Inhalants, Hallucinogens, Methamphetamine	6.5.2.3
2	Recency of Any Use: Pain Relievers, Stimulants, Sedatives, Tranquilizers	6.5.2.3
3	Recency of Any Use: OxyContin	6.5.2.3
4	Recency of (Mis)Use: Inhalants, Hallucinogens, Methamphetamine, Pain Relievers, Stimulants, Sedatives, Tranquilizers	6.5.2.3
5	Recency of Misuse: OxyContin	6.5.2.3
6	Recency of Use: LSD, PCP, Ecstasy, Ketamine, DMT/AMT/Foxy, <i>Salvia divinorum</i>	6.5.2.3
7	12-month Frequency of Use: Inhalants, Hallucinogens, Methamphetamine	6.5.2.4
8	30-day Frequency of (Mis)Use: Inhalants, Hallucinogens, Methamphetamine, Pain Relievers, Stimulants, Sedatives, Tranquilizers	6.5.2.4
9	Age at First Use: Inhalants, Hallucinogens, Methamphetamine	6.5.2.4
10	Age at First Use: LSD, PCP, Ecstasy	6.5.2.4
11	Past Year Initiation of Drug Misuse: Pain Relievers, Stimulants, Sedatives, Tranquilizers	6.5.2.3
12	Past Year Substance Dependence and Abuse: Hallucinogens	7.3.2
13	Past Year Substance Dependence and Abuse: Inhalants	7.3.2
14	Past Year Substance Dependence and Abuse: Methamphetamine, Stimulants	7.3.2
15	Past Year Substance Dependence and Abuse: Pain Relievers	7.3.2
16	Past Year Substance Dependence and Abuse: Sedatives, Tranquilizers	7.3.2

modPMN = modified predictive mean neighborhood.

3.7.1 Sequencing of Variables for modPMN

In PMN, the sequencing of variables within an imputation set is very important, because only variables earlier in the sequence are used as independent variables in the models for variables later in the sequence. Therefore, the imputation for variables later in the sequence can be said to be, in some sense, "better" than the imputation for variables earlier in the sequence. In the second cycle of modPMN, as described in Section 3.7.3, all other variables in the imputation set are used as independent variables for each variable. However, it may be said that the models are more fine-tuned for variables later in the sequence, because a larger set of independent variables was used for them in both the first cycle and the second cycle. The conclusion is that the sequencing of variables within a set is still important for modPMN but not quite as important as it is for PMN.

There are three clear candidates for the sequencing of variables within a set. The first is by descending item response rate, the method adopted by IVEware: all else being equal, imputation will have more of an impact on variables with lower item response rates. The second is based on the sequencing in the questionnaire, from beginning to end, which will tend to lead to the preservation of logical relationships and skip patterns. The third candidate approach is based on some measure of the variable's analytic importance.

Not surprisingly, this third approach is highly subjective and difficult to determine. The first and second are straightforward to determine, but the first is more difficult to implement because it requires year-to-year changes to the procedures, as item response rates tend to vary from year to year. (The sequencing of questions in the questionnaire rarely changes from year to year.) In consideration of the above, the approach taken since the 2015 NSDUH was a compromise between these three approaches:

- For imputation sets 4 and 8 in [Table 3.7](#), all other modPMN drugs preceded all prescription drugs, but within these larger drug categories, sequencing was done by descending item response rate. [Table 3.8](#) provides an example of the sequencing.
- For all other imputation sets in [Table 3.7](#), the variables were sequenced by descending item response rate.

The rationale for the former is that it allows all other modPMN drugs to be used as independent variables for the prescription drugs for both cycles. The prescription drugs were deemed more analytically important than the other modPMN drugs.

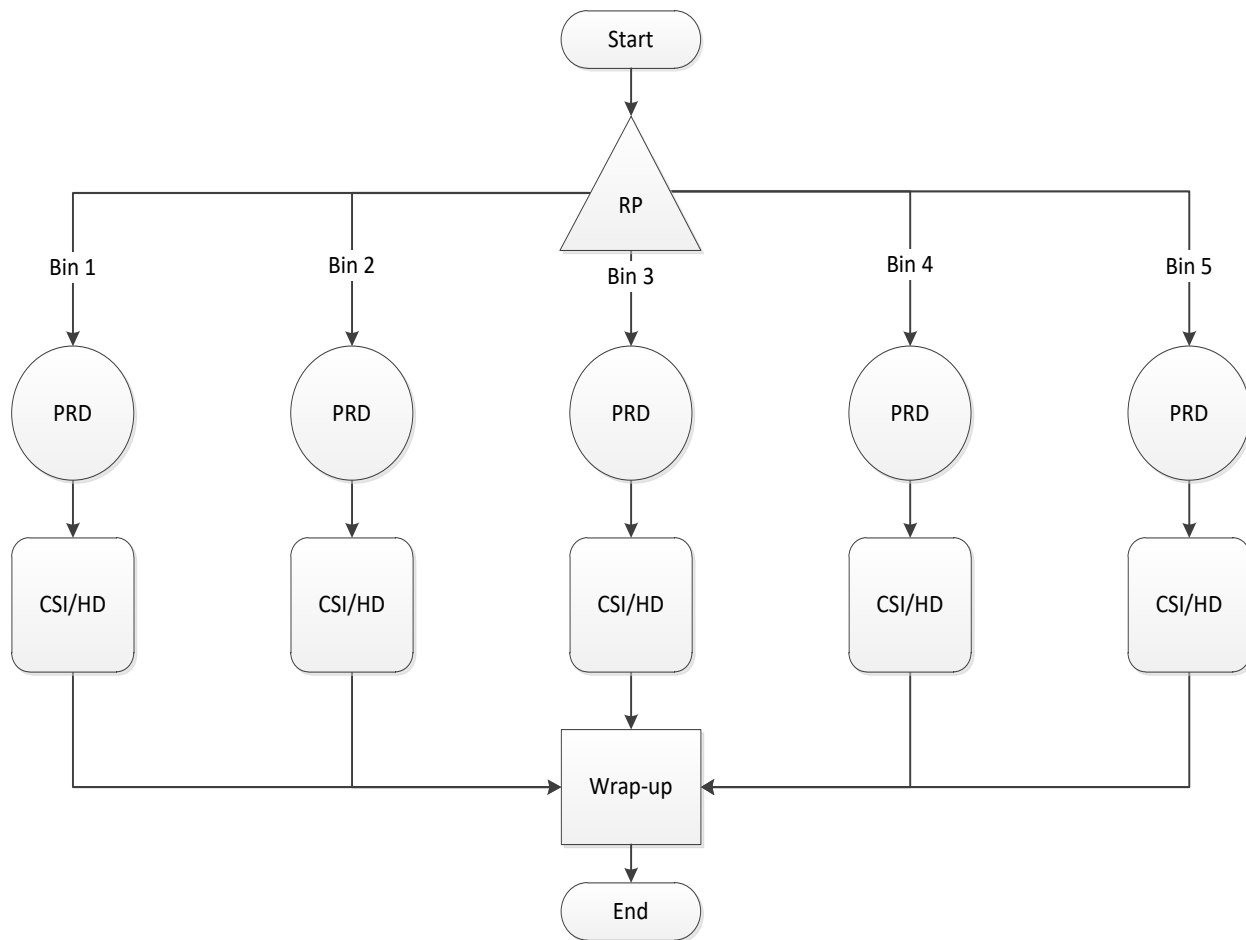
Table 3.8 Example of Sequencing for Imputation Sets 4 and 8

Drug	Drug Type	Rank by Item Response Rate (Descending)	Position in Sequence
Hallucinogens	Other	1	1
Inhalants	Other	5	2
Methamphetamine	Other	7	3
Prescription Pain Relievers	Prescription	6	7
Prescription Tranquilizers	Prescription	3	5
Prescription Stimulants	Prescription	2	4
Prescription Sedatives	Prescription	4	6

3.7.2 Application of modPMN for a Single Variable

To apply modPMN for a single variable, 12 programs are used, as shown in [Figure 3.6](#). The first program implements the RP step, which divides the domain members into five bins as described in Section 3.6.1. Within each bin, a PRD step (described in Section 3.6.2) and an imputation step (Section 3.6.3) are implemented. The imputation step involves collocated stochastic imputation for categorical variables and hot-deck imputation for continuous variables. The final program merges the five output datasets from the within-bin imputation steps into a single dataset, which contains imputed values for the variable that is undergoing imputation.

Figure 3.6 Twelve-Step Block of modPMN Programs



CSI = collocated stochastic imputation; HD = hot deck; modPMN = modified predictive mean neighborhood; PRD = prediction; RP = response propensity.

3.7.3 Cycling in modPMN

When modPMN is applied to an imputation set, the three procedural steps of modPMN (Section 3.6) are always applied in the same coordinated way, by sequencing the 12-step blocks within two cycles across variables. At the end of each 12-step block in the first cycle, the variable has undergone provisional imputation, and at the end of each 12-step block in the second cycle, the variable has undergone final imputation. For each 12-step block, provisionally imputed values for all variables are used in the RP and PRD models whenever they are available.

Table 3.9 lists the set of independent variables used in the RP and PRD models for each combination of cycle and variable number. Note that there is always a base set of independent variables, which are typically demographic variables that have either already undergone imputation or have no missing values for some other reason.

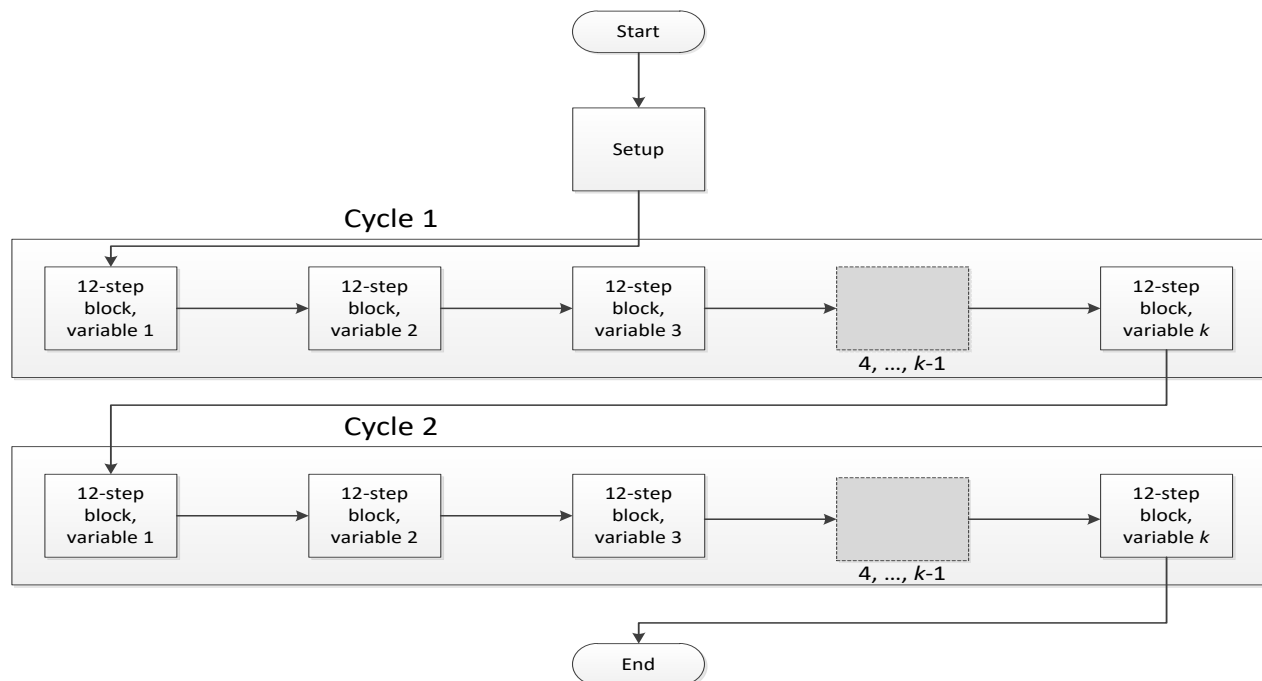
Table 3.9 Independent Variables Used in Response Propensity and Prediction Steps of modPMN

Cycle	Variable	Independent Variables
1	1	Base set
	2	Base set + provisionally imputed version of variable 1
	3	Base set + provisionally imputed versions of variables 1 and 2
	...	...
	k	Base set + provisionally imputed versions of variables 1, 2, ..., $k - 1$
2	1	Base set + provisionally imputed versions of variables 2, 3, ..., k
	2	Base set; final imputed version of variable 1; provisionally imputed versions of variables 3, 4, ..., k
	3	Base set; final imputed versions of variables 1 and 2; provisionally imputed versions of variables 4, 5, ..., k
	...	...
	k	Base set + final imputed versions of variables 1, 2, ..., $k - 1$

modPMN = modified predictive mean neighborhood.

Figure 3.7 explains how the 12-step blocks are arranged in order to complete imputation for a particular imputation set. In the setup step, all variables are collected that are needed for the imputation of the set, including the edited versions of all variables undergoing imputation, the base set of covariates, the analysis weight, and the sample design variables needed for the PRD steps. The process then moves sequentially through the 12-step blocks for each variable nested within the two cycles. At the end of the second cycle, final imputed variables have been created for all variables in the imputation set.

Figure 3.7 Arranging the 12-Step Blocks to Complete Imputation for an Imputation Set



3.8 Special Auxiliary Variables: Age Group and State Rank

The age group and state of residency auxiliary variables apply to several of the imputation sets described in Chapters 4 through 7 and 11 through 14. Across variable groups, many imputation sets are processed separately by age group, regardless of the type of PMN that was used. The state of residence is used to construct a state-rank variable, which is then used in imputation for the drug variables (Chapter 6) and the income variables (Chapter 12).

3.8.1 Age Groups

The variables related to drug use, household composition, income, and health insurance are highly correlated with age. This, along with the desire to use parallel processing to expedite the time it takes to impute all the variables, led to the decision to separate the imputation procedures for these variables into distinct age groups when using the PMN methodology. Therefore, the drug use variables for alcohol, cigarettes, cigars, pipes, smokeless tobacco, marijuana, cocaine/crack, heroin, and the lifetime use variables for hallucinogen "child" drugs (i.e., LSD, PCP, Ecstasy, ketamine, DMT/AMT/Foxy, *Salvia divinorum*, and "any other" hallucinogen) were imputed within each of three age groups: 12 to 17, 18 to 25, and 26 or older. The household composition (roster-derived), income, and health insurance variables were imputed within the following four age groups: 12 to 17, 18 to 25, 26 to 64, and 65 or older.⁵⁰ The roster pair variables (i.e., the variables related to the relationship between two respondents from the same sampled dwelling unit) were often divided into age groups depending on the ages of both pair members.

When imputation was performed using a hot-deck step, the age group restriction could be considered a likeness constraint. However, the models also were built separately within the age groups, so this restriction was not loosened unless no other options were available. Although the demographic variables did not always show a high correlation with age, the imputation of missing values in the demographic variables also was performed within age groups. This was done to maintain consistency with how the other variables were imputed and to facilitate parallel processing. The same three age groups that were used for drugs were also used for demographics. Occasionally, small sample sizes necessitated the aggregation of age groups at the modeling stage. In particular, the models for education level (highest grade completed) were fit within the 12-17 and 18-or-older age groups. In the employment status models, the 15-17 and 18-25 age groups were aggregated and models were fit within the 15-25 and 26-or-older age groups. Finally, all age groups were aggregated for the Hispanic/Latino group, marital status, and immigrant age-of-entry models.

3.8.2 State Rank

Because state-level estimates are an important product of the NSDUH, there has been interest in requiring the donor to be from the same state as the recipient. However, this could not

⁵⁰ Age groups were sometimes aggregated for the health insurance procedures. See Chapter 13 for details.

always be implemented because of insufficient pools of donors.⁵¹ In such cases, information about the state of residence of each respondent was incorporated into the modeling and (when applicable) hot-deck steps of the PMN and modPMN procedures by grouping respondents into three categories based on the ranking of their state of residence. For lifetime drug use, the states were ranked by the weighted proportion of lifetime users of the drug of interest. For recency and frequency of drug use, the states were ranked by the weighted proportion of past month users of the drug of interest. For income, the states were ranked by the weighted proportion of respondents whose personal incomes during the prior calendar year were greater than or equal to \$20,000. These state-rank variables were used as covariates in the RP and PRD steps and sometimes in likeness constraints in the hot-deck step.

⁵¹ In the hot-deck step for some of the demographic variables, a likeness constraint required the donor to be from the same segment as the item nonrespondent. Segments never cross state lines, so this can be viewed as a refined use of the state of residence. In practice, this constraint often had to be removed because many segments included only a handful of unit respondents.

4. Editing and Imputation for the NSDUH Interviewer-Administered Demographics Variables

4.1 Introduction

This chapter discusses procedures for editing and imputing the interviewer-administered questions for demographic characteristics. The interviewer-administered demographic questions include age, birth date, gender, marital status, race, Hispanic/Latino origin, Hispanic/Latino group, U.S. military status, and education level (highest grade completed). Although the interview date was not classified as a demographic variable, it is used along with a respondent's birth date (when provided by the respondent) as the starting point for determining the respondent's current age. Therefore, the editing procedures for the interview date also are included in this chapter. Though not specifically a demographic variable, a question on respondents' perceived health status is included in this interviewer-administered section at the beginning of the interview and therefore is also discussed in this chapter. In addition, editing of field interviewer (FI) debriefing variables corresponding to questions that FIs completed after the interview is discussed in this chapter. The self-administered demographic variables (including but not limited to immigrant status and employment status) are discussed in Chapter 5.

Editing was first performed on the interviewer-administered demographics variables. This editing could range from simply assigning legitimate skip codes (Section 2.4.2) to coding "OTHER, Specify" responses (Section 2.3.1.3) and resolving inconsistencies, as was the case for race and ethnicity.

After editing, the variables were processed in one of four ways:

- No imputation required: interview date, age, and gender. These are described in Sections 4.2.1, 4.2.2, and 4.2.4, respectively. The edited data contained no missing values by design.
- No imputation performed: current state residency reported by interviewers,⁵² military status,⁵³ perceived health status, and FI debriefing.⁵⁴ These are described in Sections 4.2.3, 4.2.8, 4.2.10, and 4.2.11, respectively.

⁵² The state residency that was reported by interviewers differs from the variable STATE, which contains the Federal Information Processing Standards (FIPS) codes for states and is created at the sampling stage. The variable STATE is recommended for use in state-level analyses.

⁵³ A final imputation-revised variable is not created for military service. However, a limited non-PMN imputation is performed for military service as part of the imputation procedures for health insurance, as described in Section 13.3.2.4.

⁵⁴ FIs complete the FI debriefing questions after the interview but are included in the list for completeness.

- Random assignment: birth date. This procedure is summarized in Section 4.2.5 because it is straightforward and does not involve the predictive mean neighborhood (PMN) method, which is described in Chapter 3.
- PMN: race, Hispanic/Latino origin, Hispanic/Latino group, and education level. These are described in Sections 4.2.7 and 4.2.9.

Overall, the interviewer-administered demographics variables that are discussed in this chapter either did not have missing data by design or tended to have high item response rates. Except for race, the item response rates when nonresponse was allowed (i.e., "don't know" or "refused") were more than 99 percent.⁵⁵

4.2 Editing Selected Interviewer-Administered Demographics Variables

In this section, editing procedures are described for the following:

- interview date (Section 4.2.1);
- age (Section 4.2.2);
- state residency (Section 4.2.3);
- gender (Section 4.2.4);
- birth date (Section 4.2.5);
- marital status and number of times married (if respondent had ever been married) (Section 4.2.6);
- race, Hispanic/Latino origin, and Hispanic/Latino group (Section 4.2.7);
- U.S. military service history and current military status (if respondents had ever been in the U.S. military) (Section 4.2.8);
- education level (Section 4.2.9);
- perceived health status (Section 4.2.10); and
- FI debriefing variables (Section 4.2.11).

4.2.1 Creating the Edited Interview Date Variable (INTDATE)

The program for the computer-assisted interviewing (CAI) captured information about the time and date at the start of the interview, at the start of each section to which respondents were routed, and at the completion of the interview. This information was recorded in separate variables that were included in the interview data that were transmitted from the field to the data processing staff. These variables containing the start and end dates and times of the interview and sections are referred to as time and date "stamps." For example, the time and date stamps for the beginning of the interview were captured in the variable TBEGINTR, the time and date

⁵⁵ When given the opportunity to enter a race, many respondents entered "Hispanic" or some Hispanic/Latino group such as "Mexican," resulting in a comparatively larger amount of missing data for the race question. As a result, the item response rate for the race variables is still high but tends to be about 96 to 97 percent.

stamps for the end of the interview were captured in the variable TENDINR, and the time and date stamps for the alcohol questionnaire section were captured in the variable TBEG_ALC. These time and date stamps were sequentially arranged in the dataset in the order that the questionnaire sections appeared. In most cases, the time and date stamps were compared and edited to obtain the final edited interview date (INTDATE). The editing indicator associated with the interview date variable (EIIDATE) specified the section date stamp that was used to create the edited interview date.

In some situations, a "breakoff" occurred (i.e., the respondent did not finish the interview in one sitting and resumed the interview on another date), and the respondent's birthday occurred between the beginning and the end of the interview. In these situations, the interview date was set to the end-of-interview date stamp, which was consistent with the first date stamp after the respondent's birthday. (This date stamp was indicated in the CAI.)

A date stamp was not used to set the interview date if any of the following conditions were true:

1. The date stamp was more than 14 days outside the quarter in which the interview was supposed to take place.
2. The date stamp was later in time than a subsequent date stamp.
3. The date stamp occurred before a birthday, which in turn occurred before the end of the interview.

If none of the date stamps were usable, then the interview date was imputed to be sometime in the quarter in which the interview was assumed to have taken place.

4.2.2 Creating the Edited Age Variable (AGE)

The age that respondents reported at the beginning of the interview (CALCAGE) was determined in one of two ways:

1. The respondent reported his or her birth date to the interviewer. This information was captured in the variable AGE1. The interviewer confirmed with the respondent that the birth date information had been recorded correctly and could not proceed further with the interview until the respondent verified the birth date. CALCAGE was determined from AGE1 and the date that was recorded at the start of the interview. The interviewer then confirmed with the respondent that this calculated age was correct. If the respondent indicated that the calculated age was not correct, then the interviewer could not proceed further until any additional corrections to the birth date had been made and the respondent verified both the date of birth as well as the calculated age.
2. If the respondent did not know or refused to report a birth date, the respondent was asked to provide his or her correct age, which was captured in the variable DKREFAGE. In this situation, CALCAGE was set to the age that had been recorded in DKREFAGE.

Interviews were terminated if respondents did not report their age in one of these two ways. Interviews also were terminated if CALCAGE indicated that respondents were less than 12 years old.

Even after CALCAGE has been determined for respondents at the beginning of the interview, respondents could change their age in response to consistency checks in self-administered sections pertaining to their substance use (e.g., if they reported first using a drug at an age that was greater than their current age). Any changes that respondents made to their age during the questionnaire were captured in the temporary variable CURNTAGE. Therefore, it was possible for the age that was recorded for the respondent at the beginning of the interview (CALCAGE) to be different from the value of CURNTAGE that was captured at the end of the interview and stored in the variable NEWAGE. As discussed before, NEWAGE is the final value stored in the variable CURNTAGE at the end of the interview.

The final age variable, AGE, was determined using CALCAGE and NEWAGE and three other sources: the age calculated from the final edited interview date (INTDATE) and the reported birth date (AGE1), the age corresponding to the "self" in the questionnaire household roster (if it existed) (Section 11.2.3), and the pre-interview age that was reported during screening of the dwelling unit (DU).⁵⁶ In most situations when the final edited continuous age was determined, priority was given to CALCAGE, NEWAGE, and the age that was calculated from AGE1 and INTDATE. There were occasions, however, where the age corresponding to the "self" in the household roster was used, even if it did not agree with CALCAGE and NEWAGE. If the final age (AGE) did not agree with the date of birth that had originally been entered (AGE1), the birth date also was edited. An intermediate value for age was determined in the following manner:

Intermediate value for age =

- NEWAGE, if nonmissing and exactly equal to CALCAGE, where TBEG_TUT (the interview date time stamp at the beginning of the tutorial) = INTDATE (the edited interview date) (age indicator = 1); else
- NEWAGE, if nonmissing, TBEG_TUT and INTDATE were not equal, but NEWAGE was exactly equal to CALCAGE (adjusted by Blaise⁵⁷ to a changed interview date if the interview date was changed within the questionnaire), and the respondent's birthday did not fall between the dates corresponding to TBEG_TUT and INTDATE (age indicator = 1); else
- NEWAGE, if nonmissing, TBEG_TUT and INTDATE were not equal, the respondent's birthday fell between the dates corresponding to TBEG_TUT and INTDATE, the given value of CALCAGE agreed with what it should be based on

⁵⁶ When contacting the DU, the FI asked to speak with an adult resident of the household aged 18 or older who could serve as the screening respondent. Using a handheld computer, the FI completed a 5-minute procedure with the screening respondent that involved listing all household members along with their basic demographic data (including age). The computer used the demographic data in a preprogrammed selection algorithm to select zero, one, or two sample people, depending on the composition of the household.

⁵⁷ Blaise is the computer program within the CAI instrument that was used to direct the respondent and interviewer through the questionnaire.

- INTDATE and the given birth date (i.e., EIIDATE not equal to 6), and NEWAGE and CALCAGE were exactly equal (age indicator = 1); else
- age calculated from INTDATE and the reported birth date, if the birth date was nonmissing, TBEG_TUT and INTDATE were not equal, the respondent's birthday fell between the dates corresponding to TBEG_TUT and INTDATE, and the given value of CALCAGE did not agree with what it should be based on INTDATE and the given birth date (EIIDATE = 6), where the newly calculated age based on INTDATE was exactly equal to the screener age and/or the roster age (if it existed) (age indicator = 2); else
 - NEWAGE, if NEWAGE differed from CALCAGE and NEWAGE = screener age and NEWAGE = roster age (if it existed), and the interview date at the beginning of the interview (TBEGINTR) was within the appropriate quarter (age indicator = 3); else
 - CALCAGE, if CALCAGE differed from NEWAGE and CALCAGE = screener age and CALCAGE = roster age (if it existed), and the interview date at the beginning of the interview (TBEGINTR) was within the appropriate quarter (age indicator = 4); else
 - age calculated from reported birth date and INTDATE, if EIIDATE = 5 and NEWAGE = CALCAGE (but neither was equal to the correct age) (age indicator = 5); else
 - NEWAGE, if NEWAGE differed from CALCAGE, but NEWAGE = roster age, provided roster age existed (age indicator = 6); else
 - CALCAGE, if CALCAGE differed from NEWAGE, but CALCAGE = roster age, provided roster age existed (age indicator = 7); else
 - NEWAGE, if NEWAGE differed from age calculated from reported birth date and INTDATE, but NEWAGE = CALCAGE, screener age, and roster age (if it existed) (age indicator = 8); else
 - CALCAGE, if CALCAGE differed from NEWAGE, but CALCAGE = age calculated from INTDATE and the reported birth date, and CALCAGE was within 1 year of screener age and roster age (age indicator = 9).

After these rules had been applied, this intermediate age value was compared with the age corresponding to the "self" in the household roster. For most respondents, the final edited value for the age variable (AGE) was set to this intermediate age value. There were exceptions, however, as detailed below.

By the time the interviewer reached the roster part of the questionnaire, there had been multiple opportunities to change the value of age in response to consistency checks. This value of age was called CURNTAGE by the Blaise program. One of the consistency checks in the questionnaire household roster was to verify the value of the respondent's own entry for age in the household roster (the "self" entry) against the value of CURNTAGE. If the self age differed from CURNTAGE, then the interviewer could either change the respondent's age that was entered in the roster or override the consistency check and provide an explanation for why the

roster age did not match CURNTAGE. If the consistency check for age was overridden, then the value for age corresponding to the "self" in the roster may not match the intermediate age value that was described previously. However, if sufficient explanation was given for overriding the age consistency check, other evidence pointed to the veracity of the roster age, and the difference between CURNTAGE and the roster age for self was less than 2 years, then AGE was set to the roster age, even if it disagreed with both NEWAGE and CALCAGE. In particular, all of the following conditions had to be met for this to occur:

1. The interviewer specifically indicated that the roster age was the correct one.
2. The pre-interview screener age matched the roster age.
3. The other household member's roster supported the roster age value, if another member of the household completed the interview.

It also was possible for interviewers to jump back from the household roster or elsewhere in interviewer-administered sections (i.e., after respondents had completed the audio computer-assisted self-interviewing [ACASI] section of the interview) to change the respondent's date of birth. In this situation, the values of CALCAGE and NEWAGE could be consistent, but these changes to the respondent's age between the beginning and end of the interview would not be readily apparent. Rather, this pattern would become evident in reviewing ACASI data. For example, if respondents were aged 21 or older at the beginning of the interview and they did not change their age during the ACASI portion of the interview, questions were skipped in the consumption of alcohol section about alcohol use by people aged 12 to 20. However, if their final age indicated they were aged 12 to 20, then they would have been eligible to be asked questions about alcohol use by people aged 12 to 20 (see Sections 2.4.6 and 8.5).

Three age category variables were created from the final age: CATAGE with four levels (12 to 17, 18 to 25, 26 to 34, and 35 or older); CATAG2 with three levels (12 to 17, 18 to 25, and 26 or older); and CATAG3 with five levels (12 to 17, 18 to 25, 26 to 34, 35 to 49, and 50 or older). These variables were used instead of the continuous age variables in some subsequent imputations and analyses.

4.2.3 Creating the Edited Current State Residency Variable (STATELOC)

FIs were instructed to use the FI "checkpoint" FIPE4 at the beginning of the interview to report the state where the DU was located. The term "checkpoint" refers to an item that was completed by the FI about the location of the DU or characteristics of the sample within the DU. These checkpoints were not administered to the respondents but were used to customize the wording of questions in subsequent sections of the interview, such as state-specific names for Temporary Assistance for Needy Families (TANF) programs (Section 12.2.3).

The state that interviewers entered in FIPE4 sometimes failed to match the state residence information that was used to sample a given case. These mismatches were investigated by field staff during data collection. Some of these mismatches existed for a valid reason, such as if a respondent had been selected in a DU in one state but had moved to another state. In these situations, if FIPE4 reflected the state where the respondent was currently living, the edited variable STATELOC (corresponding to FIPE4) retained the value from FIPE4. Otherwise, if the

state information in FIPE4 was entered incorrectly, STATELOC was assigned a code for bad data (Section 2.4.3). In turn, the setting of STATELOC to bad data affected the editing of other variables that were dependent on the state where the respondent was reported to be a resident.

4.2.4 Creating the Edited Gender Variable (IRSEX)

As with previous surveys since 2002, FIs in 2016 were required to enter the respondent's gender in question QD01. As a result, it was not possible to have missing values for this question. To maintain continuity with surveys prior to 2002 in which missing data could exist for gender, the variable name IRSEX was used to describe gender in the 2016 survey. However, it was not necessary to create an imputation indicator, because IRSEX and QD01 were equivalent.

As for the situation described previously for age, interviewers also could jump back to the beginning of the interview and change the respondent's gender based on roster information. This type of change to QD01 between the beginning and end of the interview would not be readily apparent. However, this pattern would become evident in reviewing ACASI data. For example, if respondents were defined as being male at the beginning of the interview, questions in the self-administered health care section about pregnancy were skipped. However, if their final gender and age indicated that they were female and aged 12 to 44, then they would have been eligible to be asked questions about pregnancy (see Sections 2.4.6 and 10.2.1).

4.2.5 Creating the Edited Birth Date Variable (BRTHDATE)

To proceed with the interview, respondents were required to provide their date of birth or current age (if they did not provide their date of birth) at the beginning of the interview. Each completed case respondent possessed a current age, although a number of cases had missing birth dates. If the birth date was nonmissing but was inconsistent with AGE and INTDATE (either in the unedited data or as a result of editing age and/or interview date), then the reported birth month and day were preserved, and the birth year was logically edited according to the interview date and age.

In cases with missing birth dates, a birth date was randomly selected from all possible birth dates, given the final age and interview date. Each date in this period (365 or 366 days, depending on whether the period includes February 29 in a leap year) had an equal probability of selection.

4.2.6 Creating the Edited Marital Status Variables (MARITAL, EDMARIT, and NOMARR)

In the 2016 questionnaire, a single question (QD07) asked about the respondent's marital status if respondents were aged 15 or older. If respondents indicated that they were currently married, widowed, or divorced or separated, they also were asked to report the number of times they had been married (QD08). The exact wording of the questions was as follows:

QD07: Are you now married, widowed, divorced or separated, or have you never married?

- 1 MARRIED
- 2 WIDOWED
- 3 DIVORCED OR SEPARATED
- 4 HAVE NEVER MARRIED

QD08: [IF QD07 = 1 OR 2 OR 3] How many times have you been married?

NUMBER OF TIMES: _____ [RANGE: 1 - 9]

MARITAL and NOMARR underwent minimal processing. Legitimate skip codes were assigned to both variables (Section 2.4.2) if respondents were aged 12 to 14. Similarly, legitimate skip codes were assigned to NOMARR if respondents had never been married. The base variable for creating an imputation-revised version of marital status was called EDMARIT. This variable was equivalent to MARITAL, with the exception that all legitimate skip codes were collapsed into a single legitimate skip code (99), and missing values were set to the SAS[®]⁵⁸ missing code (.) so that they could be properly handled by the modeling programs.

Note that these questions had been moved to the self-administered demographic section of the questionnaire in 2015 but were returned to the interviewer-administered demographic section in 2016 to improve data quality. Editing procedures remained the same across years. The variables in 2016 were determined to be comparable with the corresponding interviewer-administered variables in 2014 and prior years. Consequently, the edited variables in 2016 were given the names MARITAL and NOMARR, which were used in 2014 and earlier. However, the 2015 marital status variables in the self-administered section of the interview had different names because these variables were determined not to be comparable with the variables from 2016 and from 2014 and earlier years.

4.2.7 Creating the Edited Race and Hispanic/Latino (Origin and Group) Variables

In the 2016 questionnaire, two interviewer-administered questions focused on the respondent's ethnicity⁵⁹ (QD03 and QD04) and two focused on the respondent's race (QD05 and QD05ASIA). For those questions with multiple categories (QD04, QD05, and QD05ASIA), the respondent had the opportunity to select more than one category. Two more Hispanic/Latino group categories were added to QD04 since the 2004 survey: Dominican (from Dominican Republic) and Spanish (from Spain). These new categories were added to the survey because of the large number of "OTHER, Specify" responses in previous National Surveys on Drug Use and Health (NSDUHs) that mapped to these categories. The questions as they appear in the survey instrument are presented below.

QD03: Are you of Hispanic, Latino, or Spanish origin or descent?

- 1 YES
- 2 NO

⁵⁸ SAS[®] software is a registered trademark of SAS Institute Inc.

⁵⁹ The questions about ethnicity were limited to determining whether a respondent was Hispanic/Latino or not, and the specific Hispanic/Latino group to which a Hispanic/Latino respondent belonged.

QD04: (Asked only if QD03 = 1) Which of these Hispanic, Latino, or Spanish groups best describes you?

- 1 MEXICAN, MEXICAN AMERICAN, MEXICANO, OR CHICANO
- 2 PUERTO RICAN
- 3 CENTRAL OR SOUTH AMERICAN
- 4 CUBAN OR CUBAN AMERICAN
- 5 DOMINICAN (FROM DOMINICAN REPUBLIC)
- 6 SPANISH (FROM SPAIN)
- 7 OTHER (SPECIFY)

QD05: Which of these groups describes you?

- 1 WHITE
- 2 BLACK OR AFRICAN AMERICAN
- 3 AMERICAN INDIAN OR ALASKA NATIVE (AMERICAN INDIAN INCLUDES NORTH AMERICAN, CENTRAL AMERICAN, AND SOUTH AMERICAN INDIANS)
- 4 NATIVE HAWAIIAN
- 5 GUAMANIAN OR CHAMORRO
- 6 SAMOAN
- 7 OTHER PACIFIC ISLANDER
- 8 ASIAN (INCLUDING: ASIAN INDIAN, CHINESE, FILIPINO, JAPANESE, KOREAN, AND VIETNAMESE)
- 9 OTHER (SPECIFY)

QD05ASIA: (Asked only if QD05=8) Which of these Asian groups describes you?

- 1 ASIAN INDIAN
- 2 CHINESE
- 3 FILIPINO
- 4 JAPANESE
- 5 KOREAN
- 6 VIETNAMESE
- 7 OTHER (SPECIFY)

As stated in the guidelines from the Office of Management and Budget (OMB),⁶⁰ "Hispanic/Latino" was categorized as an ethnicity, not a race. However, when given the opportunity to enter a race, many respondents entered "Hispanic" or some Hispanic/Latino group, resulting in missing data for the race question. Even though the final drug use tables were cross-classified with a variable that combined race and ethnicity, separate variables were initially created for race and ethnicity, and the race/ethnicity variables used in the tables were derived from these separate variables.

⁶⁰ In October 1997, the OMB released a notice, "Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity" (OMB, 1997) that provides new standards for maintaining, collecting, and presenting Federal data on race and ethnicity.

Due to the relationship between Hispanicity and race reporting, Hispanicity was used in the editing of race, and vice versa. In the process of editing race, the "OTHER, Specify" response to the Hispanic/Latino group question (QD04) was consulted (if it existed) if no race information was identified in QD05 or QD05ASIA. Similarly, in the process of editing the Hispanic/Latino group, the "OTHER, Specify" responses to the race questions (QD05 and QD05ASIA) were consulted (if they existed) if no Hispanic/Latino group information was identified in QD04. Because of the interdependence of race and Hispanicity, the editing of these variables is discussed together in this section.

The procedures used to edit the race and Hispanicity variables in the surveys since 2008 differed in several ways from the procedures used in previous surveys. One of the major differences was in the handling of race for multiple-race respondents. The first procedural changes were triggered by the elimination of the QD06 question, which appeared in the survey from 1999 to 2002. QD06 asked respondents who selected more than one racial category from QD05 and QD05ASIA combined to choose the race with which they identified the most. Without this question, it was impossible to determine (directly) the single race that a given multiple-race respondent would most closely identify for himself or herself. In the 2003-2007 surveys, QD06 responses were "simulated" based on models built using true QD06 responses from the 2000-2002 surveys.⁶¹ However, because racial demographics in the United States had changed since the 2000 survey and because recent data that were needed to update these models were not available, this method was not used after 2008 and single races were not assigned for multiple-race respondents. Refer to Section 3.3 of the 2008 imputation report (Ault et al., 2010) for more details.

4.2.7.1 Categories Used in Race and Hispanic/Latino Variables

4.2.7.1.1 Race Categories

For editing purposes, the 7 specific categories in QD05 (white, black or African American, American Indian or Alaska Native, Native Hawaiian, Guamanian or Chamorro, Samoan, and Other Pacific Islander) and the 6 specific categories in QD05ASIA (Asian Indian, Chinese, Filipino, Japanese, Korean, and Vietnamese) were combined to produce 13 racial categories. Two other categories also were created: "Other Asian" (where the responses to QD5ASIA did not fit into the above category) and "Asian nonspecific" (where no response was selected to QD05ASIA, even though Asian was selected in QD05). Respondents could choose almost any subset of these categories. The only subsets that were not logically possible were those that included "Asian nonspecific" in combination with one or more specific Asian categories. Combining the information from QD05 and QD05ASIA, as well as QD04 when necessary, allowed for the creation of all the edited and imputation-revised race variables.

⁶¹ Because of the questionnaire differences between the 1999 survey and the 2000-2002 surveys, the procedure for simulating QD06 responses for the 2003-2007 surveys was made simpler by limiting the QD06 results from the 2000-2002 surveys. During the 2003-2007 surveys, for the purpose of allocating respondents into imputation classes, a model was used to select a single race for respondents who had selected more than one race (IRRACE2). The selection of a single race was based on models that were fit using data from the 2000-2002 surveys. This method is described in Appendix E of the 2007 imputation report (Ault et al., 2009).

4.2.7.1.2 *Hispanic/Latino Categories*

With the addition of two Hispanic/Latino categories since the 2004 survey, respondents were given the choice of seven categories in QD04 (Mexican/Mexican American/Mexicano/Chicano, Puerto Rican, Central or South American, Cuban/Cuban American, Dominican (from Dominican Republic), Spanish (from Spain), or some other Hispanic/Latino group),⁶² and they could choose more than one category. As with QD05, interviewers could manually enter the alternative to the choices given, which would be either coded to some subset of the existing seven categories or set to missing. The "OTHER, Specify" responses to QD05 and/or QD05ASIA, if nonmissing, were consulted if no Hispanic/Latino origin group information was available from QD04. The final imputation-revised Hispanic/Latino group variable, IRHOGP4, included all seven Hispanic/Latino group levels and a legitimate skip code (99) for respondents who were not Hispanic/Latino.

4.2.7.2 *Classification of "OTHER, Specify" Codes*

All "OTHER, Specify" responses from QD04, QD05, and QD05ASIA were assigned both a race code and a Hispanic/Latino code. Each of these codes was mapped to at least one of the categories described in Section 4.2.7 and in this section, or to some other code that was informative in the final imputation described in Sections 4.3.2 and 4.3.3. A summary of categories of "OTHER, Specify" codes and how they were handled is given in the following sections. Appendix C provides the individual "OTHER, Specify" codes and more details about how they were handled.

4.2.7.2.1 *Mapping of Race "OTHER, Specify" Codes to Edited Values*

This section describes the directly and indirectly mapped race codes. The edits following from either of these types of mapped codes resulted in values that were considered "final" in that no imputation was necessary for them.

The directly mapped codes were mapped to one or more of the categories given in the questionnaire (Section 4.2.7). There were directly mapped racial category codes and directly mapped geographic category codes. Racial category codes were exactly equivalent to one or more categories in QD05 or QD05ASIA, and were mapped directly to those categories regardless of whether the write-in response was in QD05 or QD05ASIA. (Respondents were still considered at least part Asian, even if the write-in response in QD05ASIA was non-Asian. The racial makeup of a respondent who entered a non-Asian racial category in QD05ASIA was determined on a case-by-case basis.) For example, a response such as "Han" mapped directly to a category in QD05ASIA ("Chinese"), and a response such as "mestizo" mapped directly to two categories in QD05, "white" and "Native American."

By contrast, geographic category codes corresponding to a country where census data indicated a racially homogeneous society depended on the corresponding question. For example, an entry of "Polish" in QD05 mapped to white because the Polish census data indicated nearly all Poles were white. On the other hand, an entry of "Polish" in the QD05ASIA "OTHER, Specify"

⁶² When listing the six Hispanic/Latino defined categories in QD04, they shall henceforth be listed in this chapter as Mexican, Puerto Rican, Central or South American, Cuban, Dominican, and Spanish.

mapped to "Other Asian." Geographic category codes also included ethnic groups where the racial identification was not immediately obvious. For example, a response of "Arab" would be automatically mapped to "white" if the response was a write-in response for QD05. However, as with the "Polish" entry, if the "Arab" response was a write-in response in QD05ASIA, the respondent was considered "Other Asian."

Indirect mapping was used for countries that were racially heterogeneous. A racial category was chosen by generating a random number and allocating the race based on a comparison of the random number with the proportions of races in the country's census.⁶³ For example, an entry of "Bolivian" would have a 55 percent chance of being allocated to the American Indian/Alaska Native category, because the latest Bolivian census indicated 55 percent of Bolivians were American Indian/Alaska Native. For countries where the census indicated a small proportion of some indistinct category such as "other" and the randomly generated number indicated an allocation to this proportion, the final race was left to imputation (appropriately constrained based upon the indistinct response).

If two or three heterogeneous countries were entered in the "OTHER, Specify" response (e.g., Bolivian and Peruvian), the final race was allocated using the following procedure: (1) randomly assign races based on the proportions for each country mentioned; and (2) combine the results. Exceptions to these rules occurred with the categories Mexicans, Puerto Ricans, Cubans, Dominicans, Central or South Americans (no country listed), and Spanish, which were given codes described under the next heading, with a final value determined using the formal imputation procedures described in Section 4.3.1. Starting with the 2006 survey, the imputation processing of indirectly mapped codes obtained from QD05ASIA has been simplified. In prior survey years, this type of write-in response was mapped to a race through country census information; since the 2006 NSDUH, all census-based write-in responses to the Asian race question were mapped directly to the "Other Asian" racial category.

4.2.7.2.2 Mapping of Race "OTHER, Specify" Codes to Inform Imputation

"OTHER, Specify" responses that could not be mapped definitively to a specific race category resulted in incomplete values requiring imputation. These responses were assigned two types of codes, either informative or noninformative, for the formal imputation procedures for race described in Section 4.3.1.

Responses that provided information were used to limit the final imputation. For example, a response of "mixed" resulted in an imputation among donors with two or more races, and a response of "brown" resulted in an imputation among donors who were not single race white.

⁶³ See <https://www.cia.gov/library/publications/the-world-factbook/fields/2075.html> (Ethnicity and Race by Countries) for more information.

A noninformative response (e.g., American) that was not accompanied by a response to one of the precoded categories⁶⁴ in QD05 or QD05ASIA (i.e., those other than the "OTHER, Specify" response) resulted in an unrestricted imputation.

4.2.7.2.3 Subsequent Editing of Race "OTHER, Specify" Codes

Subsequent to the initial mapping of the race "OTHER, Specify" codes, edits were sometimes implemented that revised or clarified the initial mapping before final races were allocated. These edits were necessary if multiple sources of information, including "OTHER, Specify" responses, provided conflicting or confusing information. These edits were implemented when (1) the final mapping depended upon the source question (i.e., QD04, QD05, and QD05ASIA); (2) the responses were given to both the "OTHER, Specify" and precoded categories of QD05 or QD05ASIA; or (3) the different "OTHER, Specify" responses were present in at least two of QD04, QD05, and QD05ASIA. In some cases, it was necessary to individually examine the responses to determine the appropriate mapping.

Occasionally, the final mapped value depended upon whether the "OTHER, Specify" code was in QD04, QD05, or QD05ASIA. An example from directly mapped codes is "Indian." This response would be mapped to "American Indian/Alaska Native" if the "OTHER, Specify" response was in QD05, but it would be mapped to "Asian Indian" if the "OTHER, Specify" response was in QD05ASIA. Indirectly mapped codes also could depend upon the source question. The census data from many countries included Asian categories. If the "OTHER, Specify" response was in QD05ASIA, the random imputation to a census category was limited to the Asian categories. "OTHER, Specify" responses that were not specifically Asian sometimes occurred in the "OTHER, Specify" category of QD05ASIA. These were carefully examined, but the "Asian" part of the response was always preserved.

If "OTHER, Specify" responses to QD05 or QD05ASIA accompanied responses to the precoded categories in QD05 and QD05ASIA, it was necessary to reconcile these responses. In some cases, the combination of responses mapped to one of the multiple race categories. For example, if a respondent selected "black/African American" in QD05 and wrote in "black and American Indian," then the respondent would be assigned both racial categories "black/African American" and "American Indian/Alaska Native."

There were instances, however, when the "OTHER, Specify" response was ignored because of responses to the precoded categories in QD05 and QD05ASIA. In particular, the "OTHER, Specify" response was always ignored if a precoded response category was selected, and the "OTHER, Specify" response was a geographic category code.⁶⁵ For example, if the interviewer selected the category for "black/African American" for the respondent and also wrote in "Polish," it was assumed that the respondent was a black Pole and, for racial identification purposes, was considered single-race black/African American. This was true even though the

⁶⁴ For example, precoded categories in QD05 in 2016 were 1 = White; 2 = Black or African American; 3 = American Indian or Alaska Native; 4 = Native Hawaiian; 5 = Guamanian or Chamorro; 6 = Samoan; 7 = Other Pacific Islander; and 8 = Asian. Category 9 (Other) led to the "OTHER, Specify" question.

⁶⁵ Actually, this "edit" was not "subsequent" to the initial mapping. Instead, the initial mapping was ignored under the circumstances described.

Polish census did not identify significant numbers of people with a race other than white in the Polish population.

In some instances, it was necessary to reconcile the "OTHER, Specify" responses to QD04, QD05, and QD05ASIA. In these cases, the responses were examined on an individual basis, and sometimes a new code was assigned that more accurately reflected the situation.

4.2.7.2.4 Mapping of Hispanic/Latino "OTHER, Specify" Codes

Certain Hispanic/Latino codes were considered "Definitely Hispanic." If any of these appeared in QD05 or QD05ASIA, the respondent was considered Hispanic/Latino regardless of the response to QD03. Examples included "Hispanic" and "Dominicano" (Spanish for "Dominican"). There was also a code to handle respondents who were definitely not Hispanic/Latino (i.e., the respondent reported "Not Hispanic/Latino"). If this code appeared in QD04, QD05, or QD05ASIA, then the respondent was considered non-Hispanic/Latino regardless of the response to QD03. All other Hispanic/Latino codes either mapped directly to one or more of the seven Hispanic/Latino group categories or provided no new information (e.g., Hispanic).

4.2.7.3 Edited Race Variables

4.2.7.3.1 Individual Race Categories (EDQD051-EDQD0515)

Edited variables were created that correspond to the 15 racial categories described in Section 4.2.7.1.1. These variables were called EDQD05xx, where xx represented a number between 1 and 15, corresponding to each of the 15 categories.

EDQD05xx =

- 1, if the level xx was selected by the respondent in QD05 or QD05ASIA; else
- 2, if the level xx was indicated by a directly mapped code in QD05 or QD05ASIA; else
- 3, if no EDQD05xx variables had values of 1 or 2, and the level xx was indicated by a directly mapped code in QD04 (Hispanic/Latino status); else
- 4, if (a) no EDQD05xx variables had values of 1, 2, or 3, and (b) the level xx was indicated by an indirectly mapped code in QD04, QD05, and/or QD05ASIA; else
- missing.

EDQD0515 (Asian nonspecific) was a little different from the others. In particular, there was no specific level of QD05 or QD05ASIA that corresponded to it. It was used mainly to preserve a response of "Asian" to QD05, even if the respondent selected nothing in QD05ASIA. The value of EDQD0515 was set to 1 if the respondent selected "Asian" in QD05 but mentioned

nothing that mapped to a specific Asian category in QD05ASIA. It also could have values of 2, 3, or 4, depending on the "OTHER, Specify" codes.⁶⁶

4.2.7.3.2 Broad Categories of Race (EDRACE)

EDRACE is a 25-level variable that indicates which of four broad racial categories (white, black/African American, American Indian/Alaska Native, Asian/Other Pacific Islander) were identified in QD04, QD05, and QD05ASIA, and it also has levels to indicate how the imputation should be restricted based on the race of the donor. The first three broad racial categories corresponded to EDQD051, EDQD052, and EDQD053, respectively. "Asian/Other Pacific Islander" was considered to have been identified if any of EDQD054 through EDQD0515 was nonmissing. EDRACE was created using the following rules, under five possible scenarios:

Scenario 1: If only one broad racial category was identified in QD04, QD05, and/or QD05ASIA, EDRACE =

- 1 (white only), if EDQD051 was nonmissing; else
- 2 (black/African American only), if EDQD052 was nonmissing; else
- 3 (American Indian/Alaska Native only), if EDQD053 was nonmissing; else
- 4 (Asian/Other Pacific Islander only), if any of EDQD054 through EDQD0515 were nonmissing.

Scenario 2: If two broad racial categories were identified in QD04, QD05, and/or QD05ASIA, EDRACE =

- 5 (white and black/African American only), if both EDQD051 and EDQD052 were nonmissing; else
- 6 (white and American Indian/Alaska Native only), if both EDQD051 and EDQD053 were nonmissing; else
- 7 (white and Asian/Other Pacific Islander only), if EDQD051 was nonmissing and at least one of EDQD054 through EDQD0515 were nonmissing; else
- 8 (black/African American and American Indian/Alaska Native only), if both EDQD052 and EDQD053 were nonmissing; else
- 9 (black/African American and Asian/Other Pacific Islander only), if EDQD052 was nonmissing and at least one of EDQD054 through EDQD0515 were nonmissing; else
- 10 (American Indian/Alaska Native and Asian/Other Pacific Islander only), if EDQD053 was nonmissing and at least one of EDQD054 through EDQD0515 were nonmissing.

⁶⁶ A value of 2 indicated that the respondent wrote "Asian" in the QD05 other-specify blank. A value of 3 indicated that the response was obtained from the other-specify part of the Hispanic/Latino group question (QD04). Finally, a value of 4 indicated that the respondent gave a country of origin as a response to QD05, and the census for that country had "Asian" as one of its categories.

Scenario 3: If three broad racial categories were identified in QD04, QD05, and/or QD05ASIA, EDRACE =

- 11 (white, black/African American, and American Indian/Alaska Native only), if all of EDQD051 through EDQD053 were nonmissing; else
- 12 (white, black/African American, and Asian/Other Pacific Islander only), if both EDQD051 and EDQD052 were nonmissing and at least one of EDQD054 through EDQD0515 were nonmissing; else
- 13 (white, American Indian/Alaska Native, and Asian/Other Pacific Islander only), if both EDQD051 and EDQD053 were nonmissing and at least one of EDQD054 through EDQD0515 were nonmissing; else
- 14 (black/African American, American Indian/Alaska Native, and Asian/Other Pacific Islander only), if both EDQD052 and EDQD053 were nonmissing and at least one of EDQD054 through EDQD0515 were nonmissing.

Scenario 4: If all four broad racial categories were identified in QD04, QD05, and/or QD05ASIA, EDRACE = 15.

Scenario 5: If none of the broad racial categories were identified in QD04, QD05, and/or QD05ASIA, EDRACE =

- 16 (multiple race, no other information), if an "OTHER, Specify" response such as "biracial" or "mixed" appeared in QD04, QD05, or QD05ASIA; else
- 17 (indicated race other than white, no other information), if an "OTHER, Specify" response such as "nonwhite," "brown," "tan," or similar responses in Spanish appeared in QD04, QD05, or QD05ASIA; else
- 18 (white, or both white and American Indian/Alaska Native), if the random assignment of a census data code resulted in imputation restricted to donors who were either white, or both white and American Indian/Alaska Native; else
- 19 (not American Indian/Alaska Native, in part or in full), if the random assignment of a census data code resulted in imputation restricted to donors who were not American Indian/Alaska Native, in part or in full; else
- 20 (non-Hispanic Mexican), if "Mexican" was mentioned in the QD05 and/or QD05ASIA "OTHER, Specify" responses, but QD03 = 2; else
- 21 (non-Hispanic Cuban), if "Cuban" was mentioned in the QD05 and/or QD05ASIA "OTHER, Specify" responses, but QD03 = 2; else
- 22 (non-Hispanic Central or South American), if "Central or South American" was mentioned in the QD05 and/or QD05ASIA "OTHER, Specify" responses, but QD03 = 2; else
- 23 (non-Hispanic Dominican), if "Dominican" was mentioned in the QD05 and/or QD05ASIA "OTHER, Specify" responses, but QD03 = 2; else
- 24 (non-Hispanic Spanish), if "Spanish" was mentioned in the QD05 and/or QD05ASIA "OTHER, Specify" responses, but QD03 = 2; else

- 25 (non-Hispanic Puerto Rican), if "Puerto Rican" was mentioned in the QD05 and/or QD05ASIA "OTHER, Specify" responses, but QD03 = 2; else
- missing.

4.2.7.3.3 *Finer Categories of Race (EDNWRACE)*

EDNWRACE was a 16-level edited variable used as a base variable for the imputation-revised finer racial category variable IRNWRACE. Level 16 of EDNWRACE indicated when the imputation should be restricted to Asian-specific categories. It was created using the following rules, under three possible scenarios:

Scenario 1: If only one of EDQD051 through EDQD0515 was nonmissing,

EDNWRACE =

- 16 (Asian nonspecific only), if EDQD0515 was the nonmissing variable; else
- xx (one known racial category only), where EDQD05xx was the nonmissing variable out of EDQD051 through EDQD0514.

Scenario 2: If more than one of EDQD051 through EDQD0515 was nonmissing,

EDNWRACE =

- 13 (Native Hawaiian and Other Pacific Islander only where Other Pacific Islander includes Guamanian/Chamorro and Samoan), if both EDQD054 and one or more of EDQD055, EDQD056, or EDQD057 were nonmissing, and all other EDQD05xx variables were missing; else
- 14 (Asian multiple category), if all of EDQD051 through EDQD057 were missing (i.e., at least two of the ordinary Asian categories were selected); else
- 15 (more than one race).

Scenario 3: If all of EDQD051 through EDQD0515 were missing,

EDNWRACE =

- 15 (more than one race), if EDRACE = 16; else
- missing.

4.2.7.4 *Edited Hispanic/Latino Variables*

4.2.7.4.1 *Hispanic/Latino Indicator (EDHOIND)*

An edited Hispanic/Latino indicator, EDHOIND, was created using responses to QD03 and, in rare cases, the "OTHER, Specify" responses to QD04, QD05, and/or QD05ASIA. This indicator variable was created as follows:

EDHOIND =

- 1 (Hispanic/Latino), if QD03 = 1 and no "OTHER, Specify" response stated that the respondent was definitely not Hispanic/Latino, or if the "OTHER, Specify" response to QD05 or QD05ASIA indicated that the respondent was definitely Hispanic/Latino; else
- 2 (not Hispanic/Latino), if QD03 = 2 and no "OTHER, Specify" response stated that the respondent was definitely Hispanic/Latino, or if the "OTHER, Specify" response to QD04, QD05, and/or QD05ASIA indicated that the respondent was definitely not Hispanic/Latino; else
- missing.

The race "OTHER, Specify" responses, which were considered "definitely Hispanic/Latino," and the single Hispanic/Latino "OTHER, Specify" response, which was considered "definitely not Hispanic/Latino," are listed in Appendix C.

4.2.7.4.2 Individual Hispanic/Latino Group Categories (EDQD041-EDQD047)

The edited variables EDQD041 through EDQD047 were created to match the seven Hispanic/Latino group categories described in Section 4.2.7.1.2: Mexican, Puerto Rican, Central or South American, Cuban, Dominican, Spanish, and Other Hispanic/Latino.

EDQD04xx =

- 1, if the level xx was selected by the respondent in QD04; else
- 2, if the "OTHER, Specify" response from QD04 mapped directly to level xx; else
- 3, if no EDQD04xx variables had values of 1 or 2, and the "OTHER, Specify" response from QD05 or QD05ASIA mapped directly to level xx; else
- missing.

4.2.7.4.3 Edited Hispanic/Latino Group (EDHOGRP)

The edited variable EDHOGRP was the base variable for creating an imputation-revised Hispanic/Latino group variable. It had seven levels to match the seven Hispanic/Latino group categories described in Section 4.2.7.1.2, plus several other more general Hispanic/Latino levels that could be used in a restricted imputation. Those respondents with EDHOIND = 2 were assigned EDHOGRP = 99. It was created using the following rules, under four possible scenarios:

Scenario 1: If EDHOIND = 2,

EDHOGRP = 99.

Scenario 2: If EDHOIND = 1 or missing and only one of EDQD041 through EDQD047 was nonmissing,

EDHOGRP = xx , where EDQD04 xx was the nonmissing one.

Scenario 3: If EDHOIND = 1 or missing and more than one of EDQD041 through EDQD047 was nonmissing,

EDHOGRP =

- 1 (Mexican), if EDQD041 was nonmissing; else
- 2 (Puerto Rican), if EDQD042 was nonmissing; else
- 3 (Central or South American), if EDQD043 was nonmissing; else
- 4 (Cuban), if EDQD044 was nonmissing; else
- 5 (Dominican), if EDQD045 was nonmissing; else
- 6 (Spanish), if EDQD046 was nonmissing; else
- 7 (Other), if EDQD047 was nonmissing.

For the multiple Hispanic/Latino group respondents, a priority rule similar to the one used in the surveys prior to 2004 was applied in determining a single Hispanic/Latino group. The only difference is the addition of two more Hispanic/Latino group categories since the 2004 survey, resulting in the following order: Mexican, Cuban, Puerto Rican, Central or South American, Dominican, Spanish, and Other Hispanic/Latino.

Scenario 4: If EDHOIND = 1 or missing and all of EDQD041 through EDQD047 were missing,

EDHOGRP =

- EDRACE + 7 (imputation restricted by race), if $1 \leq \text{EDRACE} \leq 14$; else
- missing.

4.2.8 Creating the Edited Military Service History and Military Veteran Status Variables

4.2.8.1 Current Service Variables (SERVICE and MILSTAT)

SERVICE and MILSTAT were created from QD09 and QD10, respectively, which asked respondents aged 17 or older about their current and past service in the U.S. military. The questions as they appear in the survey instrument are presented below.

QD09: Have you ever been in the United States Armed Forces?

- 1 YES
- 2 NO

QD10: Are you **currently** on **active** duty in the United States Armed Forces, are you in a Reserve component, or are you now separated or retired from the military?

- 1 ON ACTIVE DUTY IN THE ARMED FORCES
- 2 IN A RESERVES COMPONENT
- 3 NOW SEPARATED OR RETIRED FROM THE MILITARY

Respondents who were currently on active duty in the U.S. military were not eligible to be interviewed for NSDUH. If respondents reported in QD10 that they were currently on active military duty, the interviewers were asked to confirm this answer with the respondents. The interview was terminated if respondents confirmed that they were on active duty in the U.S. military. Consequently, there were no final respondents in the final NSDUH data who reported that they currently were on active military duty. However, some final respondents could be civilians who were currently in the military reserves or were separated or retired from the military.

Unlike the situation in most places in the interview (Section 2.4.2), responses of "don't know" or "refused" to the question about lifetime military service were treated as potential indications of military service. Thus, respondents who did not know or refused to report whether they had ever been in the U.S. military also were asked QD10 in order to determine their eligibility status for the interview.

Legitimate skip codes were assigned to SERVICE and MILSTAT if respondents were aged 12 to 16 (see Section 2.4.2). In addition, legitimate skip codes were assigned to MILSTAT if respondents were aged 17 or older but they reported in QD09 that they had never been in the U.S. armed forces.

Two questions about respondents' current service in a military reserve component have been included in the interviewer-administered demographics section since 2013, as shown below.

V2A Are you a member of a Reserve component currently serving full-time in an active duty status?

- 1 YES
- 2 NO
- DK/REF

V2B Are you currently serving full-time in a Reserve component? Full-time service does not include annual training for the Reserves or National Guard.

- 1 YES
- 2 NO
- DK/REF

Question V2A was administered if respondents reported in question QD10 that they were currently on active duty in the U.S. armed forces. Because respondents who reported in question QD10 and confirmed on follow-up that they were currently on active duty were not eligible to be

interviewed for NSDUH, no final NSDUH respondents had data for question V2A. Consequently, an edited variable was not created that corresponded to this question.

Question V2B was administered if respondents reported in question QD10 that they were serving in a reserve component. The edited variable MILRSRV corresponded to this question. As for SERVICE and MILSTAT, legitimate skip codes were assigned to MILRSRV if respondents were aged 12 to 16. Legitimate skip codes also were assigned to MILRSRV if respondents were aged 17 or older but they reported in QD09 that they had never been in the U.S. armed forces or they reported in QD10 that they were now separated or retired from the reserves or active duty.

4.2.8.2 Military Veteran Status Variables

Questions also have been included in the interviewer-administered demographics section since 2013 for respondents who had ever served in the U.S. armed forces regarding whether they ever served on active duty, the period(s) when they served, and whether they served in a military combat zone where they drew imminent danger pay or hostile fire pay. These questions are shown below.

QD10A Have you ever served on **active duty** in the United States Armed Forces or Reserve components?

Active duty does not include training for the Reserves or National Guard, but **does** include activation, for example, for a national emergency or military conflict.

- 1 YES
- 2 NO
- DK/REF

QD10B1 When did you serve on active duty in the United States Armed Forces or Reserve components? Just give me the number or numbers from the card.

TO SELECT MORE THAN ONE CATEGORY, PRESS THE SPACE BAR BETWEEN EACH CATEGORY YOU SELECT.

- 1 SEPTEMBER 2001 OR LATER
- 2 AUGUST 1990 TO AUGUST 2001 (INCLUDING PERSIAN GULF WAR)
- 3 MAY 1975 TO JULY 1990
- 4 VIETNAM ERA (MARCH 1961 TO APRIL 1975)
- 5 FEBRUARY 1955 TO FEBRUARY 1961
- 6 KOREAN WAR (JULY 1950 TO JANUARY 1955)
- 7 JANUARY 1947 TO JUNE 1950
- 8 WORLD WAR II (DECEMBER 1941 TO DECEMBER 1946)
- 9 NOVEMBER 1941 OR EARLIER

QD10C Did you ever serve on active duty in the United States Armed Forces or Reserve components in a military combat zone or an area where you drew imminent danger pay or hostile fire pay?

- 1 Yes
- 2 No
- DK/REF

In addition, some response categories in question QD10B1 were revised in 2014. In the 2013 survey, response category 4 for the Vietnam era indicated that this period was from August 1964 to April 1975, and category 5 indicated that the period between the Korean and Vietnam conflicts was from February 1955 to July 1964. The category since 2014 for the Vietnam era shows this period as starting in March 1961 (instead of August 1964), and the period between the Korean and Vietnam conflicts shows this period as ending in February 1961. Despite this change, the edited variable for service during the Vietnam era retained the name ACTDVIET that had been created for 2013. However, the variable since 2014 for service between the Korean and Vietnam conflicts has used the name ACTDB4VN (where B4VN = before Vietnam) because the previous name for the variable in 2013 (ACTD5564) defined a period of military service (i.e., from 1955 to 1964) that has no longer been applicable since 2014.⁶⁷

The edited variables corresponding to questions QD10A through QD10C were assigned legitimate skip codes if respondents were aged 12 to 16 or respondents indicated in question QD09 (SERVICE) that they had never served in the U.S. armed forces. Similarly, the edited variables for QD10B1 (ACTD2001 through ACTD1941) and QD10C (COMBATPY) were assigned legitimate skip codes if question QD10A (edited variable ACTDEVER) was answered as "no."

As noted in Section 2.4.4, the response categories for the individual response categories in the "enter all that apply" question QD10B1 differed from those that were used for similar variables prior to 2013. Documentation for these values was as follows:

- 1 = Yes, and
- 2 = No (not entered).

In addition, the editing procedures included identifying inconsistencies between respondents' ages and reports of being on active duty in a particular period based on respondents' birth dates and the particular period when they reported being on active duty. Variables for the specific period of active duty were assigned codes for bad data (Section 2.4.3) if respondents' answers would lead to the conclusion that the respondents would have been younger than 16 or older than 65 during the entire period when they reported serving on active duty. These edits were performed based on calculated upper and lower bounds for the possible ages when

⁶⁷ The 2013 variable ACT5564 also was subsequently renamed to ACTDB4VN so that data for this variable from 2013 could be included on data files that also contained the corresponding variable from 2014 and future years.

respondents could have been in the military during a given period of military service, based on their birthdates and the boundaries of the service periods in the questionnaire.

No editing was done if respondents would have been aged 16 to 65 at some point during the period for which they reported service. These edits resulted in data on periods of active military duty being set to bad data for a small number of respondents (fewer than 15) in 2016, including some respondents who reported serving on active duty in a period before they may have even been born.

4.2.9 Creating the Edited Highest Grade Completed Variables (EDUHIGHST and EDEDUC)

EDUHIGHST and EDEDUC were created using the responses to the interviewer-administered education question QD11, which asked about the highest grade in school that the respondent had completed. The question from the survey instrument appears below:

QD11: What is the highest grade or year of school you have **completed**?

- 0 NO SCHOOLING COMPLETED
- 1 1ST GRADE COMPLETED
- 2 2ND GRADE COMPLETED
- 3 3RD GRADE COMPLETED
- 4 4TH GRADE COMPLETED
- 5 5TH GRADE COMPLETED
- 6 6TH GRADE COMPLETED
- 7 7TH GRADE COMPLETED
- 8 8TH GRADE COMPLETED
- 9 9TH GRADE COMPLETED
- 10 10TH GRADE COMPLETED
- 11 11TH GRADE COMPLETED
- 12 REGULAR HIGH SCHOOL DIPLOMA
- 13 12TH GRADE, NO DIPLOMA
- 14 GED CERTIFICATE OF HIGH SCHOOL COMPLETION
- 15 SOME COLLEGE CREDIT, BUT NO DEGREE
- 16 ASSOCIATE'S DEGREE (FOR EXAMPLE, AA, AS)
- 17 BACHELOR'S DEGREE (FOR EXAMPLE, BA, BS)
- 18 MASTER'S DEGREE (FOR EXAMPLE, MA, MS, MENG, M. ED, MSW, MBA)
- 19 DOCTORATE DEGREE (FOR EXAMPLE, PHD, EDD)
- 20 PROFESSIONAL DEGREE BEYOND A BACHELOR'S DEGREE (FOR EXAMPLE, MD, DDS, DVM, LLB, JD)

Since 2015, the CAI program compared respondents' highest level of education from QD11 and their current grade or year in school (if applicable) from question QD18, which is a self-administered question. A series of consistency checks (QD18CC01 through QD18CC04) were triggered based on certain answers in QD11 and QD18. QD18 and the associated consistency checks between QD11 and QD18 are shown below.

QD18 [IF QD17 = 1] What grade or year of school are you **now** attending?

[IF QD17b = 1] What grade or year of school will you be attending when your vacation is over?

- 1 1st grade
- 2 2nd grade
- 3 3rd grade
- 4 4th grade
- 5 5th grade
- 6 6th grade
- 7 7th grade
- 8 8th grade
- 9 9th grade
- 10 10th grade
- 11 11th grade
- 12 12th grade
- 13 College or university/1st year
- 14 College or university/2nd year
- 15 College or university/3rd year
- 16 College or university/4th year
- 17 College or university/5th or higher year
- DK/REF

QD18CC01 [IF (QD11=17 AND QD18<=16 AND QD17=1) OR (QD11=13 AND QD18<=12 AND QD17=1) OR (QD11=14 AND QD18<=12 AND QD17=1)] The computer recorded that you are now attending [QD18]. Is this correct?

[IF (QD11=17 AND QD18<=16 AND QD17b=1) OR (QD11=13 AND QD18<=12 AND QD17b=1) OR (QD11=14 AND QD18<=12 AND QD17b=1)] The computer recorded that you will be attending [QD18] when your vacation is over. Is this correct?

- 4 Yes
- 6 No
- DK/REF

QD18CC02 [IF QD18CC01=4] The answers for the last question and an earlier question disagree. Which answer is correct?

- 1 *[IF QD17=1] I am now attending [QD18] [IF QD17b=1] I will be attending [QD18] when my vacation is over*
- 2 *The highest grade or degree that I completed is [QD11]*
- 3 *Neither answer is correct*
- DK/REF

*QD18CC03 [IF QD18CC02=2 OR QD18CC02=3 OR QD18CC01=6] Please answer this question again. [IF QD17 = 1] What grade or year of school are you **now** attending?*

[IF QD17b = 1] What grade or year of school will you be attending when your vacation is over?

- 1 1st grade*
- 2 2nd grade*
- 3 3rd grade*
- 4 4th grade*
- 5 5th grade*
- 6 6th grade*
- 7 7th grade*
- 8 8th grade*
- 9 9th grade*
- 10 10th grade*
- 11 11th grade*
- 12 12th grade*
- 13 College or university/1st year*
- 14 College or university/2nd year*
- 15 College or university/3rd year*
- 16 College or university/4th year*
- 17 College or university/5th or higher year*
- DK/REF*

QD18CC04 [IF QD18CC02=1 OR QD18CC02=3] Please answer this question again. What is the highest grade or year of school you have completed?

- 0 No schooling completed*
- 1 1st grade completed*
- 2 2nd grade completed*
- 3 3rd grade completed*
- 4 4th grade completed*
- 5 5th grade completed*
- 6 6th grade completed*
- 7 7th grade completed*
- 8 8th grade completed*
- 9 9th grade completed*
- 10 10th grade completed*
- 11 11th grade completed*
- 12 Regular high school diploma*
- 13 12th grade, NO DIPLOMA*
- 14 GED certificate of high school completion*
- 15 Some college credit, but no degree*
- 16 Associate's degree (for example, AA, AS)*
- 17 Bachelor's degree (for example, BA, BS)*
- 18 Master's degree (for example, MA, MS, MEng, M. Ed, MSW, MBA)*

- 19 *Doctorate degree (for example, PhD, EdD)*
- 20 *Professional degree beyond a bachelor's degree (for example, MD, DDS, DVM, LLB, JD)*
- DK/REF

Consequently, editing to create EDUHIGHST took into account data from QD18CC04. Specifically, if respondents changed their answer for their highest level of education in QD18CC04 to "don't know" or "refused," then the original answer from QD11 was replaced with the relevant code of 94 or 97, respectively, in EDUHIGHST (Section 2.2.1). Similarly, if QD18CC04 had a value from 0 to 20, then the default procedure was to assign the response from QD18CC04 to EDUHIGHST. In particular, no editing was done to EDUHIGHST if respondents reported in QD18 that they were currently in the 12th grade, confirmed that their current grade was correct, and reported that they previously completed the 12th grade but did not receive a diploma (QD18CC04 = 13). In this situation, it was assumed that the respondents were repeating the 12th grade.

However, EDUHIGHST was assigned a code of 85 for bad data (Section 2.3) based on the conditions that are described below when respondents confirmed that their current grade in QD18 was the correct answer (QD18CC02 = 1). Similar edits were implemented if respondents said that neither their current grade or year in school nor their highest level of education were correct (QD18CC02 = 3) but the updated answer for the current grade (from QD18CC03) and the highest level of education (from QD18CC04) were still inconsistent.

- The respondent was currently in the 12th grade or lower and reported completing a grade or level of education in QD18CC04 that was higher than his or her current grade from QD18 or QD18CC03. The respondent logically would not have completed a higher grade than his or her current grade.
- The respondent was currently in the 12th grade or lower, and the highest completed grade from QD18CC04 was lower than the current grade by more than two grade levels. These responses would indicate that the respondent skipped from his or her previous grade level to the current grade level by more than two levels.
- The respondent was currently in the first year of college or higher and reported in QD18CC04 that he or she completed less than the 11th grade or that he or she completed the 12th grade but did not receive a high school diploma (QD18CC04 = 13). These responses indicate explicitly that the respondent did not receive a high school diploma (when QD18CC04 = 13) or are implausible for the highest level of education if the respondent was currently at a level of schooling beyond high school.
- The respondent was currently in his or her second year of college or higher and reported that his or her highest level of education was the receipt of a general educational development (GED) certificate of high school equivalency (QD18CC04 = 14). The respondent logically would be expected not to go from the receipt of GED to being in the second year of college or higher.

These edits to set values in EDUHIGHST to bad data were implemented for about 40 respondents in 2016.

In addition, if respondents in 2016 confirmed that their current grade from QD18 was correct and reported in QD18CC04 that they received a high school diploma and that they were currently beyond their second year of college, then EDUHIGHST was assigned a code of 35, where 35 = Some college, no degree LOGICALLY ASSIGNED.⁶⁸ In this situation, the receipt of a high school diploma would technically have been correct. Because QD18CC04 was self-administered, however, it was assumed that the respondents incorrectly reported in QD18CC04 that their *highest* level of education was a high school diploma, when they may have had some further education beyond high school in order to be beyond their second year in college.

A second variable, EDUHIGHED, also was created as part of the self-administered demographics processing (Section 5.2.3.3) that took into account additional patterns in the data for QD11 and QD18. However, EDUHIGHED was not used in any published estimates involving educational attainment.

The base variable for creating an imputation-revised version of education was called EDEDUC. It was equivalent to EDUHIGHST, except that missing values that described the type of item nonresponse (i.e., "don't know" or "refused") were set to the SAS missing code (.) so that they were properly handled by the imputation-related modeling programs.

4.2.10 Creating the Perceived Health Status Variable (HEALTH)

HEALTH was created from interviewer-administered question QD12, which asked the respondent to rate his or her health. The question from the survey instrument appears below:

QD12: This question is about your overall health. Would you say your health in general is excellent, very good, good, fair, or poor?

- 1 EXCELLENT
- 2 VERY GOOD
- 3 GOOD
- 4 FAIR
- 5 POOR

No processing of this variable was done beyond that of assigning the edited variable name HEALTH instead of QD12 (the variable name obtained from the CAI).

4.2.11 Creating the FI Debriefing Variables

The FI debriefing section was completed by the interviewer after the interview to obtain information about the potential quality of the interview. This included information about factors that might have affected the quality of the data, such as the degree of privacy in the interview

⁶⁸ The corresponding edit was incorrectly not implemented in 2015 if respondents reported in QD18CC02 that neither answer from QD11 or QD18 was correct (QD18CC02 = 3), that they were currently in their third year of college or higher (QD18CC03 was greater than or equal to 15), and that their highest level of education was a high school diploma (QD18CC04 = 12). However, only a very small number of respondents in 2015 were affected, such that the highest level of education remained at the receipt of a high school diploma. In comparison, more than 13,000 adult respondents aged 18 or older in 2015 were classified as being high school graduates, and more than 17,000 adult respondents were classified as having some college education (but no degree) or an associate's degree.

setting. These questions were not read aloud to the respondent. Since 2015, these questions have been included as part of the household screener instrument (which was completed on a tablet) rather than at the end of the CAI instrument (which was completed on a laptop computer). Nevertheless, these FI debriefing variables underwent some editing.

Only minimal processing was done to the data in this section. Specifically, unedited variables were replaced with final, mnemonic variable names (e.g., PRIVACY for the variable pertaining to the interviewer's indication of how private the interview was). Where relevant, variables also were assigned legitimate skip codes based on the routing logic in this section.

Where the content of the debriefing questions remained the same (i.e., despite being completed as part of a different instrument), the names for edited variables beginning in 2015 were the same as those that were used prior to 2015. The variable TOOLONG was created for a question that was added beginning in 2015 for an interviewer to report if the respondent said that the interview was too long. In addition, the variable for the interviewer's perceptions of the respondent's understanding of the interview was changed from UNDINT in 2014 to TROUBUND (for "trouble understanding") in 2015 and subsequent years because this question was changed from a rating scale in 2014 for the respondent's level of understanding of the interview to a yes/no question in 2015 and subsequent years for whether the respondent had trouble understanding any of the interview questions.

4.3 Imputation for Selected Interviewer-Administered Demographics Variables

In this section, the imputation procedures applied to marital status, race, Hispanic/Latino origin, Hispanic/Latino group, and education level are described. These variables comprised five independent manifestations of the single response propensity (RP)/single prediction (PRD) type of PMN described in Chapter 3. Each imputation set is discussed in a separate section below.

4.3.1 Marital Status (Imputation Set 1)

The first interviewer-administered demographic variable that underwent imputation was the marital status variable. The four substantive levels of the imputation-revised variable IRMARIT matched the four answer categories of QD07 (i.e., married, widowed, divorced or separated, or never married). Respondents aged 12 to 14 were automatically assigned an IRMARIT value of 99, a "legitimate skip" code. Because this is the first variable to undergo imputation in each cycle, there were no imputation-revised variables to use as auxiliary variables. This tended to make the imputation process simple and straightforward.

In marital status imputation procedures, only one RP model and only one PRD model was fit; most other NSDUH imputation procedures are run separately within three or four age groups. Single models were used across all age groups to ensure adequate sample size for response categories that would be rare within certain age groups (e.g., the "widowed" category for younger age groups). To account for the correlation between age and marital status, AGE was used in both the RP and PRD model steps and in a likeness constraint in the hot-deck step.

4.3.1.1 Response Propensity Step

The RP model for imputation set 1 utilized the preliminary analysis weight, PANALWT. The marital status question QD07 was only asked of respondents aged 15 or older. Therefore, the domain contained unit respondents with AGE greater than or equal to 15. The creation of the AGE variable is described in Section 4.2.2. Unit respondents in the domain with nonmissing EDMARIT values were considered item respondents. The EDMARIT variable is described in Section 5.2.6. See [Table D.1](#) in Appendix D for details of the covariates used in the RP model for this variable.

4.3.1.2 Prediction Step

Using the adjusted weights that are outputs of the RP step, the marital status variable was modeled using polytomous logistic regression as implemented by the MULTILOG procedure in SUDAAN[®].⁶⁹ The outcome variable had four levels, which mapped to the four answer categories of QD07. The four predicted means used in the subsequent hot-deck step were the predicted probabilities that the respondent selected each of the four answer categories of QD07.

4.3.1.3 Hot-Deck Step

The predicted means from the PRD step play a central role in the donor selection algorithm applied in the hot-deck step, but unlike the RP and PRD steps, the hot-deck steps for marital status were run separately within three age groups: 15 to 17, 18 to 25, and 26 or older. This was done to allow parallel processing, which decreases the time required for implementation. No logical constraints were used, and the only likeness constraint other than the delta constraint involved the continuous AGE variable. The few unit nonrespondents requiring imputation for this variable are usually handled in the first attempt to find a donor, due to the mild set of constraints and large domain. The only imputation-revised variable created in the hot-deck step was IRMARIT.

4.3.2 Race Variables (Imputation Set 2)

As discussed in Section 4.2.7, race and Hispanicity were closely related. Therefore, race was used in the imputation of Hispanic/Latino origin and Hispanic/Latino group, and Hispanicity was used in the imputation of race. Since race underwent imputation first, imputation-revised versions of the Hispanic/Latino indicator and the Hispanic/Latino group were not available. This precluded their usage in race models. However, they were used extensively in constraints in the hot-deck step. The RP, PRD, and hot-deck steps were all run separately within three age groups: 12 to 17, 18 to 25, and 26 or older.

The race questions had low response rates relative to other questions in the NSDUH, due to high item nonresponse among Hispanic/Latino respondents. Nearly all the race nonrespondents reported being of Hispanic/Latino origin ([Table 4.1](#)). The likeness constraints involving Hispanic/Latino group strongly influenced the final imputed values.

⁶⁹ SUDAAN[®] is a registered trademark of Research Triangle Institute.

4.3.2.1 Response Propensity Step

The response propensity models for imputation set 1 utilized the preliminary analysis weight, PANALWT. The domain for the RP models included all unit respondents. Item respondents were those with EDRACE values from 1 to 15 and EDNWRACE values from 1 to 15. See [Table D.1](#) in Appendix D for details of the covariates used in the RP models for this variable.

4.3.2.2 Prediction Step

Using the adjusted weights that are outputs of the RP step, the race variables were modeled using polytomous logistic regression as implemented by the MULTLOG procedure in SUDAAN. The outcome variable was the five-level variable EDRACEFORMODEL, which had the following levels:

1. White Only
2. Black/African American Only
3. American Indian/Alaska Native Only
4. Asian/Other Pacific Islander Only
5. Multiple Race

In survey years prior to 2008, multiple-race respondents were assigned to one of the first four categories above. An edited variable that did not include a category for more than one race was useful in the past because (1) the multiple race cell contained a small number of respondents, making imputation models difficult to fit, and (2) it was necessary to be used as a base variable for the final imputation-revised variable that did not include a category for more than one race (between 2003 and 2007, called IRRACE2). On the first point, the multiple race category has become less sparse over time (refer to Section 3.3 of the 2008 imputation report [Ault et al., 2010] for more details). On the second point, because multiple-race respondents were classified as a separate category starting in 2008, a decision was made to cease to create IRRACE2, where multiple-race respondents were assigned a single race as shown in the first four categories above. It was replaced in most cases with the variable RACE4. The variable RACE4 is described in Section 4.3.3.4.

EDRACEFORMODEL is a recode of the variable EDRACE, described in Section 4.2.7.3.2:

EDRACEFORMODEL =

- EDRACE, if $1 \leq \text{EDRACE} \leq 4$; else
- 5, if $5 \leq \text{EDRACE} \leq 16$; else
- missing.

The five predicted means used in the subsequent hot-deck step were the predicted probabilities that the respondent had each value of EDRACEFORMODEL.

4.3.2.3 Hot-Deck Step

Each item nonrespondent in the hot-deck step was assigned one of 13 missingness patterns. Twelve of the missingness patterns, all rare, were set up to handle cases where something was known about the race categories such as "known to be Asian." The 13th missingness pattern, by far the most common, handled cases where nothing was known about the race categories. For a description of these missingness patterns, see [Table E.5](#). Logical constraints applied to the cases where something was known about the race categories. Otherwise, for the cases where nothing was known, only likeness constraints were used. Sometimes, what was "known" about the race categories came from a random assignment for indirectly mapped codes, as described in Section 4.2.7.2.1.

Besides the segment and delta likeness constraints, the likeness constraints based on Hispanic/Latino group were important determinants of the final imputed value, because the vast majority of the item nonrespondents for race were Hispanic/Latino. In 2016, 17.9 percent of the overall respondent pool were of Hispanic/Latino origin, but 97.55 percent of the item nonrespondents for race were of Hispanic/Latino origin. [Table 4.1](#) reports the distribution of Hispanic/Latino group among race item nonrespondents in 2016. Almost all are Hispanic/Latino and most (more than two thirds) of the Hispanic/Latino nonrespondents are Mexican only.

Table 4.1 Hispanic/Latino Status of Item Nonrespondents for Race

Hispanic/Latino Status	Item Nonrespondents for Race	
	Number	Percentage
Not Hispanic/Latino or Missing Hispanic/Latino Indicator	64	2.45
Hispanic/Latino	2,546	97.55
Mexican Only	1,787	68.47
Puerto Rican Only	240	9.20
Central/South American Only	226	8.66
Dominican Only	140	5.36
Other/Unknown	153	5.86
Total	2,610	100.00

Depending on the missingness pattern, the item nonrespondent received values from the selected donor for some subset of EDNWRACE, EDQD051-EDQD057, and a collapsed version of EDQD058-EDQD0515. The collapsed version of EDQD058-EDQD0515 is simply the minimum of these variables and is an indicator of whether the respondent was Asian. Most receive values for all variables. Item nonrespondents in missingness pattern 2 (known to be Asian but missing an Asian finer category) received values for only EDNWRACE, and item nonrespondents in missingness pattern 3 (known to be multiple race, but no other information) received values for everything except EDNWRACE. The imputation-revised versions of these variables are shown in [Table 4.2](#).

Table 4.2 Edited Race Variables and Their Imputation-Revised Counterparts

Edited Race Variable	Imputation-Revised Race Variable
EDQD051	IRRACEWH
EDQD052	IRRACEBK
EDQD053	IRRACENA
EDQD054	IRRACENH
EDQD055	IRRACEG
EDQD056	IRRACESM
EDQD057	IRRACPI2
EDQD058-EDQD0515 (collapsed)	IRRACEAS
EDNWRACE	IRNWRACE

Due to the strict constraints, the delta constraint had to be dropped sometimes. However, the likeness constraints related to Hispanic/Latino group were never dropped.

4.3.3 Hispanic/Latino Origin Variable (Imputation Set 3)

For the Hispanic/Latino origin indicator, the RP, PRD, and hot-deck steps were all run separately within three age groups: 12 to 17, 18 to 25, and 26 or older. Details on the procedures are given in the next four sections. The base variable for imputation, EDHOIND, is described in Section 4.2.7.4.1. The item response rate for this variable was much higher than for race.

4.3.3.1 Response Propensity Step

The response propensity models for imputation set 3 utilized the preliminary analysis weight, PANALWT. The domain indicator for the RP model included all unit respondents. Item respondents were those with a nonmissing value for EDHOIND. See [Table D.1](#) in Appendix D for details of the covariates used in the RP models for this variable.

4.3.3.2 Prediction Step

Using the adjusted weights that are outputs of the RP step, the Hispanic/Latino origin indicator was modeled using dichotomous logistic regression as implemented by the LOGISTIC procedure in SUDAAN.⁷⁰ The outcome variable was EDHOIND. The single predicted mean was the predicted probability that the respondent was of Hispanic/Latino origin.

4.3.3.3 Hot-Deck Step

The hot-deck step for the Hispanic/Latino origin indicator included a single predicted mean from the PRD step, no logical constraints, and only the segment and delta likeness constraints. EDHOIND is the base variable for imputation, and the imputation-revised version is

⁷⁰ In SAS-callable SUDAAN, this is the RLOGIST procedure to avoid confusion with SAS's own LOGISTIC procedure.

called IRHOIND. Details on the hot-deck step, including the likeness constraints, are available in [Tables E.6](#) and [E.7](#) in Appendix E.

4.3.3.4 Recodes for Additional Race/Ethnicity Variables

The imputation-revised race (IRNWRACE) and imputation-revised Hispanic/Latino indicator (IRHOIND) variables were used to create several additional combined race/ethnicity variables. One of these (RACE4) was used in the subsequent processing of imputation-revised variables and had four levels: non-Hispanic/Latino white, non-Hispanic/Latino black/African American, Hispanic/Latino, and non-Hispanic/Latino other/multiple race. The NEWRACE1 and NEWRACE2 variables also were created from IRNWRACE and IRHOIND and were used extensively in the production of the 2016 detailed tables (Center for Behavioral Health Statistics and Quality, 2017f).

4.3.4 Hispanic/Latino Group Variable (Imputation Set 4)

The edited variable EDHOGRP, described in Section 4.2.7.4.3, categorized Hispanic/Latino respondents into Hispanic/Latino groups. These categories were directly mapped to the same categories in the imputation-revised variable, IRHOGRP4, which had eight possible values: Puerto Rican, Mexican, Cuban, Central or South American, Dominican, Spanish, Other Hispanic/Latino, and not Hispanic/Latino. The closely-related imputation-revised variable IRHOG RPM was also created to identify respondents who selected more than one Hispanic/Latino group; recall that a priority rule is used to assign a single group to multiple-group respondents in the creation of EDHOGRP (and therefore IRHOGRP4).

Imputations were not conducted separately within age groups. The Hispanic/Latino group variables were created only for respondents of Hispanic/Latino origin as defined by IRHOIND. This results in a small domain. The models were likely to be better when age groups were combined because (1) none of the response categories were sparsely populated and (2) sufficiently large donor pools were ensured in the hot-deck step.

4.3.4.1 Response Propensity Step

The response propensity models for imputation set 4 utilized the preliminary analysis weight, PANALWT. The domain indicator included all respondents of Hispanic/Latino origin as defined by IRHOIND. Item respondents were those with a nonmissing value for EDHOGRP who selected only a single Hispanic/Latino group. The multiple-group respondents whose EDHOGRP was assigned by the priority rule (Scenario 3, described in Section 4.2.7.4.3) were not used to fit the PRD model in the next step. See [Table D.1](#) in Appendix D for details of the covariates used in the RP models for these variables.

4.3.4.2 Prediction Step

Because the model would have been much more difficult to fit if all seven levels were used, the EDHOGRP variable was collapsed into a four-level categorical variable (EDHOGRP2). [Table 4.3](#) shows the mapping of EDHOGRP levels to EDHOGRP2 levels. Using the adjusted weights that are outputs of the RP step, EDHOGRP2 was then modeled using polytomous logistic regression as implemented by the MULTLOG procedure in SUDAAN. The

four predicted means used in the subsequent hot-deck step were the predicted probabilities that the respondent had each value of EDHOGRP2.

Table 4.3 Mapping of EDHOGRP Levels to EDHOGRP2 Levels

EDHOGRP (Base Variable)	EDHOGRP2 (Modeled Variable)
Mexican	Mexican
Puerto Rican	Puerto Rican
Central or South American	Other Hispanic/Latino
Cuban	Cuban
Dominican	Other Hispanic/Latino
Spanish	Other Hispanic/Latino
Other Hispanic/Latino	Other Hispanic/Latino
Not Hispanic	Not Hispanic

4.3.4.3 Hot-Deck Step

The hot-deck step for the Hispanic/Latino group variables was straightforward. Besides the segment and delta likeness constraints, the most notable feature was a likeness constraint involving race. A five-level race variable was used as a covariate in the RP and PRD models with the following levels: White Only, Black/African American Only, American Indian/Alaska Native Only, Asian Only, and Multiple Race. To further exploit the relationship between race and Hispanic/Latino group, a likeness constraint required the donor's selected race(s) to match a subset of the racial categories mentioned by the recipient. The constraint did not apply if the recipient was an item nonrespondent for race.

IRHOGRP4 was the imputation-revised version of EDHOGRP. The other imputation-revised variable IRHOGRP4M was set equal to 8 (more than one Hispanic/Latino group) if either the respondent reported membership in more than one group, or the donor for a particular item nonrespondent reported membership in more than one group. Otherwise, IRHOGRP4M was set equal to IRHOGRP4.

The Hispanic/Latino group variables generally have low imputation rates. The number of cases with missing data is usually fewer than 300 in each survey year.

4.3.4.4 Recodes for Additional Analyses

Among the recoded variables that were created from IRHOGRP4, the variable HISPGRP2 was used in subsequent processing and was created by collapsing the levels of IRHOGRP4 into four levels: Puerto Rican, Mexican, Other Hispanic/Latino (includes Cuban, Central or South American, Dominican, Spanish, and Other Hispanic/Latino), and not Hispanic/Latino.

4.3.5 Education Level Variable (Imputation Set 5)

The imputation-revised education level variable was similar to the imputation-revised Hispanic/Latino group variable in that it was categorical with numerous levels, and as with the

Hispanic/Latino group, the response variable for the PRD model was collapsed into fewer levels for ease of modeling. There were generally very few missing cases for this variable, fewer than 50, so the application of the method tended to be straightforward. Two age groups were used for RP and PRD modeling: 12 to 17 and 18 or older. However, the hot-deck step was implemented separately for three age groups: 12 to 17, 18 to 25, and 26 or older.

4.3.5.1 Response Propensity Step

The response propensity models for imputation set 5 utilized the preliminary analysis weight, PANALWT. The domain indicator for each of the two RP models included all unit respondents. Item respondents were those with a nonmissing value for EDEDUC. See [Table D.1](#) in Appendix D for details of the covariates used in the RP models for this variable.

4.3.5.2 Prediction Step

EDEDUC was collapsed into fewer levels for modeling. The response variables were different for the two PRD models: the response variable for the 12-17 age group had five levels, and the one for the 18-or-older age group had four. The mapping of EDEDUC to the response variable RESPEDUC is shown in [Table 4.4](#).

Table 4.4 Mapping of EDEDUC Levels to RESPEDUC Levels

EDEDUC (Base Variable)	RESPEDUC (Modeled Variable)	
	12-17	18+
No Schooling Completed	Less than Elementary School	Less than High School
1 st grade completed		
2 nd grade completed		
3 rd grade completed		
4 th grade completed		
5 th grade completed		
6 th grade completed	Elementary School	
7 th grade completed		
8 th grade completed	Middle School	
9 th grade completed		
10 th grade completed	Some High School	
11 th grade completed		
12 th grade, No Diploma		
Regular High School Diploma	High School	High School
GED Certificate of High School Completion		Some College/Associate's Degree
Some College Credit, but No Degree		
Associate's Degree		College Graduate
Bachelor's Degree		
Master's Degree		
Doctorate Degree		
Professional Degree Beyond a Bachelor's Degree		

Using the adjusted weights that are outputs of the RP step, both PRD models were fit using polytomous logistic regression as implemented by the MULTILOG procedure in SUDAAN. The predicted means matched the levels of the response variable, so there were five predicted means for the 12-17 hot-deck step and four for the 18-25 and 26-or-older hot-deck steps.

4.3.5.3 Hot-Deck Step

The hot-deck step for the education level variable was straightforward and implemented separately for three age groups: 12 to 17, 18 to 25, and 26 or older. The only base variable was EDEDUC, and the imputation-revised version was called IREDUHIGHST. Both variables are based on the detailed 18-level variable, as compared with the simplified RESPEDUC variable used in the RP and PRD steps. No logical constraints were required. In addition to the segment and delta likeness constraints, the third likeness constraint required the donor to be the same age as the recipient. This was an especially important constraint for the 12-17 age group, because the age covariate often had to be dropped from the PRD model due to near-empty cells when the variables were cross-tabulated, causing instability in the estimates.

4.3.5.4 Recode for Additional Education Variable

EDUHIGHCAT, a recoded education variable, was created using the imputation-revised highest grade completed variable (IREDUHIGHST). EDUHIGHCAT had five levels (less than high school and aged 18 or older, high school graduate and aged 18 or older, some college/Associate's Degree and aged 18 or older, college graduate and aged 18 or older, or 12 to 17 years old).

If a respondent was 12 to 17 years old, then EDUHIGHCAT was assigned that level. If IREDUHIGHST was "No Schooling Completed" or any grade level up to "12th Grade, No Diploma," then EDUHIGHCAT was designated "Less than High School." If IREDUHIGHST was either "High School Diploma" or "GED," then EDUHIGHCAT was assigned "High School." If IREDUHIGHST was "Some College Credit, but No Degree" or "Associate's Degree," then EDUHIGHCAT was designated "Some College/Associate's Degree." Any degree beyond "Associate's Degree" was categorized as "College Graduate." EDUHIGHCAT was often used as a covariate in later imputation models.

5. Editing and Imputation for the NSDUH Self-Administered Demographics Variables

5.1 Introduction

This chapter discusses procedures for editing and imputing the self-administered demographic variables. The interviewer-administered demographics variables that were edited and imputed in the 2016 National Survey on Drug Use and Health (NSDUH) were discussed in Chapter 4.

This chapter discusses editing procedures for the following self-administered demographics sections:

- moves in the past 12 months,
- immigrant status,
- adult sexual orientation,
- English language proficiency,
- disability,
- current education,
- military families, and
- employment and workplace.

This chapter also discusses imputation procedures for the variables pertaining to immigrant status and current employment status. Other edited variables for self-administered demographics that are discussed in this chapter were not imputed.

5.2 Editing the Self-Administered Demographics Variables

This section documents the editing procedures for the self-administered demographics sections that were listed in Section 5.1. As noted in Section 1.1, the self-administered demographics sections of the interview were administered by the interviewers using audio computer-assisted self-interviewing (ACASI).

5.2.1 Editing for Patterned Responses and Questionable Interview Dates

As discussed in Section 2.3.2.2, edits for patterned responses were implemented in the self-administered demographic questions beginning in 2016. These edits resulted in data in self-administered demographics variables being set to "bad data" (Section 2.3) for slightly more than 10 respondents in 2016 who keyed responses of only 1 (or variants such as 11) or 2 (or variants such as 22) wherever possible in the self-administered demographics questions.

However, the new edits in 2016 did not identify patterned responses in the self-administered demographic data for about 15 additional respondents aged 15 or older. Further investigation indicated that keeping the data for these additional respondents instead of setting the responses to "bad data" would have minimal effects on estimates. Corrected checks for identifying patterned responses in this section of the interview have been implemented for 2017.

As discussed in Section 2.3.3, a small number of respondents in 2016 had reported interview dates that were outside of the quarter in which the interviews had been completed. Because the demographics variables described in this chapter were self-administered, the edits for setting date-dependent variables (e.g., number of moves in the past 12 months) to bad data (Section 2.3) applied to the variables in this chapter.

5.2.2 Moves in the Past 12 Months

The questions pertaining to moves in the past 12 months and the state where respondents lived a year ago as they appear in the survey instrument are presented below.

QD13 How many times in the **past 12 months** have you moved? Please include moves from one residence to another within the same city/town as well as those from one city/town to another.

NUMBER OF TIMES: _____ [RANGE: 0 - 365]

DK/REF

PROGRAMMER: SHOW 12 MONTH CALENDAR

QD13A [IF QD13 NE 0] In what state did you live on [FILL PAST 1 YEAR DATE], that is, one year ago today?

1 Alabama	27 Montana
2 Alaska	28 Nebraska
3 Arizona	29 Nevada
4 Arkansas	30 New Hampshire
5 California	31 New Jersey
6 Colorado	32 New Mexico
7 Connecticut	33 New York
8 Delaware	34 North Carolina
9 The District Of Columbia (Washington, Dc)	35 North Dakota
10 Florida	36 Ohio
11 Georgia	37 Oklahoma

12	Hawaii	38	Oregon
13	Idaho	39	Pennsylvania
14	Illinois	40	Rhode Island
15	Indiana	41	South Carolina
16	Iowa	42	South Dakota
17	Kansas	43	Tennessee
18	Kentucky	44	Texas
19	Louisiana	45	Utah
20	Maine	46	Vermont
21	Maryland	47	Virginia
22	Massachusetts	48	Washington
23	Michigan	49	West Virginia
24	Minnesota	50	Wisconsin
25	Mississippi	51	Wyoming
26	Missouri	52	Outside Of U.S.

DK/REF

In 2013 and 2014, question QD13 was asked only if respondents reported one or more moves in the past 5 years in the social environment section (for adults aged 18 or older) or the youth experiences section (for adolescents aged 12 to 17). Prior to 2013, all respondents were asked QD13. Because the questions about moves in the past 5 years were deleted beginning in 2015 from the respective sections for adults and adolescents, all respondents in 2016 were eligible to be asked QD13, as was the situation prior to 2013.

Although the logic for asking QD13 reverted back to all respondents being eligible to be asked the question, the corresponding variable was not renamed back to the variable that existed in 2012 and prior years because QD13 was self-administered in 2015 (i.e., instead of being interviewer administered). The variable corresponding to QD13 was named MOVSPNPYR in 2015 and subsequent years.

If respondents moved at least once in the past 12 months, they were asked in question QD13A (edited variable MOVFRSTAT) to report the state where they were living a year prior to their interview date. If respondents reported zero moves in the past 12 months in QD13, then MOVFRSTAT was assigned a legitimate skip code (Section 2.4.2). If QD13A had been skipped

because QD13 had a missing value (e.g., "don't know," "refused"), then MOVFRSTAT retained a code of 98 (blank).

5.2.3 Immigrant Status

The questions pertaining to immigrant status are presented below.

QD14 Were you born in the United States?

- 1 Yes
- 2 No
- DK/REF

QD15 [IF QD14 = 2] In what country or U.S. territory were you born?

COUNTRY OR US TERRITORY: _____
DK/REF

QD16A [IF QD14 = 2 and (QD13 > 0 OR QD13 = DK/REF)] Have you lived in the United States for at least one year?

- 1 Yes
- 2 No
- DK/REF

QD16B [IF QD16A = 1 OR (QD14=2 AND QD13=0 AND QD13 NE DK/REF)] For how many years have you lived in the United States? Please estimate the total number of years you have lived here over your lifetime.

NUMBER OF YEARS: _____ [RANGE: 1-90]
DK/REF

QD16C [IF QD16A = 2] For how many months have you lived in the United States? Please estimate the total number of months you have lived here over your lifetime. If you have lived in the United States for less than one month, please enter 0.

NUMBER OF MONTHS: _____ [RANGE: 0-12]
DK/REF

5.2.3.1 Births within or Outside of the United States

Respondents were asked in question QD15 to report the country or U.S. territory where they were born if they reported in QD14 that they were not born in the United States. With the change to ACASI beginning in 2015, instances increased of respondents reporting in QD14 that they were not born in the United States but then specifying a state or the District of Columbia in QD15. However, respondents were not logically inferred to have been born in the United States if they specified in QD15 that they were born in a U.S. military installation outside of the United States.

The move of these questions to ACASI beginning in 2015 also increased the occurrence of missing data in QD14 and QD15. This increase included increases in answers of "don't know" or "refused" in QD14 and QD15. In addition, responses in QD15 for the other country of birth that were coded as "bad data" (Section 2.3) increased after these questions became self-administered. The editing and imputation report in the 2015 NSDUH methodological resource book (MRB; Center for Behavioral Health Statistics and Quality, 2017a) documents the changes in occurrences of inconsistent or missing data between 2014 and 2015 for QD14 and QD15.

When respondents reported in QD14 that they were not born in the United States but they specified a state or the District of Columbia in QD15, the edited variable IMMBORNUS corresponding to QD14 was assigned a code of 3, where 3 = Yes LOGICALLY ASSIGNED. Because a large number of respondents since 2015 have specified a state or the District of Columbia in QD15, the data in QD15 (edited variable IMMBORNUS) were not overwritten with codes of 989 (LEGITIMATE SKIP Logically assigned; Section 2.4.2) to allow analysts to see the responses in QD15 that led to the logical inference that respondents were born in the United States.

Otherwise, when respondents reported in QD14 that they were born in the United States, IMMBORNUS was assigned a legitimate skip code. This assignment of a legitimate skip code did not include situations in which respondents had missing data for where they were born (i.e., don't know, refused, or bad data). However, if respondents refused to answer QD14, then that refusal was propagated to IMMBORNUS.

5.2.3.2 Length of Time Living in the United States

The logic for asking questions QD16A through QD16C about the length of time that respondents have lived in the United States (if QD14 was answered as "no") was updated beginning in 2015 to take into account the number of moves in the past 12 months from QD13. Whether a respondent had lived in the United States for at least 1 year (QD16A) was asked in 2016 only if respondents who were not born in the United States reported at least one move in the past 12 months in question QD13 or if they answered QD13 as "don't know" or "refused." The assumption was that immigrants who reported no moves in the past 12 months logically have lived in the United States for at least 1 year. The number of years that immigrant respondents have lived in the United States (QD16B) was asked in 2016 if they reported in QD16A that they had lived in the United States for at least 1 year or they reported no moves in the past 12 months in QD13.

Editing of the variables for the length of time that immigrant respondents had lived in the United States principally involved assigning legitimate skip codes (Section 2.4.2) when the associated questions did not apply. In particular, if respondents reported in QD14 that they had been born in the United States (IMMBORNUS = 1), then the edited variables IMMYRMORE and IMMOSNUM (corresponding to QD16A and QD16C, respectively) were assigned a code of 99 and IMMYRSNUM (corresponding to QD16B) was assigned a code of 999. If the edits described in Section 5.2.3.1 resulted in respondents being logically inferred to have been born in the United States (IMMBORNUS = 3), then any nonblank values in IMMYRMORE and IMMOSNUM were overwritten with a code of 89 (LEGITIMATE SKIP Logically assigned) and nonblank values in IMMYRSNUM were overwritten with a code of 989.

Similarly, a legitimate skip code was assigned to IMMYRMORE and IMMOSNUM if immigrant respondents reported no moves in the past 12 months in QD13. If respondents were asked QD16A, then a legitimate skip code was assigned to IMMOSNUM if respondents reported in QD16A that they had lived in the United States for at least 1 year or to IMMYRSNUM if respondents reported that they had not lived in the United States for at least 1 year.

The number of years that respondents reported living in the United States in QD16B also was compared with respondents' ages. If the number of years in QD16B was greater than a respondent's current age, then IMMYRSNUM was assigned a code for bad data (Section 2.3) to indicate that the length of time that the respondent reported living in the United States was inconsistent with his or her age. No editing was done to IMMYRSNUM if QD16B was equal to the respondent's current age.

5.2.4 Adult Sexual Orientation

Since 2015, questions QD62 (edited variable SEXATTRACT) and QD63 (edited variable SEXIDENT) have been added to the NSDUH questionnaire to ask adult respondents aged 18 or older about their sexual attraction and sexual identity, respectively. The wording and structure of these questions closely resemble the wording and structure of questions that have been used in the National Survey of Family Growth. Lesbian, gay, bisexual, and transgender individuals are an important subpopulation for the Substance Abuse and Mental Health Service Administration's (SAMHSA's) Strategic Initiative for the prevention of substance use and mental disorders.

QD62 [IF CURNTAGE = 18 OR OLDER] People are different in their sexual attraction to other people. Which statement best describes your feelings?

[IF QD01=5]

- 1 I am only attracted to females
 - 2 I am mostly attracted to females
 - 3 I am equally attracted to females and males
 - 4 I am mostly attracted to males
 - 5 I am only attracted to males
 - 6 I am not sure
- DK/REF

[IF QD01=9]

- 1 I am only attracted to males
- 2 I am mostly attracted to males
- 3 I am equally attracted to males and females
- 4 I am mostly attracted to females
- 5 I am only attracted to females
- 6 I am not sure

QD63 [IF CURNTAGE = 18 OR OLDER] Which one of the following do you consider yourself to be?

- 1 Heterosexual, that is, straight
- 2 [IF QD01=9 THEN "Lesbian or] Gay
- 3 Bisexual
- DK/REF

Because only adults were asked to report their sexual attraction and sexual identity, adolescent respondents aged 12 to 17 were skipped out of these questions. No editing was done to SEXATTRACT and SEXIDENT other than to assign legitimate skip codes if respondents were aged 12 to 17.

5.2.5 English Language Proficiency and Disability

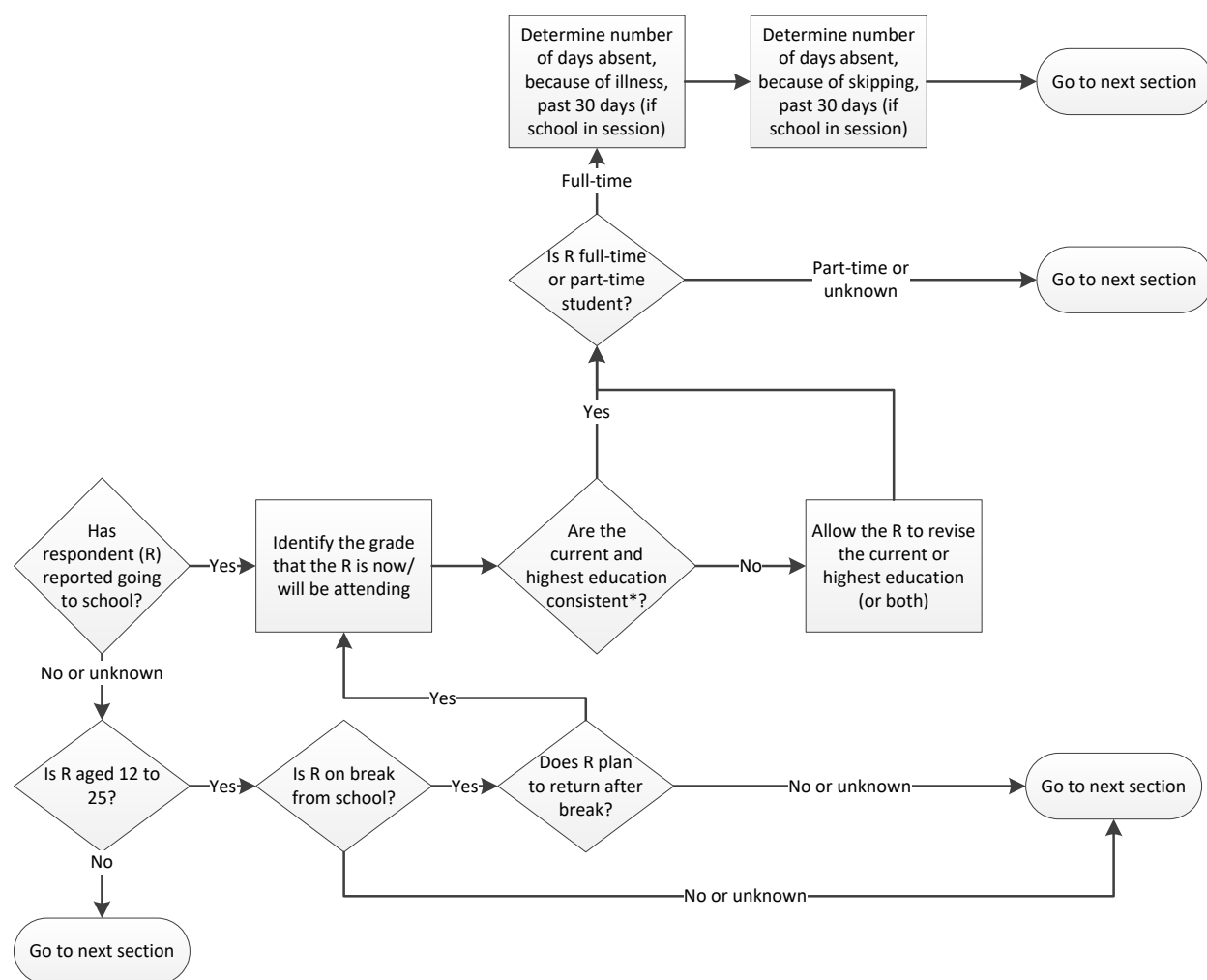
Since 2015, questions QD55 through QD61 on disability status and English language proficiency have been included in the NSDUH questionnaire. These items are intended to help monitor health disparities in populations with disabilities or limited proficiency with English.

Except where indicated, all respondents were asked these questions. No editing was done to the associated variables, other than to assign final mnemonic names to the variables (e.g., SPEAKENGL, corresponding to question QD55 for how well respondents speak English) and to assign legitimate skip codes to one variable for respondents who were outside of the eligible age range to be asked the corresponding question. Specifically, question QD61 (edited variable DIFFERAND) asked respondents aged 14 or older whether they had difficulty doing errands alone because of a physical, mental, or emotional condition. Thus, legitimate skip codes were assigned to DIFFERAND for respondents aged 12 or 13.

5.2.6 Current Education

The questions for current education assessed whether respondents were currently enrolled in school. If so, respondents were asked to report their current grade in school (or the grade they would be in once they returned from school break), whether they were a full-time or part-time student, and if they were full-time students, the number of days that they missed school in the past 30 days because they were sick or because they skipped school (questions QD18 through QD21). [Figure 5.1](#) shows the logic for asking follow-up questions if respondents reported that they were enrolled in school.

Figure 5.1 Follow-Up Question Logic for Current Education



*A consistency check between the current and highest education was triggered in 2016 if the respondent (R) (1) received a bachelor's degree but was (or will be) in the 4th year of college or lower, (2) completed the 12th grade without receiving a diploma but was (or will be) in the 12th grade or lower, or (3) received a general educational development certificate of high school completion but was (or will be) in the 12th grade or lower.

Question QD17 underwent changes for 2016. First, the question "Are you now attending or are you currently enrolled in school?" was changed to "Do you go to school?" The question about going to school also was moved to follow the explanation of what was meant by "school." In addition, an instruction was included that respondents should answer QD17 as "yes" if they were on a holiday or break from school and planned to return to school when their break was over. Although only QD17 changed for 2016, these changes were determined to have affected the comparability of the entire self-administered education data between 2015 and 2016. Therefore, the education variables in this section were renamed for 2016. Table 5.1 shows the names of the edited variables that corresponded to the questions in the current education section for 2016 and the corresponding names from the 2015 survey.

Table 5.1 Names in 2016 and 2015 for Current Education Variables

Question	Measure	2016 Name	2015 Name
(None)	Additional edited variable for the highest level of education (in addition to edits from the consistency checks)	EDUHGHEDT	EDUHIGHED
QD17, QD17A, QD17B	Currently going to school	EDUSCHLGO	EDUENROLL
QD18	Current or upcoming grade in school	EDUSCHGRD	EDUGRDNOW
QD19	Full-time or part-time student status	EDUFULPAR	EDUFULPRT
QD20	Number of days missed school in the past 30 days because of sickness or injury	EDUSCKMON	EDUSICKMO
QD21	Number of days missed school in the past 30 days because of skipping or "cutting" school	EDUSKPMON	EDUSKIPMO

NOTE: Questions QD18 through QD21 applied if respondents reported in QD17 through QD17B that they were currently in school or planned to return to school when their break was over. See [Figure 5.1](#) for additional information on routing within the current education section.

As described in Section 1.4.1.12, the following additional changes were made to the current education questions beginning in 2015.

- A consistency check was added that checked respondents' answers for their highest level of education and their current grade or year in school. In addition to data from the consistency check being used in editing the variable for the highest level of education (EDUHIGHST), these data were used in editing the variable for respondents' current (or upcoming) grade in school (EDUSCHGRD).
- For respondents who were not attending school, questions were dropped for whether they received a high school diploma or a general educational development (GED) certificate of high school equivalency because this information was captured in the revised question QD11 for the highest level of education.
- For respondents who were not in school and had not received a high school diploma or GED, questions were dropped for respondents' reasons for leaving school without receiving a high school diploma and the age when they left school.

The remainder of this section discusses the following issues that were relevant to editing of the current education variables:

- overall routing logic for the current education section (Section 5.2.6.1);
- edits involving respondents' school enrollment status (Section 5.2.6.2);
- general edits involving the last and current grade for respondents who were in school (Section 5.2.6.3);
- specific edits involving the last and current grade for respondents aged 12 to 17 who were in school (Section 5.2.6.4);

- specific edits involving the last and current grade for respondents aged 18 or older who were in school (Section 5.2.6.5); and
- miscellaneous edits involving the current education variables (Section 5.2.6.6).

Table B.1 in Appendix B describes edits for the current education variables pertaining to current school enrollment, last and current grade, and associated variables if respondents were currently in school. Table B.1 also is intended as a companion to the discussion of the editing procedures for the current education variables in Sections 5.2.6.2 through 5.2.6.5.

5.2.6.1 Routing Logic for Current Education

As shown in Figure 5.1, determining whether respondents were currently going to school was critical for determining whether they would be asked further questions about their education or they would be routed to the next relevant set of questions (e.g., QD10D for family members currently serving in the United States military). However, respondents could answer the question about going to school (QD17) as "no" if they were on a holiday break or vacation from school. Therefore, respondents who did not report in question QD17 that they were currently going to school were asked follow-up questions (if they were aged 12 to 25 and their highest reported level of education from question QD11 was an associate's degree or lower) to determine if they were on a holiday or vacation break from school (question QD17B), and if so, whether they intended to return to school once their break was over (question QD17B).

If respondents reported in QD17 that they were not currently going to school (QD17 = 2) but they were on break from school (QD17A = 1) and they intended to return to school once their vacation or break was over (QD17B = 1), the edited variable EDUSCHLGO for going to school (based on data from QD17, QD17A, and QD17B) was set to a value of 1 ("yes") to indicate that the respondents should be considered to be going to school. Otherwise, the response from QD17 was carried over to EDUSCHLGO. That included situations in which respondents reported in QD17A that they were not on vacation break from school, or who reported in QD17B that they were on break but did not intend to return to school once their break was over.

If respondents were currently going to school or were on break but intended to return, the wording of subsequent questions varied according to their current status. For question QD18, respondents who reported in QD17 that they were currently going to school (QD17 = 1) were asked to report their current grade in school (i.e., at the time of the interview). For respondents who were on vacation break from school but intended to return to school once their break was over (QD17B = 1), question QD18 asked for the grade that they would be in once they returned from their vacation break. Similarly, for the question about full-time or part-time student status (question QD19) respondents who were currently going to school were asked "Are you a full-time student or a part-time student?" Those who were on break from school but intended to return to school were asked "Will you be a full-time or a part-time student?"

Based on the logic shown in Figure 5.1, a key aspect of editing the current education variables involved assigning legitimate skip codes (Section 2.4.2) based on whether respondents were currently going to school and respondents' ages. If respondents were not currently going to school, the edited variables corresponding to questions QD18 through QD21 (EDUSCHGRD, EDUFULPAR, EDUSCKMON, and EDUSKPMON) were assigned legitimate skip codes.

5.2.6.2 Edits Involving Current School Enrollment

Because the routing logic in the current education section was based on respondents' reports of whether they were going to school (or not), an important aspect of editing the variables in this section involved editing of the variable EDUSCHLGO for going to school. If respondents were not currently going to school (EDUSCHLGO = 2), then EDUSCHGRD, EDUFULPAR, EDUSCKMON, and EDUSKPMON were assigned legitimate skip codes (Section 2.4.2).

5.2.6.3 General Edits Involving the Last and Current Grade

The current school grade from question QD18 could be inconsistent with the highest level of education that the respondent reported completing in question QD11. Although the consistency check questions QD18CC01 to QD18CC04 that have been included since 2015 give respondents the opportunity to resolve inconsistencies between QD11 and QD18, the consistency checks did not cover all possible inconsistencies between these two questions. Therefore, edits were implemented to cover other inconsistencies in the data.

In particular, the changes to the response options in QD11 for levels beyond the 11th grade (Section 4.2.8) affected the editing of data between QD11 and QD18. For respondents who had completed the 11th grade or lower, one might expect the current grade in QD18 to be one grade level higher than the response in QD11. In addition, no editing was done when the current grade reported in QD18 was the 11th grade or lower and was the same as the highest grade reported in QD11 because respondents could have been repeating a grade.

As noted in Section 4.2.8, editing of the interviewer-administered education variable EDUHIGHST (highest grade completed, corresponding to question QD11) with respect to data in QD18 for a respondents' current grade was limited to the consistency checks between QD11 and QD18 because QD18 was in a separate section of the interview. However, a second variable, EDUHGHEDT, was created as part of the editing of the current education section in which the highest level of education underwent further edits. Consequently, the interviewer-administered education variable EDUHIGHST would not be affected by future changes that might occur in the content of current education variables. Nevertheless, the EDUHGHEDT variable might in some situations be a more accurate reflection of the highest level of education that respondents had completed.

The general education edits discussed in the remainder of this section were invoked if an inconsistency occurred in the answers from QD11 and QD18 that had not triggered a consistency check between these two questions. An algorithm that had been developed in previous years to handle these types of inconsistencies was updated beginning in 2015 to restrict the edits to respondents aged 12 to 17. This algorithm is discussed in detail below.

For respondents aged 12 to 17, a series of arrays was set up that mapped out the highest grade and current grade that would be *expected* relative to a respondent's current age, assuming an orderly progression from one grade level to the next highest level. [Table 5.2](#) shows a matrix that maps the current age with expected grades in the United States.

Table 5.2 Mapping of Current Age with Expected Grades

Current Age	12	13	14	15	16	17
Expected Completed Grade	6	7	8	9	10	11
Expected Current Grade	7	8	9	10	11	12

For example, one might expect most people in the United States to have completed the 6th grade by the time they are 12 years old. It would therefore not be unreasonable for someone to be aged 12 and to be currently in the 7th grade, depending on when the respondent was interviewed. An upper age limit was set at 17 because in 2016, response options in QD11 for levels beyond the 11th grade no longer referred to single years of education. For example, levels 12 and 13 in QD11 in 2014 referred to completion of the 12th grade and completion of the first year of college, respectively. Since 2015, level 12 in QD11 has referred to the receipt of a high school diploma, and level 13 has referred to completion of the 12th grade without the receipt of a diploma.

In addition, the algorithm allowed for some deviation relative to the expected ages. For example, if a respondent was aged 12 and was currently in the 9th grade, this was still considered to be an acceptable pattern because the respondent might have received accelerated education beyond his or her expected grade.

Most edits were done separately depending on whether a respondent was aged 12 to 17 or 18 or older. The rationale for doing edits separately for these two age groups was that most respondents aged 12 to 17 would not be expected to have graduated from high school. In contrast, adults aged 18 or older could report in question QD11 that they had received a high school diploma, some education beyond high school but no degree, an associate's degree, a bachelor's degree, or a graduate or professional degree.

5.2.6.4 Edits Involving the Last and Current Grade for Respondents Aged 12 to 17

For respondents aged 12 to 17, the highest grade completed and the current grade were considered to be consistent with the respondent's age if what was reported was within 2 years of the grades shown in [Table 5.2](#). Thus, if a respondent was aged 12, the algorithm considered completion of any grades from the 4th through the 8th to be sufficiently consistent with the respondent's age. Similarly, for respondents aged 12, the algorithm considered current enrollment in any grades from the 5th through the 9th to be sufficiently consistent with the respondent's age.

Therefore, the following four data combinations were possible:

- both the completed grade and the current grade were considered to be consistent with the respondent's age;
- the highest completed grade was consistent with the respondent's age, but the current grade was not;

- the current grade was consistent with the respondent's age, but the highest completed grade was not; or
- neither the highest completed grade nor the current grade was consistent with the respondent's age.

Separate edits were done according to the four combinations of data patterns described immediately above. The following edits were done if both the completed grade and current grade were considered to be consistent with the respondent's age:

- If the current grade was more than two grade levels higher than the highest completed grade, the current grade was edited to be consistent with the highest grade because the latter was an interviewer-administered variable. For example, if a respondent was aged 12, reported completing the 6th grade, and reported currently being in the 9th grade (QD18 = 9), then the edits logically inferred that the respondent currently was in the 7th grade because this answer was more consistent with both the respondent's current age and the highest completed grade. The edited variable for current grade (EDUSCHGRD) was assigned a code of 27 (i.e., 7th grade LOGICALLY ASSIGNED).
- If the current grade was lower than the highest completed grade, the editing code gave precedence to the reported grade that would yield the most consistent result relative to the respondent's age. In particular, if accepting the report of the highest grade and inferring that the respondent was currently in the next highest grade would yield a current grade that was inconsistent with the respondent's age, then the edited variable EDUHGHEDT for the highest level of education was assigned a value consistent with the current grade. Suppose, for example, that a 12-year-old respondent reported currently being in the 6th grade but completed the 7th grade. Accepting the answer that the respondent was currently in the 6th grade and had completed the 5th grade would be more consistent with the respondent's current age than would be the converse (i.e., accepting that this respondent had completed the 7th grade and inferring that he or she was currently in the 8th grade). In this example, EDUHGHEDT would be assigned a code of 25 (i.e., 5th grade LOGICALLY ASSIGNED).

If the highest completed grade was consistent with the respondent's age but the current grade was not, the highest completed grade was accepted by default. This was done if the current grade was lower than the highest completed grade or the current grade was more than two grade levels higher than the highest completed grade. The edited current grade EDUSCHGRD was therefore assigned a value to indicate a current grade level that was 1 year higher than the highest completed grade. For example, if the respondent reported completing the 10th grade, EDUSCHGRD would be assigned a code of 31 (i.e., 11th grade LOGICALLY ASSIGNED).

If the current grade was consistent with the respondent's age but the highest completed grade was not, the editing procedures accepted the current grade by default. Thus, if a 12-year-old respondent reported last completing the 3rd grade and reported currently being in the 6th grade, this edit would identify the current 6th grade as being consistent with an age of 12; that is, being in the 6th grade would be more consistent with the respondent's current age compared with

completing the 3rd grade. In this example, EDUHGHEDT would be assigned a code of 25 (i.e., 5th grade LOGICALLY ASSIGNED).

If neither the current grade nor the reported highest grade was consistent with the respondent's age, the following was done:

- If the current grade was lower than the highest grade that was reported, the algorithm picked the answer that was closest to the expected grade, based on the matrix shown in [Table 5.2](#). The variable with the more inconsistent data was assigned a bad data code. This edit allowed for situations where respondents may have fallen behind where they would be expected to be grade-wise (e.g., if they had been held back a year).
- If the current grade was exactly two grade levels higher than the reported highest completed grade and the highest completed grade was higher than what would be expected for the respondent's age, no further editing was done. Otherwise, the created current variable EDUHGHEDT was assigned a bad data code. This edit was designed to allow for situations where a respondent might be on an accelerated track.
- If the current grade was more than two grade levels higher than the reported highest completed grade and it was lower than the expected current grade, then the value was retained for the current grade. The variable EDUHGHEDT was assigned a bad data code. In other situations, both EDUHGHEDT and EDUSCHGRD (i.e., the edited current grade) were assigned codes of bad data. The rationale for the first edit was that, if EDUSCHGRD was lower than the expected current grade, EDUSCHGRD would be more consistent with the *expected* current grade and the respondent's age than what the reported highest completed grade would be.

5.2.6.5 Edits Involving the Last and Current Grade for Respondents Aged 18 or Older

Editing of EDUHGHEDT and EDUSCHGRD for adult respondents aged 18 or older took into account the changes beginning in 2015 to levels 13 to 20 in question QD11 to refer to the receipt of a high school diploma, college degree, or partial completion of the requirements for a diploma or degree; prior to 2015, levels in QD11 beyond the 11th grade referred to the completion of the 12th grade or a number of years of college beyond the 12th grade. Editing of EDUHGHEDT and EDUSCHGRD for adult respondents also took into account the respondents' current age. For example, if a respondent aged 18 or 19 reported completing the 10th grade or lower and also reported that he or she was currently in the first or second year of college, then EDUHGHEDT was coded as bad data, because being in the first or second year of college was consistent with an age of 18 or 19.

5.2.6.6 Miscellaneous Edits Involving the Current Education Variables

[Table B.1](#) discusses additional editing of variables in the current education section other than those pertaining to current enrollment and grade level. These included situations in which the data were consistent with respondents' status as going to school but other potential inconsistent responses were observed, such as responses for absences from school in the past 30 days.

5.2.7 Military Families

Questions QD10D and QD10E were added to the NSDUH questionnaire in 2015 to identify whether respondents had immediate family members who were currently serving in the U.S. military (i.e., individuals serving in one of the branches of the U.S. Armed Forces, in a Reserve component, or in the National Guard). Questions about family members who are currently serving in the U.S. military are relevant to a SAMHSA strategic initiative to support the well-being of military families.

These questions underwent further revision for 2016. The revised questions for 2016 are shown below. In QD10D, a list of specific family members was presented to respondents for 2016, and respondents were asked to consider only those family members. Respondents then were asked whether any of these family members were currently serving in the United States military. In 2015, respondents were asked in QD10D if anyone in their "immediate family" was current serving.

As in 2015, QD10E in 2016 was asked if respondents reported in QD10D that any of their family members were currently serving in the United States military. In 2016, however, respondents who were asked QD10E could choose only those family members that were listed in QD10D. In particular, "Another member of my immediate family" (category 9 in question QD10E in 2015) and the associated "OTHER, Specify" question QD10F were deleted in 2016.

QD10D For this question, please consider only the family members listed below, whether or not they live with you.

- Your spouse
- Your unmarried partner
- Your mother
- Your father
- Your son or sons
- Your daughter or daughters
- Your brother or brothers
- Your sister or sisters

Please include only these family members, whether they are biological, step or adoptive. Are any of these family members currently serving in the United States military? Press F2 to see and hear a definition of "military."

- 1 Yes
- 2 No
- DK/REF

IF F2 IS PRESSED, PLEASE DISPLAY:

To serve in the military is to be employed as a member of one of the branches of the United States Armed Forces, Reserve components or National Guard, including the Coast Guard.

Press Enter to close this box. Then, type in your response to the question.

QD10E [IF QD10D = 1] Which member or members of your family are currently in the United States military?

To select more than one category, press the space bar between each category you select.

- 1 Your spouse
- 2 Your unmarried partner
- 3 Your mother
- 4 Your father
- 5 Your son or sons
- 6 Your daughter or daughters
- 7 Your brother or brothers
- 8 Your sister or sisters

These changes were determined to have affected the comparability of the military family data between 2015 and 2016. Therefore, the military family variables were renamed for 2016. [Table 5.3](#) shows the names of the edited variables that corresponded to the questions for military family members in 2015 and 2016. The edited variables MILOTHFAM and MILOTHREL from 2015 (corresponding to QD10E9 and QD10F, respectively) were no longer included for 2016.

Table 5.3 Names in 2016 and 2015 for Military Family Variables

Question	Measure	2016 Name	2015 Name
QD10D	Any immediate family member currently serving in the U.S. military	MILTFAMILY	MILFAMILY
QD10E1	Spouse is currently serving in the U.S. military	MILTSPOUS	MILSPOUSE
QD10E2	Unmarried partner is currently serving in the U.S. military	MILTPARNR	MILPARTNR
QD10E3	Mother is currently serving in the U.S. military	MILTMOTHR	MILMOTHER
QD10E4	Father is currently serving in the U.S. military	MILTFATHR	MILFATHER
QD10E5	Son or sons is/are currently serving in the U.S. military	MILTSON	MILSON
QD10E6	Daughter or daughters is/are currently serving in the U.S. military	MILTDAUGH	MILDAUGHT
QD10E7	Brother or brothers is/are currently serving in the U.S. military	MILTBROTH	MILBROTHR
QD10E8	Sister or sisters is/are currently serving in the U.S. military	MILTSISTR	MILSISTER

As noted previously, if respondents answered QD10D as "yes," they were asked to report which members of their immediate family (as specified in QD10D) were currently serving in the U.S. military. Because QD10E was an "enter all that apply" question, separate variables were created for the family relationships in the question. Codes of 1 or 2 (Yes, or No not entered, respectively; Section 2.4.4) were assigned to the edited variables corresponding to QD10E. If respondents answered QD10D as "no," then legitimate skip codes were assigned to MILTSPOUS through MILOTHREL (Section 2.4.2).

Because "OTHER, Specify" data were not available in 2016 for editing the military family relationship variables in QD10E, no respondents were assigned codes of 3 (Yes LOGICALLY ASSIGNED) in MILTSPOUS through MILTSISTR beginning in 2016. In comparison, more than 40 respondents in 2015 did not report that their fathers were currently serving in the military, but it was logically inferred that their fathers were currently serving based on "OTHER, Specify" data from QD10F.

It also was no longer possible in 2016 for respondents to report that the only member of their immediate family who was currently serving in the U.S. military was "another family member" and then to specify in QD10F that "no one else" was serving. Consequently, no respondents were assigned codes of 4 (No LOGICALLY ASSIGNED) in MILTFAMILY beginning in 2016. However, this edit was implemented for only a small number of respondents (i.e., fewer than five) in 2015.

Two additional edits were implemented for MILTMOTHR, MILTFATHR, MILTSON, and MILTDAUGH when respondents' ages were considered to be inconsistent with their reports of one of these family members currently serving in the U.S. military.

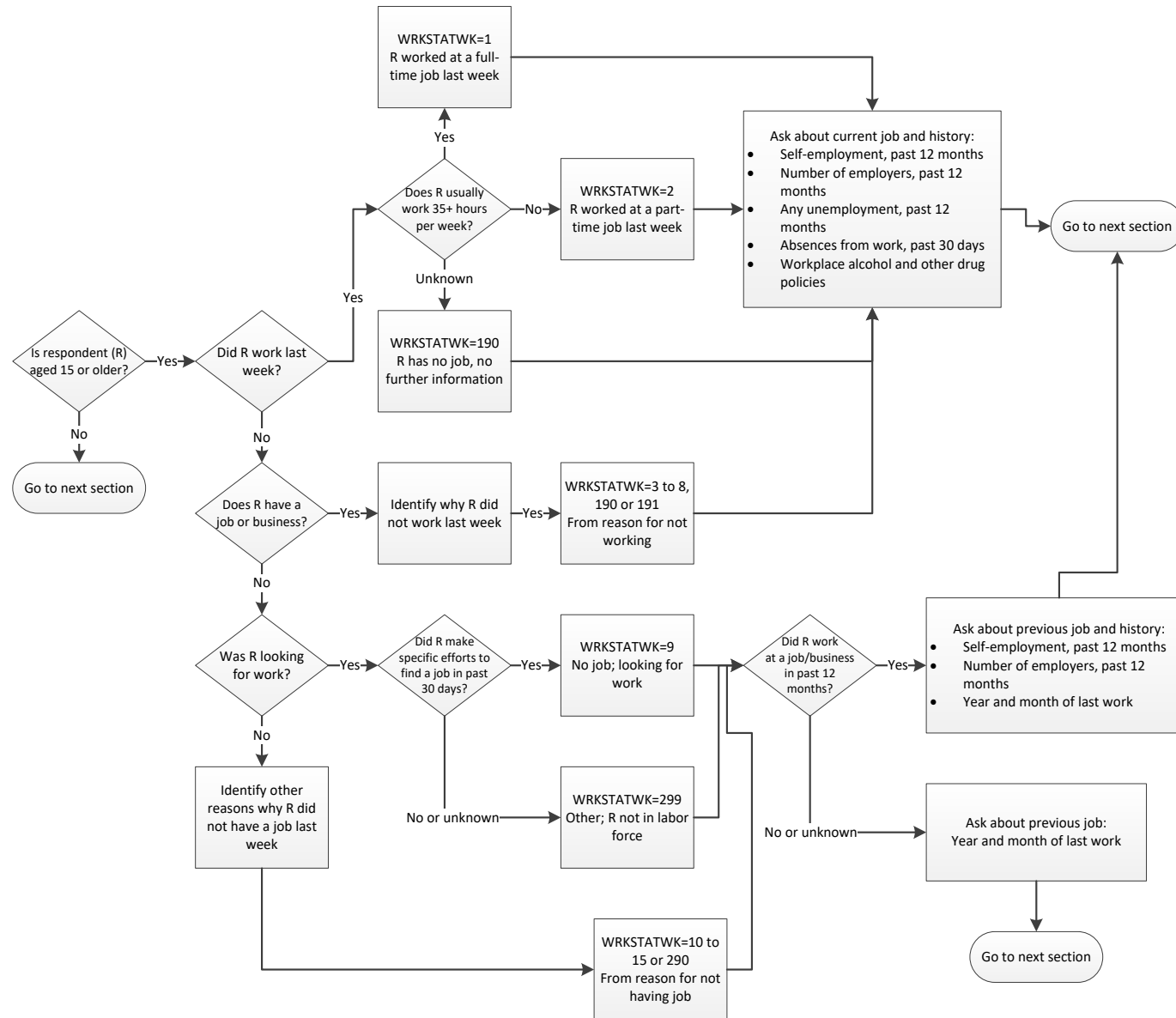
- If respondents were aged 50 or older and reported that their mothers or fathers were currently serving, then MILTMOTHR or MILTFATHR were assigned codes of 85 for bad data, as applicable.
- If respondents were aged 12 to 25 and reported that sons or daughters were currently serving, then MILTSON or MILTDAUGH were assigned codes of 85, as applicable.

5.2.8 Employment and Workplace

As mentioned previously, the employment questions were self-administered ACASI questions in 2016. Industry and occupation questions for respondents' current or most recent employment also were no longer included in the questionnaire beginning in 2015. Otherwise, the basic logic for administering the employment questions in 2016 was the same as in 2014.

Respondents aged 15 or older were asked questions about their current employment, employment history, and characteristics of their workplace (if applicable). Question QD26 asked whether respondents worked in the week prior to the interview. If respondents reported that they did not work in the past week, they were asked in question QD27 whether they had a job or business. Respondents then were routed through different branches of work-related questions depending on how they answered these two key questions, as shown in [Figure 5.2](#). For example, respondents who worked in the past week were asked questions to determine full-time or part-time work status (e.g., whether they usually worked 35 or more hours per week), whether they ever had a period of unemployment in the past 12 months, the number of days they missed work in the past 30 days because they were sick or because they did not want to be at their workplace, and characteristics of their workplace, particularly with respect to alcohol and other drug policies at their workplace. Similarly, respondents who did not work in the past week and did not have a job were routed into questions relevant for people who currently were not working, such as why they did not have a job, whether they made specific efforts to find work in the past 30 days, and the month and year when they last worked for pay.

Figure 5.2 Routing Logic for Respondents in the Employment Status Section



The employment and workplace questions and logic underwent important changes prior to the start of the new NSDUH baseline in 2002 and in 2005 for the question about the respondent's month of last employment (question QD39B). Documentation of historical changes to the employment and workplace questions can be found in the report on editing of interviewer-administered data that was prepared for the 2011 NSDUH MRB (Kroutil, Chien, Handley, & Bradshaw, 2013). Documentation of the change to the logic for QD39B can be found in the 2014 NSDUH MRB editing and imputation report (Center for Behavioral Health Statistics and Quality, 2016a).

An important aspect of editing the work-related variables involved identification of situations where questions had been legitimately skipped (Section 2.4.2). A second key aspect of processing the work-related variables was to use the data to establish respondents' current work status (Figure 5.2). As noted previously, a single, recoded work status variable named WRKSTATWK was created that served as the starting point for creation of a simplified edited variable (EDWRKSTAT) and a final, statistically imputed employment status variable (IRWRKSTAT). WRKSTATWK was created from the following final variables: WRKDPSTWK (whether the respondent worked in the past week), WRKHADJOB (whether the respondent had a job if he or she did not work in the past week), WRK35WKUS (whether the respondent usually worked 35 or more hours per week), WRKRSNNOT (reason for not working in the past week despite having a job), WRKRSNJOB (reason for not having a job in the past week), WRKEFFORT (made specific efforts to find work in the past 30 days), and WRKDPSTYR (whether the respondent had a job in the past 12 months). Based on the data in these variables, respondents aged 15 or older were assigned to one of the categories in WRKSTATWK that are listed in Table 5.4.

Table 5.4 Categories of WRKSTATWK

Code	Employment Situation	Code	Employment Situation
1	Worked at full-time job, past week	11	No job: keeping house full time
2	Worked at part-time job, past week	12	No job: in school/training
3	Has job but out: vacation/sick/temp absence	13	No job: retired
4	Has job but out: layoff, looking for work	14	No job: disabled for work
5	Has job but out: layoff, not looking for work	15	No job: didn't want a job
6	Has job but out: waiting to report to new job	190	Has full-time job, reason for not working unknown
7	Has job but out: self-employed, no business past week	191	Has part-time job, reason for not working unknown
8	Has job but out: in school/training	199	Has job, no further information
9	No job: looking for work	290	No job, no further information
10	No job: layoff, not looking for work	299	Other, not in labor force

NOTE: Additional codes for WRKSTATWK in the 900 series have their standard meanings in NSDUH: Don't know (994), Refused (997), Blank (998), and Legitimate skip (999).

In addition, respondents who reported in question QD31 that they did not have a job but were looking for work were not classified as being unemployed unless they reported in WRKEFFORT that they had made specific efforts in the past 30 days to find work (such as making contacts with someone about a job, sending out resumes or job applications, or placing or answering ads). If respondents reported that they did not have a job but were looking for work but WRKEFFORT was not answered as "yes," they were classified as not in the labor force (code 299) in WRKSTATWK.

If respondents did not know or refused to report whether they worked in the past week, WRKDPSTYR was checked for indications of whether respondents worked in the past year. Respondents who indicated in WRKDPSTYR that they did not work in the past 12 months were classified as not having a job (WRKSTATWK = 290). Otherwise, if respondents did not provide information on whether they worked in the past week (i.e., QD26 answered as "don't know" or "refused"), WRKSTATWK was assigned the corresponding code of "don't know" or "refused."

Table B.2 in Appendix B describes edits for employment variables pertaining to whether respondents had a job in the past week, the number of days absent from work in the past 30 days, the number of weeks without a job in the past 12 months, the number of hours worked in the past week, the year and month that respondents last worked for pay, and self-employment. For example, if respondents reported in question QD39A that they never worked for pay, interviewers were instructed to enter a response of 9991. When the month question QD39B had been skipped because a response of 9991 had been entered in QD39A, the edited month variable WRKLASTMO was assigned a code of 91. Documentation for 9991 (or 91) was as follows: 9991 = NEVER WORKED AT A JOB OR BUSINESS.

A refinement to the editing procedures for the employment and workplace section also has been implemented since 2003 for respondents who reported that they did not work in the past week (WRKDPSTWK = 2). The variable pertaining to the number of hours that respondents worked in the past week (WRKDHRSWK) was assigned a legitimate skip code regardless of how respondents answered the question about having a job (QD27). Prior to 2003, a legitimate skip code was assigned to WRKDHRSWK only if respondents reported that they did not work in the past week (WRKDPSTWK = 2) and QD27 was answered as "no"; the prior logic did not assign legitimate skip codes to WRKDHRSWK if respondents answered QD27 as "don't know" or "refused." Logically, however, if respondents reported that they did not work in the past week, they would not have worked any hours at a job during that period, regardless of how they answered question QD27.

A final procedure that is discussed for the editing of the employment variables (not presented in Table B.2) concerns the creation of the base variable EDWRKSTAT, which was used to create the imputation-revised employment status variable. EDWRKSTAT was derived from WRKSTATWK and WRK35WKUS. WRK35WKUS was used in some situations to determine whether employed respondents were employed full time or part time. Specifically, EDWRKSTAT was derived in the following manner:

EDWRKSTAT =

- 99, if the respondent is 12 to 14 years old; else
- 1 (full time), if WRKSTATWK = 1 or 190, or if WRKSTATWK = 3, 6, 7, 8, or 199 and WRK35WKUS = 1; else
- 2 (part time), if WRKSTATWK = 2 or 191, or if WRKSTATWK = 3, 6, 7, 8, or 199 and WRK35WKUS = 2; else
- 3 (unemployed), if WRKSTATWK = 4, 5, 9, or 10; else
- 4 (other), if WRKSTATWK = 11, 12, 13, 14, 15, 290, or 299; else

- 5 (part time or full time), if WRKSTATWK = 3, 6, 7, 8, or 199 and WRK35WKUS was missing (i.e., greater than 2); else
- missing.

5.3 Imputation for Self-Administered Demographics Variables

For the self-administered demographics section of the 2016 NSDUH, three imputation-revised variables were created from the base variables EDWRKSTAT, IMMBORNUS, and ENTRYAG2: the first was an employment status variable, IRWRKSTAT; the second, IRIMMBORNUS, was an indicator of whether the respondent was born in the United States; and the third, IRIMMENTAGE, recorded the age at which the respondent entered the United States. As with their edited counterparts, the names for each of these imputation-revised variables were new beginning in 2015. The name changes signify that the variables were not comparable with previous years due to questionnaire changes. Beginning in 2015, employment status moved to the ACASI self-administered section, and the immigrant status variables had a change in skip logic.

These variables were processed in three separate, single-member imputation sets, using the single response propensity (RP)/single prediction (PRD) type of the predictive mean neighborhood (PMN) method described in Chapter 3. Employment status and the born in the United States indicator have item response rates around 99 percent. Age at which the respondent entered the United States has a response rate around 87 percent. See [Table A.26](#) in Appendix A for details on the rates of missingness among the self-administered demographics variables.

5.3.1 Employment Status (Imputation Set 1)

The first self-administered demographic variable that underwent imputation was the employment status variable. The imputation process was straightforward except for one feature: the handling of cases with EDWRKSTAT = 5, where it was known that the respondents were employed but it was not known whether they were employed part time versus full time. These cases were handled in the hot-deck step in a separate missingness pattern, with a single logical constraint and a modified predictive mean vector. The final imputation-revised variable IRWRKSTAT had five levels: employed full time, employed part time, unemployed, other, and a skip code for respondents aged 12 to 14. Two age groups were used for RP and PRD modeling: 15 to 25 and 26 or older. The hot-deck step was implemented separately for three age groups: 15 to 17, 18 to 25, and 26 or older.

5.3.1.1 Response Propensity Step

The response propensity models for imputation set 2 utilized the preliminary analysis weight, PANALWT. For the first RP model, the domain included all unit respondents aged 15 to 25. For the second RP model, the domain included all respondents aged 26 or older. In both situations, item respondents were those with EDWRKSTAT values of 1, 2, 3, or 4. See [Table D.1](#) in Appendix D for details of the covariates used in the RP models for this variable.

5.3.1.2 Prediction Step

Using the adjusted weights that are outputs of the RP step, the employment status variable was modeled using polytomous logistic regression as implemented by the MULTLOG procedure in SUDAAN. For both age groups, the outcome variable had four levels, which mapped to the first four levels of EDWRKSTAT.

5.3.1.3 Hot-Deck Step

The predicted means from the prediction step play a central role in the donor selection algorithm applied in the hot-deck step, but unlike the RP and PRD steps, the hot-deck steps were run separately within three age groups: 15 to 17, 18 to 25, and 26 or older. This was done to allow for parallel processing, which decreases the time required for implementation. Each item nonrespondent in the hot-deck step was assigned one of two missingness patterns. Item nonrespondents with a missing value for EDWRKSTAT were handled in the first missingness pattern, which used the full predictive mean vector and no logical constraints. Item nonrespondents with EDWRKSTAT = 5 were handled in the second missingness pattern, which applied a logical constraint to ensure that the donor was employed (either full time or part time). Also, conditional probabilities were used to take advantage of the partial information that was available. Instead of using the model's predicted probabilities directly, a single predicted mean was derived using a conditional probability, which was the probability that the respondent was employed full time, given that the respondent was employed. In addition to the segment and delta likeness constraints, a third likeness constraint (i.e., donor's age must be within 5 years of recipient's age) was applied in the hot-deck step. See Appendix E for more details on missingness patterns and constraints for employment status.

5.3.1.4 Recodes for Additional Analyses

IRWRKSTAT18 was a direct recode of IRWRKSTAT and AGE. For respondents who were younger than 15 or older than 17, IRWRKSTAT18 and IRWRKSTAT were equivalent. For 15- to 17-year-olds, responses for IRWRKSTAT were overwritten with a code indicating that the respondent was too young to have his or her employment status recorded for the variable. This was the same code that was used for 12- to 14-year-olds for IRWRKSTAT (and IRWRKSTAT18).

5.3.2 Immigrant Status, Born-in-U.S. Indicator (Imputation Set 2)

The second self-administered demographic variable that underwent imputation was the born-in-U.S. variable, IMMBORNUS. This was a dichotomous variable with very few missing responses. The RP, PRD, and hot-deck steps were all run separately within three age groups: 12 to 17, 18 to 25, and 26 or older. The imputation procedure was straightforward and is described in the next three sections.

5.3.2.1 Response Propensity Step

The response propensity models for imputation set 3 utilized the preliminary analysis weight, PANALWT. The domain indicator for the RP model included all unit respondents. Item

respondents were those with a nonmissing value for IMMBORNUS. See [Table D.1](#) in Appendix D for details of the covariates used in the RP models for this variable.

5.3.2.2 Prediction Step

Using the adjusted weights that are outputs of the RP step, the born-in-U.S. indicator was modeled using dichotomous logistic regression as implemented by the LOGISTIC procedure in SUDAAN.⁷¹ The outcome variable was IMMBORNUS. The single predicted mean was the predicted probability that the respondent was born in the United States.

5.3.2.3 Hot-Deck Step

The hot-deck step for the born-in-U.S. indicator included a single predicted mean from the prediction step, no logical constraints, and only the segment and delta likeness constraints. IMMBORNUS was the base variable for imputation and the imputation-revised version was called IRIMMBORNUS. Details on the hot-deck step are available in Appendix E.

5.3.3 Immigrant Status, Age of Entry (Imputation Set 3)

The age of entry variable was created only for respondents who were not born in the United States as defined by IRIMMBORNUS. This results in a small domain. As a result, imputations were not conducted separately within age groups. The models were likely to be improved when age groups were combined because (1) none of the response categories were sparsely populated, leading to more robust models and (2) sufficiently large donor pools were ensured in the hot-deck step. Details on the procedures applied to the age of entry variable are explained in the next three sections.

5.3.3.1 Response Propensity Step

The response propensity model for imputation set 4 utilized the preliminary analysis weight, PANALWT. The domain indicator for the RP model included all respondents who were not born in the United States as defined by IRIMMBORNUS. Item respondents were those domain members with a nonmissing value for ENTRYAG2. See [Table D.1](#) in Appendix D for details of the covariates used in the RP model for this variable.

5.3.3.2 Prediction Step

Using the adjusted weights that are outputs of the RP step, the predicted mean for an immigrant's age of entry was estimated using a linear regression model, as implemented by the REGRESS procedure in SUDAAN. To control the upper and lower bounds of predicted means for age of entry, it was necessary to perform a logit transformation on the response variable. The response variable in the model was the immigrant age at entry as a proportion of the continuous version of current age IMMCONTAGE, as described in Section 5.2.3. The expression of the proportion is $P_i = Y_i/N_i$, where $Y_i = \text{Age at Entry}_i$ and $N_i = \text{Continuous Age}_i$ (IMMCONTAGE).

⁷¹ In SAS-callable SUDAAN, this is the RLOGIST procedure to avoid confusion with SAS's own LOGISTIC procedure.

After the weight adjustment, the following empirical logit transformation was used as the response variable in a weighted linear univariate regression:

$$\log\left[(Y_i + 0.5)/(N_i - Y_i + 0.5)\right].$$

This transformation was nearly equivalent to the standard logit transformation,

$$Y_i^* = \log\left[P_i/(1 - P_i)\right],$$

which was not used because this transformation is unstable for respondents who entered the country at their current age (such that $P_i = 1$).

5.3.3.3 Hot-Deck Step

Two logical constraints were utilized in the hot-deck step for the age of entry variable. Both involved the respondent's age. One required that the donor's age of entry be less than or equal to the recipient's current age. The other required that the difference between the recipient's current age and the donor's age of entry be less than 1 year if the recipient lived in the United States for less than 1 year (as indicated by QD16A) or greater than 1 year if the recipient lived in the United States for more than 1 year. The only base variable was ENTRYAG2 and its imputation-revised counterpart was IRIMMENTAGE. The segment and delta likeness constraints were applied in the hot-deck step. Details on the hot-deck step are available in [Tables E.17 through E.19](#) in Appendix E.

6. Editing and Imputation for the NSDUH Substance Use Variables

6.1 Introduction

This chapter provides documentation of procedures for the editing and statistical imputation (subsequently referred to as "imputation") of variables in the 2016 National Survey on Drug Use and Health (NSDUH) that were used to estimate substance use in the lifetime, past year, and past month periods for NSDUH tables and reports.⁷² As a prerequisite for reviewing this chapter, readers are encouraged to familiarize themselves with the content and terminology in Chapters 1 through 3 of this report because these three chapters provide an overview to the NSDUH interview and the general approach to editing and imputation of NSDUH data, with information on specific processes that are common across the survey in Chapters 2 and 3.

For all substance use sections (i.e., tobacco through the prescription psychotherapeutic drugs), variables corresponding to questions for key substance use measures first underwent editing to (1) replace missing values with codes that indicated that the questions did not apply, (2) replace missing values with nonmissing values that could logically be inferred from other data, (3) make logical inferences when answers to related questions were inconsistent, and (4) identify inconsistent or ambiguous responses to be resolved through imputation. Except where indicated, edited variables for these measures then underwent imputation to replace missing values with nonmissing values or to replace ambiguous responses (e.g., use at some point in a respondent's lifetime but no definite period for most recent use) with specific ones.

6.1.1 Edited (but Not Imputed) Substance Use Variables

Some substance use variables in 2016 underwent editing but no further imputation. Many of these variables pertain to prescription psychotherapeutic drugs and tobacco products. In particular, the following groups of prescription psychotherapeutic drug variables were edited but did not undergo further imputation:

- any use or misuse (if applicable) in the past 12 months of specific prescription drugs;
- the age at first misuse (AFM), year of first misuse (YFM), and month of first misuse (MFM) for specific prescription drugs that were misused in the past 12 months;
- misuse of any prescription psychotherapeutic drug within an overall category (e.g., pain relievers) within the past 30 days when the respondent was drinking alcohol or within a couple of hours of when the respondent drank alcohol;

⁷² The past year refers to the period within 12 months of the interview date (i.e., past 12 months), and the past month refers to the period within 30 days of the interview date (i.e., past 30 days). Past month use also is referred to as "current" use in NSDUH reports.

- ways in which the respondent misused prescription psychotherapeutic drugs within an overall category in the past 12 months (e.g., without a prescription of the respondent's own);
- the specific prescription drug within a given category that a respondent last misused in the past 12 months;
- use of a needle to inject prescription stimulants in the past 12 months or past 30 days;
- reasons for misusing the last prescription drug within a given category in the past 12 months; and
- how the respondent (or a friend or relative of the respondent, if applicable) obtained the last prescription drug within a given category that the respondent misused in the past 12 months.

In addition, Section 6.2.7 discusses editing procedures for tobacco variables that did not undergo further imputation. Also, the variable corresponding to the usual number of drinks of alcoholic beverages that respondents consumed on days when they used alcohol in the past 30 days (corresponding to question AL07) is used in editing other variables that subsequently undergo imputation (e.g., frequency of consumption of four or more drinks for females or five or more drinks for males on a single occasion in the past 30 days, also referred to as the frequency of "binge" alcohol use) but the variable corresponding to question AL07 itself does not undergo imputation. Editing procedures for the variable corresponding to question AL07 are discussed in Section 6.2.4.7.

6.1.2 "Parent" and "Child" Drug Use Categories

In this chapter, some of the substance use measures refer to a general drug category (e.g., hallucinogens), and other measures refer to one or more subcategories within a general category (e.g., LSD, PCP, or Ecstasy as types of hallucinogens). These drug categories are described using the terms "parent drug" for the general drug category and "child drug" for the drug subcategory. Parent/child drug pairs often (but not always) occurred in sections that included multiple gate questions.⁷³ For example, the hallucinogens section includes questions about lifetime use of specific hallucinogens to determine the respondent's overall lifetime use or nonuse of any hallucinogen. Included in these questions are specific gate questions for LSD (LS01A); PCP (LS01B); Ecstasy (LS01F); ketamine (LS01I); the tryptamine hallucinogens DMT, AMT, or "Foxy" (LS01J); and *Salvia divinorum* (LS01K). However, parent/child drug pairs also could appear in separate sections, such as for any use of cocaine and specific use of crack cocaine (i.e., if respondents reported lifetime use of any form of cocaine). Table 6.1 shows the drugs with parent/child relationships in 2016 and the data that were collected for them.

⁷³ See Section 2.2.3 for a discussion of the gate question terminology.

Table 6.1 Substances in the 2016 NSDUH with a Parent/Child Relationship

Parent Drug	Child Drugs	Estimates from Parent Data	Estimates from Child Data	"Other" Lifetime Use Indicator¹
Cocaine	Crack	Initiation, recency, 12-month frequency, 30-day frequency	Initiation, recency, 12-month frequency, 30-day frequency	No
Hallucinogens	LSD, PCP, Ecstasy, Ketamine, DMT/AMT/"Foxy," <i>Salvia divinorum</i>	Initiation, recency, 12-month frequency, 30-day frequency	Initiation (LSD, PCP, and Ecstasy only), recency (all child drugs)	Yes

¹ See Section 6.4.1.

Although questions were asked in 2016 for the use and misuse of the specific pain reliever OxyContin[®] in the past 12 months, the relationship between any pain reliever and OxyContin[®] is not considered to be a parent/child relationship beginning in 2015 because respondents were not asked directly about the use or misuse of "any" prescription pain reliever in the past 12 months. Rather, the use or misuse of any prescription pain reliever was determined by reviewing the data for the use or misuse of all prescription pain relievers in the questionnaire.

In addition, methamphetamine prior to 2015 was in the section of the interview for prescription stimulants and was considered to be a "child" drug for prescription stimulants (i.e., the "parent" drug). However, methamphetamine is no longer considered to be a "child" drug for prescription stimulants because of the presence of a separate interview section for methamphetamine beginning in 2015. Consequently, the types of parent/child edits that applied to the substances listed in [Table 6.1](#) no longer apply to methamphetamine and prescription stimulants.

6.2 Editing Procedures for the Substance Use Variables Other than Prescription Drugs

Consistent with prior years, the substance use measures collected in the 2016 NSDUH for tobacco, alcohol, marijuana, cocaine, crack cocaine, heroin, hallucinogens, inhalants, and methamphetamine included (but were not limited to) lifetime use (or nonuse); initiation of use (i.e., age at first use [AFU], and for recent initiates, the month of first use [MFU] and year of first use [YFU]); most recent use; frequency of use in the past 12 months (for alcohol, marijuana, cocaine, crack cocaine, heroin, hallucinogens, inhalants, and methamphetamine); and frequency of use in the past 30 days (for tobacco products except for pipe tobacco, alcohol, marijuana, cocaine, crack cocaine, heroin, hallucinogens, inhalants, and methamphetamine). This section discusses procedures for editing the substance use variables for tobacco (i.e., cigarettes, smokeless tobacco, cigars, and pipe tobacco), alcohol, marijuana, cocaine in any form, crack cocaine, heroin, hallucinogens, inhalants, and methamphetamine.

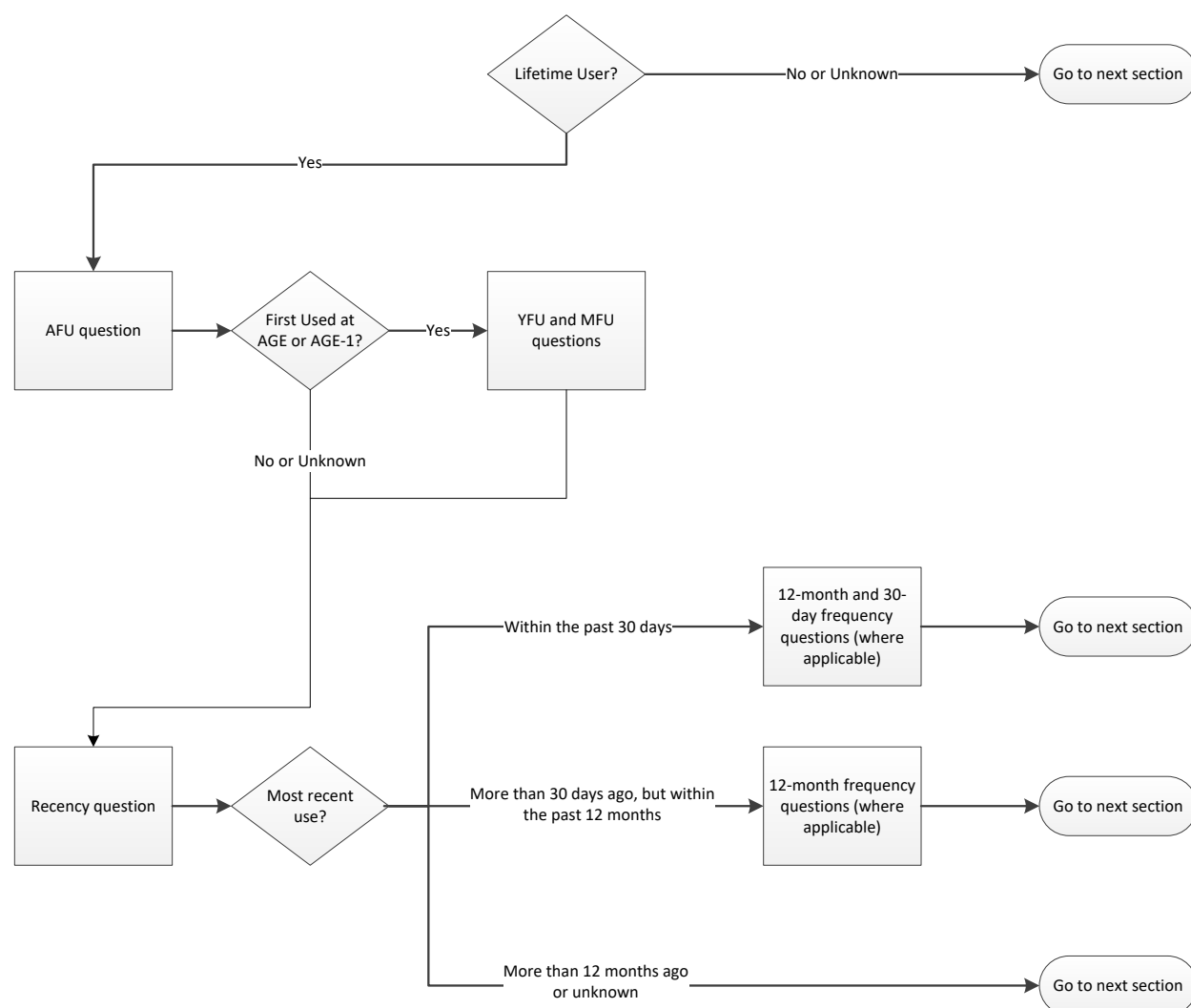
Figure 6.1 provides an overview of the general routing logic for the substance use sections of the interview other than for prescription drugs.⁷⁴ Edits for the substance use variables in these sections encompassed the following key activities:

- edits of the lifetime use variables (i.e., gate questions), where respondents indicated whether they have ever used the substance of interest (Section 6.2.1);
- edits of the recency-of-use variables, where respondents who indicated lifetime use of the substance indicated when they last used it (Section 6.2.2);
- edits involving users of only "child" drugs (Section 6.2.3);
- edits of the 12-month and 30-day frequency variables, where respondents who indicated use of a substance in the 12 months or 30 days prior to the interview indicated the number of days they used that substance in the period of interest (Section 6.2.4);
- edits for variables pertaining to the initiation of use, where respondents who indicated lifetime use of a substance indicated when they first used it (Section 6.2.5);
- edits for lifetime daily cigarette use (Section 6.2.6);
- edits for substance use variables that do not undergo imputation (Section 6.2.7); and
- edits that were applied as part of the imputation processing (Section 6.4).

In connection with each of these edits, the discussion in the remainder of Section 6.2 focuses on relevant issues or inconsistencies in the data that needed to be addressed through logical editing. As noted in Chapter 1, the skip logic in the computer-assisted interviewing (CAI) instrument limited the chances for respondents to be routed to questions where they could give answers that were inconsistent with their answers to previous questions. For example, if respondents reported never using marijuana in question MJ01, the CAI program did not ask additional questions that would presume use of marijuana at least once. However, the CAI program did not completely eliminate all opportunities for respondents to provide inconsistent answers. Nevertheless, most processing of the CAI data was relatively straightforward, and the issues discussed in this section were not widespread relative to the total number of 67,942 respondents in 2016.

⁷⁴ For additional information about the content and logic in the drug section of the questionnaire, see the computer-assisted interviewing (CAI) specifications in the 2016 NSDUH methodological resource book at <https://www.samhsa.gov/data/>.

Figure 6.1 Routing Logic for Respondents in the Substance Use Sections Other than for Prescription Drugs



AFU = age at first use; MFU = month of first use; YFU = year of first use.

6.2.1 Edits of Lifetime Use Variables

As discussed in connection with the usable case criteria (Section 2.2.3), respondents were asked first whether they had ever used tobacco products (cigarettes, smokeless tobacco, cigars, or pipe tobacco), alcohol, marijuana, cocaine (and crack cocaine, if they had ever used cocaine in any form), heroin, hallucinogens, inhalants, or methamphetamine. For hallucinogens and inhalants, respondents were asked a series of questions to establish whether they had ever used one or more specific types of drugs within that category (e.g., LSD, PCP, peyote). Only those respondents who indicated lifetime use of that drug (or lifetime use of one or more specific drugs within the respective sections for hallucinogens or inhalants) were asked more detailed questions about that substance, or any substance in that category (including situations in which respondents initially refused to answer a question about their lifetime use of a substance but then changed their answer to "yes" on follow-up; see Sections 2.2.3 and 2.4.2). For hallucinogens and inhalants, however, there was no overall question about lifetime use or nonuse of any drug in

these categories. Rather, lifetime use of any drug in the overall category for these two sections could be deduced from one or more affirmative answers regarding lifetime use of any of the specific drugs in that category. Because the sections for hallucinogens and inhalants included a question about the lifetime use of "any other" drug in that category, answers of "no" for lifetime use of all of the specific drugs in that category (i.e., including lifetime use of "any other") logically indicated that the respondent had never used any drugs in that category.

Processing of the gate variables for tobacco through methamphetamine established whether (1) respondents had used a substance of interest at least once (or any substance in that category, in the case of hallucinogens or inhalants), (2) they had never used the substance (or any substance in that category), or (3) lifetime use or nonuse of the substance could not be determined. In addition to answering these gate questions as "yes" or "no," respondents could answer them as "don't know" or "refused." As noted in Section 2.4.2, final responses of "don't know" and "refused" were treated by the CAI skip logic as equivalent to situations where respondents never used the drug of interest. For hallucinogens and inhalants, the CAI skip logic treated combinations of responses of "no," "don't know," and "refused" to the questions about individual drugs in the absence of any affirmative response to these questions in the same manner as if the respondent had answered all of these questions negatively. In these situations where a gate question was answered as "don't know" or "refused," the respondent's lifetime use or nonuse status was treated as unknown because these responses did not provide conclusive information one way or the other.⁷⁵ Cases with unknown lifetime use/nonuse status were subsequently imputed to be lifetime users or nonusers, as described in Section 6.5.

This step of the editing procedures also involved assignment of "bad data" codes to lifetime drug use variables (i.e., equivalent to missing data) if potential patterned responses previously had been identified (Sections 2.3.2 and 2.4.3). For the substance use sections pertaining to smokeless tobacco through heroin and for methamphetamine,⁷⁶ these edits involved assignment of codes for bad data for the lifetime drug question and the follow-up probe, if respondents previously had refused the lead question (e.g., questions AL01 and ALREF for alcohol). For hallucinogens and inhalants, these edits involved assignment of codes for bad data to all lifetime variables in that section (e.g., data in questions LS01A through LS01K for hallucinogens, including question LS01H for any other hallucinogen and associated "OTHER, Specify" data for hallucinogens).

6.2.1.1 Creation of Lifetime Use Variables for Tobacco Products through Heroin and for Methamphetamine

Creation of edited variables for lifetime use or nonuse of tobacco products through heroin and for methamphetamine typically was straightforward because there were overall questions about lifetime use or nonuse. In most instances, therefore, codes corresponding to respondents' original answers simply were assigned to the relevant variable (e.g., 1 = Yes; 2 = No). When respondents were routed to a follow-up question because they initially refused to answer the first

⁷⁵ For multiple gate drugs, this included situations where respondents answered one or more of the gate questions as "don't know" or "refused" but did not answer any of the other gate questions affirmatively.

⁷⁶ Because of the requirement of the usable case criteria that respondents had to have defined data for lifetime use or nonuse of cigarettes (Section 2.2.3), assignment of the missing value of bad data for lifetime use of cigarettes would result in a case being reclassified as not usable.

question about lifetime use or nonuse (Section 2.4.2), their response to this follow-up question was taken as the final response for the lifetime use variable.

The exception to these procedures involved lifetime use or nonuse of crack cocaine. Specifically, the logic to question CC01, regarding lifetime use of any form of cocaine, governed whether respondents were asked questions about crack cocaine. Therefore, the lifetime crack use variable CRKEVER (corresponding to question CK01) was assigned a code of 91 (i.e., never used cocaine) when COCEVER (corresponding to CC01) was answered as "no." If respondents reported lifetime use of cocaine but answered question CK01 as "no," then CRKEVER was assigned a code of 2.

6.2.1.2 Creation of Lifetime Use Variables for Hallucinogens and Inhalants

For the variables in the hallucinogens and inhalants sections that indicated lifetime use or nonuse of specific drugs, if respondents answered all gate questions in the series as "no," the edit procedures assigned a code of 91 to the entire series of variables (e.g., LSD through any other hallucinogen).⁷⁷ This was done to indicate that the respondents had never used any of the drugs in that category. In contrast, a code of 2 (i.e., "no") in edited variables for lifetime use of specific drugs within a broader category had the following meaning:

- the respondent was a user of at least one drug in the category but had never used the specific drug of interest (e.g., if a respondent was a lifetime user of LSD but not PCP); or
- the respondent reported never using the particular drug of interest but answered other questions in the series as "don't know" or "refused" (e.g., if a respondent did not know whether he or she had ever used LSD but definitely knew that he or she had never used PCP).

Variables also were created to indicate whether respondents had ever used one or more drugs within the overall drug categories. These variables were called HALLUCEVR (ever used any hallucinogen) and INHALEVER (ever used any inhalant). These summary variables could take on the following values:

- 1 ("Yes"). The summary variable for a given category was coded as 1 (i.e., "yes") if respondents answered "yes" for lifetime use of at least one specific drug in the overall category. The summary variable also was coded as 1 if respondents initially refused to answer all lifetime questions within one of these sections and then reported on follow-up that they had used some drug in that category at least once. For example, if at least one affirmative answer appeared in questions about lifetime use of inhalants (including the follow-up question INREF), INHALEVER was coded as 1.
- 91 ("Never Used"). This code of 91 could be assigned to the overall summary variable for one of two reasons: (1) if respondents answered all questions about lifetime use of individual drugs within the category as "no" or (2) if respondents were routed to a follow-up question because they refused to answer *all* questions about

⁷⁷ In the hallucinogens section, this included situations in which respondents initially refused to report whether they had ever used LSD, PCP, or Ecstasy but then changed their answer(s) to "no" on follow-up.

lifetime use of individual drugs within a category and then they reported that they never used that type of drug. As noted previously, this code of 91 also was assigned to all of the individual lifetime use variables within that section. For example, if respondents initially refused to answer all questions about lifetime use of specific inhalants, they were asked whether they had ever used any type of inhalant. If respondents answered this follow-up question as "no," the summary variable INHALEVER was assigned a code of 91, and all individual lifetime inhalant variables were assigned a code of 91 as well.

- 94 ("Don't know"). This edit applied only to the summary variable INHALEVER for inhalants. If respondents were routed to INREF because they refused all of the gate questions for inhalants and they answered INREF as "don't know," then INHALEVER was assigned a final code of 94.
- 97 ("Refused"). The summary variable was assigned a code of 97 if respondents initially refused to answer *all* questions about lifetime use for hallucinogens or inhalants and then continued to refuse when asked the follow-up questions about lifetime use. In this case, the edited lifetime use variables for the individual drugs within that section also retained a code of 97. For example, INHALEVER was assigned a code of 97 if a respondent refused to answer questions IN01A through IN01L for inhalants and also refused to answer the follow-up question INREF.
- 98 ("Blank"). The summary variable was assigned a code of 98 if respondents had one or more answers of "don't know" or "refused" to questions about lifetime use of specific drugs within the category (i.e., other than the previous situation in which respondents refused all questions) and they had no reports of lifetime use of any of these drugs. In this situation, the code of 98 indicated that lifetime use or nonuse for the overall category was unknown, even if one or more questions about lifetime use of specific drugs was answered as "no." For example, INHALEVER was assigned a code of 98 if some questions in IN01A through IN01L were answered as "no" but other questions were answered as "don't know" or "refused." Note that for hallucinogens, respondents were also assigned a code of 98 in the summary variable HALLUCEVR if they answered all of the lifetime questions in LS01A through LS01H as "don't know."

6.2.1.3 Editing of "OTHER, Specify" Drug Variables for Nonusers of Hallucinogens or Inhalants

If respondents had never used any of the drugs in the series of multiple gate questions for hallucinogens or inhalants, they would not have been routed to questions where they could specify the use of some other drug in that overall category. In this situation, blank values in the unedited "OTHER, Specify" drug variables were replaced with a code of 9991 to indicate that the questions were skipped because the respondents had never used that class of drugs.

Similarly, if respondents reported in the preceding multiple gate questions for hallucinogens or inhalants that they used at least one drug in the category but they never used some other drug besides the ones they were asked about, they were legitimately skipped out of the questions that asked them to specify what "other" drug they had used. In this situation, blank

values in the unedited "OTHER, Specify" drug variables were replaced with a code of 9999 (LEGITIMATE SKIP).

Functionally, the codes of 9991 and 9999 both indicate that respondents legitimately skipped out of the "OTHER, Specify" questions in that drug's section. However, the 9991 code provides for analysts the extra level of detail that the respondents were legitimately skipped out of these questions because they had never used anything within that category of drugs.

6.2.1.4 Use of "OTHER, Specify" Drug Data to Edit Lifetime Use Variables for Hallucinogens and Inhalants

When respondents specified the use of other hallucinogens or inhalants, these "OTHER, Specify" data were used to edit related variables *within the same section* (Sections 2.4.1 and 2.4.5). For example, if a respondent did not report lifetime use of a specific drug in a given section (e.g., PCP) but specified using it as another drug in the category (e.g., "marijuana laced with PCP" as another hallucinogen), the respondent was logically inferred to be a lifetime user of that drug within the relevant section where the "OTHER, Specify" response occurred (e.g., PCP for "OTHER, Specify" data within the hallucinogens section). The following code was assigned to the edited lifetime use variable for that drug: 3 = Yes LOGICALLY ASSIGNED.

This edit was especially relevant to the following hallucinogens: LSD; PCP; Ecstasy; ketamine; the tryptamines DMT, AMT, or "Foxy"; and *Salvia divinorum*. For these hallucinogens, lifetime use or nonuse was imputed if respondents had missing data for the lifetime use questions. Consequently, no imputation would be needed for these drugs if a respondent reported not knowing whether he or she had used these drugs or refused to report use of these drugs when asked directly but the "OTHER, Specify" data in that section indicated use.

With few exceptions, however, "OTHER, Specify" data that referred to street names or slang terms were not used in editing the lifetime variables for hallucinogens or inhalants because the meaning of a particular slang term could vary by region, could apply to more than one drug, or could change over time. Use of street or slang terms to infer lifetime use of a specific hallucinogen or inhalant was generally restricted to situations where that term was supplied to respondents as a synonym for that drug. For example, "angel dust" was listed in question LS01B as an alternative term for PCP. Thus, situations where respondents specified "angel dust" as some other hallucinogen were used to infer lifetime use of PCP, if question LS01B had not already been answered as "yes." Similarly, "Molly" was added as a slang term for Ecstasy in 2015. Thus, beginning in 2015, respondents who specified "Molly" (with no other information) as some other hallucinogen but who did not report lifetime Ecstasy use were inferred to be Ecstasy users; however, reports of "Molly" that could be associated with another drug (e.g., the stimulant slang terms "white molly" or "black molly") were not used to infer Ecstasy use. Other slang terms that were classified with particular hallucinogens included "shrooms" (coded as psilocybin/ mushrooms), "X," and "X-TC" (the latter two coded as Ecstasy).

In addition, as noted in Section 2.4.1, "OTHER, Specify" data for hallucinogens were not used to edit variables for inhalants, or vice versa. For example, if a respondent specified an inhalant (e.g., gasoline) as some other "hallucinogen" that he or she had ever used, then this report was not used in editing the data for lifetime inhalation of gasoline or lighter fluid from the

inhalants section. However, the "OTHER, Specify" data for hallucinogens in this example would indicate to analysts when respondents specified the use of a drug that fits another category. This approach also assumed that some respondents may specify the use of other drugs according to their *functional* properties. For example, the definition of hallucinogens (i.e., drugs that "often cause people to see or experience things that are not real") could result in respondents thinking of other drugs besides hallucinogens that alter one's perception of reality.

6.2.2 Edits of Recency-of-Use Variables for Tobacco through Methamphetamine

Edits of the variables that establish when respondents last used a substance of interest are generally the most critical.⁷⁸ The recency-of-use variables for tobacco through methamphetamine were the precursors for the final measures that established the prevalence of use in the past 30 days, past 12 months, and lifetime.

The skip logic in the CAI instrument for tobacco through methamphetamine limited the kinds of information that were available for use in editing the recency-of-use variables. In particular, respondents who answered a gate question (or all multiple gate questions) for a given substance as "no" (i.e., never used that substance) were not given the opportunity to answer additional questions as though they were users of that substance. Similarly, respondents who reported that they last used a substance "more than 12 months ago" were not given the opportunity to answer further questions in that section about use in the past 12 months or past 30 days, as though they were more recent users than what they had originally indicated (Figure 6.1).

6.2.2.1 Edits to Recency Variables for Nonusers

As noted in Sections 2.2.3 and 2.4.2 in Chapter 2, respondents were skipped out of all remaining questions about use of a particular substance when they answered "no" to a given gate question for cigarettes through heroin and for methamphetamine or when they answered "no" to all of the individual gate questions for hallucinogens and inhalants.

A code of 91 (i.e., "never used") typically was assigned to the recency variables when it was determined unambiguously that respondents had never used the substance of interest. For crack cocaine, this situation also held if the lifetime variable CRKEVER was coded as 91 because the respondent reported never using cocaine in any form. Similarly, for hallucinogens and inhalants, the recency variables were assigned a code of 91 if all of the lifetime use variables corresponding to the gate questions had been assigned a code of 91 because respondents had never used any of the individual drugs (Section 6.2.1.2).

6.2.2.2 Recency Periods in the Questionnaire

In the sections for alcohol through methamphetamine, the following standard codes for recency applied, based on the available responses to the question(s) about most recent use:

⁷⁸ The recency-of-use variables in this section for tobacco through methamphetamine were based on direct questions that asked respondents to report the last time that they used a particular substance (e.g., "How long has it been since you last used marijuana or hashish?").

- 1 = Within the past 30 days,
- 2 = More than 30 days ago but within the past 12 months, or
- 3 = More than 12 months ago.

In the tobacco sections for cigarettes, smokeless tobacco, and cigars, lifetime users first were asked whether they used the relevant tobacco product in the past 30 days (e.g., question CG05 for cigarettes). If respondents reported that they did not use the particular tobacco product in that period, they were asked to report most recent use prior to the past 30 days (e.g., question CG06 for cigarettes). Recency questions for these tobacco products included categories for most recent use (1) more than 30 days ago but within the past 12 months, (2) more than 12 months ago but within the past 3 years, and (3) more than 3 years ago.⁷⁹

For these tobacco products, single edited recency variables were created that combined the data for the questions about use in the past 30 days and most recent use in later periods, if applicable. Codes of 1 and 2 in the edited recency variables for these tobacco products had the same meaning as for alcohol through sedatives. In addition, standard codes of 3 and 4 in these edited tobacco recency variables had the following meanings based on respondents' answers to the follow-up question (e.g., question CG06 for cigarettes) if they did not report use in the past 30 days:

- 3 = More than 12 months ago but within the past 3 years, and
- 4 = More than 3 years ago.

The CAI instrument included follow-up probes for respondents who were lifetime users of a given substance but did not know or refused to report when they last used it.⁸⁰ Respondents who initially did not know when they last used a substance were asked to give their "best guess" of when they last used it. Respondents who initially refused to report when they last used a substance were asked to reconsider answering the question. If respondents changed their initial answer of "don't know" or "refused" to report a definite period when they last used the substance of interest, then that information served as the starting point for subsequent editing of the substance's recency variable. The standard recency codes described previously in this section were assigned according to the specific period of most recent use that was reported in these follow-up probes. In the absence of any inconsistencies between the recency-of-use answers in the follow-up probes and other data within a given substance's section, these answers from the follow-up probes were accepted as final and were incorporated within the edited recency variable.

6.2.2.3 Logically Assigned Recency Periods among Lifetime Users

In the edited recency variables for tobacco through methamphetamine, respondents were assigned a code of 9 (Used at some point in the lifetime LOGICALLY ASSIGNED) if they

⁷⁹ A recency variable was not created for pipe tobacco because lifetime users of pipe tobacco were asked only if they smoked tobacco in a pipe in the past 30 days.

⁸⁰ For cigarettes, smokeless tobacco, and cigars, these follow-up probes were asked if respondents reported that they did not use a given tobacco product in the past 30 days but they answered "don't know" or "refused" to the follow-up question about most recent use more than 30 days ago.

reported lifetime use of a substance but continued on follow-up to answer "don't know" or "refused" regarding when they last used it.

These respondents were eligible to be statistically imputed to be users in any period, including in the past 30 days, more than 30 days ago but within the past 12 months, or more than 12 months ago. Similarly, for respondents whose recency-of-use questions for LSD, PCP, Ecstasy, ketamine, DMT/AMT/"Foxy," or *Salvia divinorum* had been skipped but they were logically inferred to be lifetime users of these drugs based on their responses to "OTHER, Specify" items (Section 6.2.1.4), their edited recency variables at least initially were assigned a code of 9.

In the recency variables for cigarettes, smokeless tobacco, and cigars, respondents were assigned a code of 19 (Used more than 30 days ago LOGICALLY ASSIGNED) if they reported that they did not use the tobacco product in the past 30 days but they had missing data for the period more than 30 days ago when they last used it. These respondents were eligible to be statistically imputed to have most recently used a given tobacco product more than 30 days ago but within the past 12 months, more than 12 months ago but within the past 3 years, or more than 3 years ago.

The following additional codes could be assigned to the edited recency variables when respondents reported lifetime or more recent use:

- 8 = Used at some point in the past 12 months LOGICALLY ASSIGNED,
- 11 = Used in the past 30 days LOGICALLY ASSIGNED,
- 12 = Used more than 30 days ago but within the past 12 months LOGICALLY ASSIGNED,
- 14 = Used more than 12 months ago LOGICALLY ASSIGNED,⁸¹
- 29 = Used more than 30 days ago but within the past 3 years LOGICALLY ASSIGNED, and
- 39 = Used at some point within the past 3 years LOGICALLY ASSIGNED.

Sections 6.2.2.5 and 6.2.2.6 discuss situations in which these codes were assigned based on the editing procedures. In particular, codes of 11 or 12 could be assigned to a parent recency variable (e.g., any hallucinogen) based on data from a child recency variable (e.g., LSD). Recency variables did not require further statistical imputation when codes of 11 or 12 were assigned.

Recency variables that were assigned codes of 8, 9, 14, 19, 29, or 39 underwent further statistical imputation to assign a specific period of most recent use. These codes were used to place constraints during the subsequent imputation process for assigning the period of most recent use. For example, assigning a code of 8 to an edited recency variable constrained the imputation for the final recency to be within the past 30 days or more than 30 days ago but within the past 12 months.

⁸¹ Codes of 14, 29, and 39 applied only to the recency variables for cigarettes, smokeless tobacco, and cigars.

6.2.2.4 Application of "Flag and Impute" Procedures to Editing of Recency-of-Use Variables for Tobacco through Methamphetamine

The procedures used to edit the CAI recency-of-use variables for tobacco through methamphetamine were referred to as the "flag and impute" procedures (Section 2.4.7). Under these procedures, the limited situations where potential inconsistencies existed between a respondent's answer to a substance's recency question and other data in that section were identified and flagged.

Table 6.2 lists the usual types of inconsistencies that could occur between a substance's recency variable and other variables in that section, and how these inconsistencies were handled through the flag and impute procedures. In the situations described in Table 6.2, these inconsistencies were handled by statistically imputing final values for the affected recency variable and the other variable(s) where the data were inconsistent with the respondent's original answer to the recency question. Most of the codes that were described in Section 6.2.2.3 were assigned to the edited recency variables as a result of these edits. For example, if a respondent reported first use of a marijuana at his or her current age but also reported most recent use more than 12 months ago, both answers logically cannot be true. In this situation, the edited recency variable MJREC was assigned a code of 9 (Used at some point in the lifetime LOGICALLY ASSIGNED) and the inconsistent initiation variables were assigned codes for bad data (Sections 2.3 and 2.4.3).

Table 6.2 How the Flag and Impute Edit Procedures Handled Usual Inconsistencies Involving the CAI Recency Variables

Type of Inconsistency	Edits Implemented
Recency originally indicates use in the past 30 days, but use on 0 days in the past 30 days is confirmed (suggesting less recent use).	The edited recency variable was assigned a code of 8 (i.e., Used at some point in the past 12 months LOGICALLY ASSIGNED) to indicate that the respondent (R) was at least a user in the past 12 months and potentially a user in the past 30 days (Section 6.2.2.3). The 30-day frequency was set to missing (i.e., bad data).
Recency originally indicates use more than 30 days ago but within the past 12 months, but the 12-month frequency indicates use on more than 335 days in that period (suggesting past month use).	The edited recency variable was assigned a code of 8. The 12-month frequency data were set to missing.
Recency does not indicate use in the past 30 days, but the R reports first using the substance (or smoking cigarettes daily) in the same month as the interview took place (suggesting past month use).	If the recency originally indicated use more than 30 days ago but within the past 12 months, it was assigned a code of 8. If the recency originally indicated use more than 12 months ago (or was missing), it was assigned a code of 9 (i.e., Used at some point in the lifetime LOGICALLY ASSIGNED) to indicate that the R was at least a lifetime user (and potentially a user in the past 12 months or past 30 days). The values in the month of first use (MFU) and year of first use (YFU) that triggered the inconsistency with the recency-of-use answer were overwritten with bad data codes. The MFU and YFU variables also were set to bad data if the recency was missing and had been assigned a code of 9 because the initiation data would have suggested use in the past 30 days.

Table 6.2 How the Flag and Impute Edit Procedures Handled Usual Inconsistencies Involving the CAI Recency Variables (continued)

Type of Inconsistency	Edits Implemented
Recency does not indicate use in the past 30 days, but the R has other data suggesting initiation of use in the past 30 days (e.g., if first use was indicated at the R's current age and the R's last birthday was fewer than 30 days ago, or based on a comparison of the 12-month frequency and the maximum number of days that the R could have used the substance).	The edited recency variable was assigned a code of 11 (Used in the past 30 days LOGICALLY ASSIGNED) to infer that the R was a past month user. No further editing was done to the variables indicating use in the past 30 days. This edit was an exception to the general "flag and impute" procedures.
Recency does not indicate use in the past 12 months, but the age at first use (AFU; or age at first daily use of cigarettes) equals the R's current age (suggesting past year use).	For alcohol through methamphetamine, the edited recency variable was assigned a code of 9. For tobacco products, if the lead 30-day question (e.g., CG05 for cigarettes) indicated that the R did not use in the past 30 days, then the recency was assigned a code of 19 (Used more than 30 days ago LOGICALLY ASSIGNED) to indicate that the R was a user more than 30 days ago (but was potentially a user in the past 12 months). Otherwise, the relevant tobacco recency variable was assigned a code of 9. The AFU that triggered the inconsistency with the recency-of-use answer was set to bad data. If the MFU and YFU were answered (i.e., not blank), the values in these variables were overwritten with bad data codes. (Month and year data were considered to be linked with the AFU data. Therefore, if the AFU was questionable, then the month and year were considered to be questionable as well.)
Recency does not indicate use in the past 12 months, but the R reported first using in a month and year that falls within 12 months of the interview date (including data for the month and year when the R reported first smoking cigarettes daily).	For alcohol through methamphetamine, the edited recency variable was assigned a code of 9. For tobacco products, if the lead 30-day question (e.g., CG05 for cigarettes) indicated that the R did not use in the past 30 days, then the recency was assigned a code of 19 (Used more than 30 days ago LOGICALLY ASSIGNED) to indicate that the R was a user more than 30 days ago (but was potentially a user in the past 12 months). Otherwise, the relevant tobacco recency variable was assigned a code of 9. The values in the MFU and YFU that triggered the inconsistency with the recency-of-use answer were overwritten with bad data codes.
For tobacco products, recency indicates use more than 3 years ago, but AFU (or first daily use of cigarettes) indicates that the first use was within 2 years of the R's current age (suggesting use within the past 3 years). (Note that edits checking for indications of use at the R's current age were given priority over the condition described here because the former response pattern suggests use in the past year.)	The recency was assigned a code of 14 (Used more than 12 months ago LOGICALLY ASSIGNED) to indicate that the R last used at some point more than 12 months ago (but potentially in the past 3 years). The AFU that triggered the inconsistency with the recency-of-use answer was set to bad data. If the MFU and YFU were answered (i.e., not blank), the values in these variables were overwritten with bad data codes.

Table 6.2 How the Flag and Impute Edit Procedures Handled Usual Inconsistencies Involving the CAI Recency Variables (continued)

Type of Inconsistency	Edits Implemented
For tobacco products, the R did not use in the past 30 days but did not report a specific period of most recent use more than 30 days ago. However, the R reported first use within 1 or 2 years of his or her current age (i.e., but not at the R's current age).	The recency was assigned a code of 29 (Used more than 30 days ago but within the past 3 years LOGICALLY ASSIGNED) to indicate that the R last used at some point more than 30 days ago but within the past 3 years (but potentially within the past 12 months).
For tobacco products, the R did not indicate whether he or she used in the past 30 days. However, the R reported first use within 1 or 2 years of his or her current age (i.e., but not at the R's current age).	The recency was assigned a code of 39 (Used at some point in the past 3 years LOGICALLY ASSIGNED) to indicate that the R last used at some point within the past 3 years (but potentially within the past 30 days or past 12 months).

CAI = computer-assisted interviewing.

NOTE: Indications of most recent use include answers from follow-up probes for the recency questions.

Prior to implementing these flag and impute rules, initial edits checked for situations where a respondent's original answer to an AFU question might have been inconsistent with his or her recency of use, but a revised AFU was not. For example, if a respondent reported first using a substance at his or her current age, the respondent was asked to verify this AFU. If the respondent reported that this AFU was *not* correct but then on follow-up did not know at what age he or she first used, or refused to answer, the edits updated the AFU to reflect this "don't know" or refusal response. The rationale for this edit was that the respondent indicated that the initial answer was not correct. A final answer of "don't know" or "refused" to an AFU question would not necessarily be inconsistent with a reported recency of use more than 12 months ago.

6.2.2.5 Edits to Parent and Child Recency Variables

Important exceptions to the general flag and impute principles involved situations where inconsistencies existed between parent and child recency variables (any cocaine and crack cocaine; any hallucinogen use and LSD, PCP, Ecstasy, ketamine, DMT/AMT/"Foxy," or *Salvia divinorum* use). These are presented in [Table 6.3](#), along with a description of how the data were edited when specific types of inconsistencies occurred between related recency variables. In these special situations, indications of use of the child drug (e.g., crack cocaine) that were more recent than that indicated for the parent drug category (e.g., cocaine in any form) were used to logically infer more recent use of the parent drug category. For example, not all respondents might make the connection that crack cocaine fits within the broader category of cocaine in general.

Therefore, if a respondent reported last using any cocaine more than 30 days ago and also reported last using crack cocaine in the past 30 days, the edit procedures assigned a code of 11 to the edited cocaine recency variable COCREC (Section 6.2.2.3). This edit indicated logical inference that the respondent had used cocaine in any form in the past 30 days. Overall, however, imputation played a more prominent role than editing in resolving inconsistencies with respect to the most recent use of a drug.

Table 6.3 How the Flag and Impute Edit Procedures Handled Inconsistencies between Parent and Child Recency Variables

Child Recency (i.e., crack, LSD, PCP, Ecstasy, ketamine, DMT/AMT/"Foxy," <i>Salvia divinorum</i>) Reported by Respondent	Parent Recency (i.e., any cocaine, any hallucinogen) Reported by Respondent	Edited Child Recency (i.e., crack, LSD, PCP, Ecstasy, ketamine, DMT/AMT/"Foxy," <i>Salvia divinorum</i>)	Edited Parent Recency (i.e., any cocaine, any hallucinogen)
(1) Indicates use in past month.	Indicates use that is less recent than the past month.	Retains the recency reported by the respondent (R).	Logically infers the R to be a past month user. Assigns a code of 11 (Used in the past 30 days LOGICALLY ASSIGNED).
(2) Indicates use more than 30 days ago but within the past 12 months.	Coded as 8, indicating use at some point in the past 12 months (Section 6.2.2.3).	Retains the recency reported by the R.	Retains the code of 8 to indicate that the R has used at some point in the past 12 months.
	Indicates use more than 12 months ago.	Retains the recency reported by the R.	Logically infers the R to have last used more than 30 days ago but within the past 12 months. Assigns a code of 12 (Used more than 30 days ago but within the past 12 months LOGICALLY ASSIGNED).
	Coded as 9, indicating use at some point in the lifetime (Section 6.2.2.3).	Retains the recency reported by the R.	Logically infers the R to be <i>at least</i> a past year user. Assigns a code of 8.
(3) Indicates use more than 12 months ago.	Coded as 8, indicating use at some point in the past 12 months.	Retains the recency reported by the R.	Retains the code of 8 to indicate that the R is <i>at least</i> a past year user.
	Coded as 9, indicating use at some point in the lifetime.	Retains the recency reported by the R.	Retains the code of 9 to indicate that the R is <i>at least</i> a lifetime user.
(4) Coded as 8, indicating use at some point in the past 12 months (see Table 6.2).	Indicates use more than 30 days ago but within the past 12 months.	Retains the code of 8 to indicate use at some point in the past 12 months.	Assigns a code of 8 to indicate use at some point in the past 12 months.
	Coded as 8, indicating use at some point in the past 12 months.	Retains the code of 8 to indicate use at some point in the past 12 months.	Retains the code of 8 to indicate use at some point in the past 12 months.
	Coded as 9, indicating use at some point in the lifetime.	Retains the code of 8 to indicate use at some point in the past 12 months.	Assigns a code of 8 to indicate use at some point in the past 12 months.

Table 6.3 How the Flag and Impute Edit Procedures Handled Inconsistencies between Parent and Child Recency Variables (continued)

Child Recency (i.e., crack, LSD, PCP, Ecstasy, ketamine, DMT/AMT/"Foxy," <i>Salvia divinorum</i>) Reported by Respondent	Parent Recency (i.e., any cocaine, any hallucinogen) Reported by Respondent	Edited Child Recency (i.e., crack, LSD, PCP, Ecstasy, ketamine, DMT/AMT/"Foxy," <i>Salvia divinorum</i>)	Edited Parent Recency (i.e., any cocaine, any hallucinogen)
(5) Coded as 9, indicating use at some point in the lifetime.	Indicates use more than 30 days ago but within the past 12 months.	Retains the code of 9 to indicate use at some point in the lifetime.	Assigns a code of 8 to indicate use at some point in the past 12 months.
	Indicates use more than 12 months ago.	Retains the code of 9 to indicate use at some point in the lifetime.	Assigns a code of 9 to indicate use at some point in the lifetime.
	Coded as 9, indicating use at some point in the lifetime.	Retains the code of 9 to indicate use at some point in the lifetime.	Retains the code of 9 to indicate use at some point in the lifetime.

NOTE: These edits take place after inconsistencies have been identified between a recency variable and nonrecency variable (e.g., between the recency and the age at first use). For hallucinogens and the child drugs in that section, these edits also take place after the R has revised one or more answers in response to a consistency check. Further, for hallucinogens and the child drugs in that section, any inconsistencies that remain between a given recency variable and other nonrecency variables following inconsistency resolution are transferred back into the recency variables prior to implementation of these edits. For example, if the original answer to the hallucinogen recency disagreed with the age at first use but the revised recency in response to the consistency check did not, then the recency would be updated to reflect the revised value. Prior to implementation of the edits shown in the table, however, if the revised recency still disagreed with the age at first use, then the recency would be edited further to reflect the fact that the previous inconsistency still remained.

In addition, special patterns could remain in the data for cocaine and hallucinogens after most inconsistencies had been addressed and some related data elements were missing. Specifically, respondents could indicate that they first used any cocaine or any hallucinogen within 12 months of the interview date (e.g., first use at their current age) and indicate that they last used that drug more than 30 days ago but within the past 12 months (e.g., HALLUCREC = 2 for any hallucinogen use). Logically, then, if respondents who had used a child drug within a given category (e.g., LSD, PCP, Ecstasy, ketamine, DMT/AMT/"Foxy," or *Salvia divinorum*) had missing data on first use for one or more of their child drugs, it followed not only that they had to have *first* used a child drug at some point in the past 12 months, but they also had to have *last* used the child drug at some point within the past 12 months. In this situation, both the parent recency (e.g., HALLUCREC) and child recency variable(s) (e.g., LSDREC, PCPREC, ECSREC) were set to values of 8 (Used at some point in the past 12 months LOGICALLY ASSIGNED) so that they would be imputed consistently. If the parent recency indicated past month use (e.g., HALLUCREC = 1) when initiation of the parent occurred in the past 12 months but a child recency (e.g., LSDREC) did not indicate use in the past 12 months, only the child recency was set to a value of 8.

Similarly, respondents could indicate that they first used any cocaine, or any hallucinogen, in the same month that they were interviewed and indicate that they last used the drug in the past 30 days but have missing data on first use for the child drug(s) within a category.

In this situation, respondents who had used a child drug within that category (e.g., LSD) also were inferred to have last used that drug in the past 30 days. The edits assigned a code of 11 (Used in the past 30 days LOGICALLY ASSIGNED) to the child recency variables (e.g., LSDREC = 11).

6.2.3 Edits Involving Users of Only Child Drugs for Hallucinogens

In addition to the situations described in Tables 6.2 and 6.3, special edits were applied in situations for hallucinogens in which respondents reported lifetime use of only child drugs. The following specific patterns indicated that respondents were users of only the child drugs:

- one or more gate questions for the child drugs (i.e., LSD, PCP, Ecstasy, ketamine, DMT/AMT/"Foxy," or *Salvia divinorum*) were answered as "yes" and all other gate questions for hallucinogens were answered as "no"; or
- only the gate question for use of any other hallucinogen was answered as "yes" (i.e., including the gate questions for the child drugs) but the *only* report in the "OTHER, Specify" data for that section was for use of a child drug.⁸²

Respondents were *not* considered to be users of only child drugs if they reported lifetime use of only child drugs but they had responses of "don't know" or "refused" for any of the other gate questions. For example, if a respondent answered the question about lifetime use of LSD as "yes" and had some responses of "don't know" or "refused" for other hallucinogen gate questions and no reports of lifetime use of hallucinogens from the other gate questions, the respondent could have been a lifetime user of any of the specific hallucinogens for which he or she had missing data.

Users of only one child hallucinogen who answered the gate question for one of the child drugs as "yes" and answered all remaining gate questions as "no" were asked questions about initiation and recency for the overall parent category. The corresponding questions were skipped for the child drug.⁸³ Corresponding patterns occurred in the data if respondents reported lifetime use only of "other" hallucinogens and specified lifetime use of only a single child drug.

6.2.3.1 Edits to Child Hallucinogen Variables for Users of Only One Child Drug

If respondents reported use of only one child hallucinogen, the values from the parent variables were assigned to the corresponding variables that had been skipped for the child drug. For example, if a respondent reported lifetime use of only LSD, the value from the edited AFU variable for hallucinogens (HALLUCAGE) and the value from the edited hallucinogen recency

⁸² Respondents were considered to be users of only the child drug if they had codes only for "blank" or "bad data" in their "OTHER, Specify" data (see Sections 2.2.1 and 2.3) in addition to the report for a child drug. However, respondents were considered *not* to be users of only the child drug if they had codes for "don't know" or "refused" in their "OTHER, Specify" data in addition to the report for the child drug; responses of "don't know" or "refused" were considered to be potential indications of use of another drug besides the child drug.

⁸³ Because the questionnaire included initiation and recency questions for LSD, PCP, and Ecstasy, the corresponding questions were skipped if respondents reported lifetime use of only one of these child drugs. Because the questionnaire included only recency questions for the child drugs ketamine, DMT/AMT/"Foxy," or *Salvia divinorum*, the corresponding recency questions were skipped if respondents reported lifetime use of only one of these child drugs.

variable (HALLUCREC) were assigned to the corresponding variables for LSD (LSDAGE and LSDREC, respectively). Similarly, any values for the MFU and YFU for recent initiates of any hallucinogen (HALLUCYFU and HALLUCMFU, respectively) were assigned to the corresponding year and month variables for LSD. These values also were retained in the edited variables for use of any hallucinogen. Section 6.2.5 provides further details about the content of the questions in the substance use sections for initiation of use.

6.2.3.2 Edits for Lifetime Users of Multiple Child Drugs (but Only Child Drugs)

A new set of edits was implemented beginning in 2015 for respondents who reported lifetime use of more than one of the six child hallucinogens (but *only* child drugs).⁸⁴ These respondents were asked the questions for any hallucinogen and for the child drugs.

Some of these respondents reported more recent use of any hallucinogen than they reported for the individual child drugs. This pattern would be inconsistent for respondents who were lifetime users of only child drugs. However, this pattern would logically be plausible if the respondents were lifetime users of other hallucinogens in addition to the child drugs. In this situation, therefore, the response of 2 (i.e., no) in the variable for the lifetime use of any other hallucinogen (HALLUCOTH, corresponding to question LS01H) was replaced with a code of 5. This code had the following meaning: 5 = Yes (specific hallucinogen unknown) LOGICALLY ASSIGNED. That is, it was inferred that these respondents logically were lifetime users of some other hallucinogen, but the specific other hallucinogen was unknown. The corresponding "OTHER, Specify" variables for hallucinogens were assigned a code of 9998 (i.e., blank).

Similar edits are described in Section 6.2.5 according to respondents' answers to the initiation questions for any hallucinogen and for the child drugs LSD, PCP, and Ecstasy. Edits based on the initiation data did not apply to ketamine, DMT/AMT/"Foxy," and *Salvia divinorum* because these child drugs did not include initiation questions.

6.2.4 Edits of Frequency-of-Use Variables

Sections of the interview for tobacco through methamphetamine included questions about the number of days that respondents used different drugs in the past 30 days or past 12 months (or the average number of days per week or days per month that they used in the past 12 months). These are referred to in this section as 30-day and 12-month frequency variables, respectively. Data from these frequency questions can be used to distinguish between occasional and more frequent users of a drug. For example, frequent users of alcohol and illicit drugs may represent a group who are potentially in need of substance abuse treatment or other services for their substance use. Similarly, regular users of tobacco products, such as people who smoked cigarettes every day in the past 30 days, probably represent a group that would have greater difficulty stopping their use of tobacco. In addition, the alcohol section included a question about the number of days that respondents consumed four or more drinks (for females) or five or more

⁸⁴ Changes to the hallucinogens section of the interview in 2015 (Section 1.4.1) affected the comparability of hallucinogen estimates between 2015 and prior years, such that 2015 marked the start of a new baseline for estimates of hallucinogen use. Consequently, new edits could be implemented for hallucinogens in 2015 without concerns that doing so would disrupt trends for hallucinogen estimates.

drinks (for males)⁸⁵ per occasion in the past 30 days, a question which is used to construct measures of binge and heavy alcohol use in that period.⁸⁶

6.2.4.1 Preferred Ways of Reporting 12-Month Frequency

For the 12-month frequency determinations, respondents first were asked how they preferred to report their frequency of use in the past 12 months. Respondents could indicate a preference to report their frequency of use in one of three ways: (1) use on an average number of days per week in the past 12 months, (2) use on an average number of days per month in the past 12 months, and (3) the total number of days they used in the past 12 months. In particular, respondents who used a drug regularly in the past 12 months might find it easier to report their frequency of use in one of the first two ways as opposed to figuring the total number of days they used in that entire period. Conversely, respondents who used on only a few days in the past 12 months might prefer the third reporting method.

Individual variables (subsequently referred to as "source variables") were created for the preferred way of reporting the 12-month frequency and for the associated follow-up questions (i.e., average number of days per week, average number of days per month, or total number of days in the past 12 months). An overall measure of the number of days that respondents used a substance in the past 12 months was created from these individual source variables.

For respondents who chose to report a total number of days that they used a substance, the 12-month frequency was the actual number of days that the respondent reported using the drug in the past 12 months (assuming no inconsistency with the 30-day frequency; these inconsistencies are discussed in Section 6.2.4.4). For respondents who chose to report an average use in days per week or days per month, the overall number of days that they used in the past 12 months was a calculated value. Specifically, answers in terms of the average number of days used per week in the past 12 months were multiplied by 52, and answers in terms of the average number of days used per month in the past 12 months were multiplied by 12 in order to yield a calculated 12-month frequency. Because these latter two response options were averaged responses over the past 12 months, no further adjustments were made to the calculated 12-month frequency value when respondents used the drug more than 30 days ago but within the past 12 months, and they did not initiate use at some point in the past 12 months.

6.2.4.2 Editing of 12-Month and 30-Day Frequency Variables for Nonusers or Less Recent Users

If the lifetime gate question(s) and edited recency-of-use variable indicated that the respondent had never used the substance of interest,⁸⁷ then edits at this step assigned a code of 91

⁸⁵ Prior to 2015, the threshold for both males and females was consumption of five or more drinks on a single occasion in the past 30 days. Beginning in 2015, the threshold for females was lowered to consumption of four or more drinks on a single occasion; the threshold for males remained at consumption of five or more drinks.

⁸⁶ Binge alcohol use since 2015 refers to the consumption of four or more drinks for females or five or more drinks for males on a single occasion on at least 1 day in the past 30 days. Heavy alcohol use refers to binge alcohol use on 5 or more days in that period.

⁸⁷ For hallucinogens and inhalants, this meant that the respondent had never used any of the drugs in that category.

to the 30-day frequency variable (where applicable)⁸⁸ and a code of 991 to the final 12-month frequency variable (where applicable). For questions on drugs where respondents were asked to report their frequency of use in the past 12 months, a code of 91 (or 991) was assigned to the source variables pertaining to the preferred method of reporting the 12-month frequency (i.e., average number of days per week, average number of days per month, or total number of days used in the past 12 months), the average number of days per week, the average number of days per month, and total number of days used in the past 12 months.

Similarly, if the edited recency of use indicated that the respondent had used the drug but not in the period of interest, edits at this step assigned a code of 93 to the 30-day frequency variable and a code of 93 (or 993) to the 12-month frequency variable and related source variables that were used to create the 12-month frequency.

6.2.4.3 Editing of 12-Month and 30-Day Frequency Variables When the Recency Was Indefinite

If the respondent was potentially a user in the period of interest (i.e., there was some question about when the respondent last used the drug) and the CAI program had skipped the 30-day or 12-month frequency questions, then the skipped variables retained a blank code. For example, if respondents reported lifetime use of a substance but they did not indicate a specific period for most recent use (Section 6.2.2.3), the frequency-of-use questions retained missing values. If the recency was imputed to indicate use in the past 30 days or past 12 months, the imputation procedures also replaced the missing data in these frequency-of-use variables with nonmissing values (Section 6.5).

Also, for alcohol through methamphetamine, if a respondent reported last using a substance more than 12 months ago, the CAI program skipped the questions pertaining to the frequency of use in the past 12 months and past 30 days. However, reporting first use of the substance at the respondent's current age would be inconsistent with the reported recency. As discussed in Section 6.2.2.4 and [Table 6.2](#), the flag and impute edit rules assigned a value to the recency variable to indicate that this respondent was a user at some point in his or her lifetime, which could include use in the past 30 days or past 12 months. In this situation, the 12-month and 30-day frequency variables that had been skipped retained a blank value in case subsequent imputation might assign the respondent to a more recent category.

In addition, as discussed in Section 2.4.3, if respondents refused the lifetime gate question(s) and were skipped out of the 12-month and 30-day frequency questions (where applicable),⁸⁹ the edits at this step assigned refusal codes to the skipped frequency questions (i.e., the refusal was propagated). However, if respondents were skipped out of the 12-month and 30-day frequency questions because they answered the lifetime gate question(s) as "don't know," the edits retained codes of "blank" in the frequency variables, for the reasons given in Section 2.4.3.

⁸⁸ For alcohol, this edit also applied to other 30-day variables, including the variable on the number of days in the past 30 days that respondents had four or more drinks (for females) or five or more drinks (for males) on a single occasion.

⁸⁹ For the tobacco variables through heroin, such a situation would occur if respondents initially refused the gate question and then refused again on follow-up.

6.2.4.4 Editing of the Frequency-of-Use Variables Based on Consistency Checks

Sections that contained both 12-month and 30-day frequency variables (i.e., alcohol through methamphetamine) included consistency checks between these variables. A consistency check was triggered in situations where the number of days that respondents reported using the substance in the past 30 days exceeded the number of days that the respondent used in the past 12 months.

If respondents did not know how many days they used a substance in the past 30 days or refused to give an answer, they have been asked to give their best estimate of the number of days that they used. Respondents could estimate their 30-day frequency by choosing the category most likely to contain the number of days they used the substance: 1 or 2 days, 3 to 5 days, 6 to 9 days, 10 to 19 days, 20 to 29 days, or all 30 days. A consistency check also was triggered if the number of days that respondents reported using a substance in the past 12 months was lower than the minimum value for the number of days that respondents estimated using that drug in the past 30 days. For example, it would be inconsistent for a respondent to report using marijuana on 6 to 9 days in the past 30 days and also to report using it on fewer than 6 days in the past 12 months.

If the respondent revised either the 12-month or 30-day frequency data (or both) to make them consistent (i.e., such that the 12-month frequency was greater than or equal to the 30-day frequency following any updates done by the respondent), data from the consistency checks were taken as final. This included situations in which respondents resolved inconsistencies between their 12-month frequency and the minimum value for their estimated frequency of use in the past 30 days. If the 30-day frequency (or the minimum value for an estimated 30-day frequency) still was greater than the computed 12-month frequency despite a consistency check having been triggered, then the 12-month frequency was assigned a bad data code.

These edits based on data from consistency checks also applied to data that respondents entered in the follow-up questions for the 12-month and 30-day frequencies for hallucinogens. For example, when respondents were asked follow-up questions about their 30-day frequency of use (Section 6.2.4.6), consistency checks existed between the 12-month hallucinogen frequency and answers to the follow-up questions for the hallucinogen 30-day frequency. Thus, if respondents were routed to the follow-up questions for the hallucinogen 30-day frequency and the resulting 30-day frequency continued to be greater than the 12-month frequency after respondents were prompted to resolve the inconsistency, the hallucinogen 12-month frequency was assigned a bad data code.

In addition, a consistency check was triggered in the alcohol section if respondents reported that they binge drank (i.e., four or more drinks for females or five or more drinks for males in a single occasion) on more days in the past 30 days than they reported for the number of days in which they drank any alcohol in that period. If respondents made their overall 30-day frequency of alcohol use and frequency of consumption of five or more drinks consistent with one another, then these answers were taken as final.

6.2.4.5 Editing of the Frequency-of-Use Variables in Response to Data Patterns Not Involving Recency, Parent/Child Data, or Binge Alcohol Use

Table B.3 in Appendix B lists detailed edits for the 12-month and 30-day frequency variables. The edits that are described pertain to data patterns that do not involve inconsistencies (1) between the frequency data and most recent use, (2) between frequency-of-use data for parent and child data for any cocaine and crack cocaine, or (3) involving the frequency of binge alcohol use in the past 30 days. The edits in Table B.3 also pertain to any inconsistent reports that did not trigger a consistency check during the interview or situations in which a consistency check was triggered but respondents did not resolve the inconsistency.

For example, if the value for the 12-month frequency fell within the range of a respondent's estimate for the 30-day frequency (e.g., if a respondent reported using on 8 days in the past 12 months and on 6 to 9 days in the past 30 days), maximum and minimum values were created for the estimated 30-day frequency. In this example where a respondent reported use on 8 days in the past 12 months but estimated using the drug on "6 to 9" days in the past 30 days, use on 6 to 8 days (as opposed to 6 to 9 days) in the past 30 days would be consistent with the respondent's answer to the 12-month frequency. Information on the maximum and minimum possible number of days that a respondent could have used a drug in the past 30 days was used subsequently during statistical imputation to assign a final value to the 30-day frequency.

6.2.4.6 Editing of the Frequency-of-Use Variables in Response to Parent/Child Data Patterns

Table B.4 in Appendix B lists detailed edits for the 12-month and 30-day frequency variables for the parent/child pair of cocaine and crack cocaine. Because the hallucinogens section did not include questions on the frequency of use in the past 12 months or past 30 days for the child hallucinogens, inconsistencies in parent/child data for the frequency of use did not occur for hallucinogens.

For example, respondents could report use of crack cocaine on a number of days in the past 12 months that was greater than the number of days they reported using any cocaine in the past 12 months. In these situations, the higher value from the 12-month frequency for crack was assigned to the 12-month frequency for any cocaine. The source variables for the 12-month frequency for any cocaine also were edited to indicate the movement of data from the 12-month frequency of crack cocaine use to the 12-month frequency for any cocaine.

Also, since 2003, the hallucinogens section has included follow-up questions for the 12-month frequency variables. These follow-up questions were asked if respondents originally reported that they last used any hallucinogen more than 12 months ago but subsequently reported more recent use of any hallucinogen or any of the child drugs in response to consistency checks between the recency questions for any hallucinogen and the child drugs. Respondents' original answer of use of the parent more than 12 months ago would cause them to be skipped out of the 12-month frequency-of-use questions. Therefore, when respondents gave some updated indication of use in the past 12 months, they were asked to fill in previously missing information about their frequency of use in the past 12 months.

In addition to the 12-month frequency follow-up questions, the hallucinogens section had similar follow-up questions for the 30-day frequency of use. These questions were asked when respondents originally indicated that they last used any hallucinogen more than 30 days ago but subsequently reported that they last used any hallucinogen or any of the child drugs within the past 30 days. Thus, since 2003, data on the frequency of hallucinogen use in the past 30 days have been intended to be supplied by respondents (instead of through statistical imputation) if they subsequently indicated some hallucinogen use in the past 30 days but did not originally report using hallucinogens in the past 30 days.

6.2.4.7 Editing of the Frequency-of-Use and Related Variables Involving Binge Alcohol Use

[Table B.5](#) in Appendix B includes edits related to the data for the frequency of binge alcohol use in the past 30 days. For example, question AL07 in the alcohol section (edited variable ALCUS30D in 2016) asked respondents to report the usual number of drinks that they consumed in a given day in the past 30 days. Although this is not a frequency variable per se, information from this variable was used to edit the 30-day frequency data for alcohol and the frequency data for binge alcohol use. For example, if a male respondent reported having five or more drinks per occasion on *exactly* the same number of days that he reported drinking any alcohol in the past 30 days, then it would logically follow that the respondent's usual number of drinks per day had to have been five or more; a similar principle applied to females who reported having four or more drinks per occasion on exactly the same number of days that they drank any alcohol in the past 30 days. If respondents reported usually having fewer than four drinks (for females) or fewer than five drinks (for males) on the days when they drank alcohol in the past 30 days, ALCUS30D was assigned a special code of 975 (AT LEAST 4 OR 5 Logically assigned) to indicate usual consumption of a number of drinks that was at or above the threshold for binge alcohol use for males or females.

Similarly, if a respondent drank on only 1 day in the past 30 days and reported having fewer than four drinks (for females) or five drinks (for males) on that 1 day in question AL07, but the respondent answered question AL08 as "don't know" or "refused," it would logically follow that the respondent could not have been a binge alcohol user in the past 30 days. When this occurred, the edited variable ALCBNG30D (corresponding to question AL08) was assigned a special code of 80 (NO OCCASIONS OF 4+/5+ DRINKS IN THE PAST 30 DAYS Logically assigned) to denote that the respondent logically could be inferred not to have been a binge drinker in the past 30 days.

In addition, respondents who drank alcohol on more than 1 day in the past 30 days could report usual consumption of four or more drinks (for females) or five or more drinks (for males) but answer AL08 as "don't know" or "refused." These respondents logically would be binge drinkers on at least 1 day in the past 30 days, but the exact number of days on which they engaged in binge drinking would be unknown. Prior to 2015, these missing values could be replaced with values of 0 (i.e., binge drank on 0 days in the past 30 days) as part of the imputation procedures. No editing was done to ALCBNG30D beginning in 2015 to replace the responses of "don't know" or "refused" with a nonmissing value. However, a flag was set to restrict donors for the imputed binge alcohol frequency to respondents who were binge alcohol

users on at least 1 day in the past 30 days. Consequently, since 2015, it is no longer possible for binge alcohol use on 0 days to be imputed for these respondents.

6.2.5 Initiation of Use (Age at First Use, Month of First Use, and Year of First Use)

In sections for tobacco through methamphetamine (except for pipe tobacco), respondents who reported lifetime use were asked how old they were when they first used the substance of interest. If respondents reported first using the drug within 1 year of their current age, they were asked to report the specific month and year when they first used, with the allowable years ranging from 2014 to 2016. If respondents reported first using the drug at their current age and their birth month was earlier than the interview month (i.e., they reached their current age in the same year that they were interviewed), the CAI program assumed that the first use of the drug occurred in the current year (i.e., 2016). These respondents were asked only for the month that they first used in the current year. The remaining respondents who first used a drug within 1 year of their current age could be routed to one of two possible questions on the specific year they first used. They then were routed to a question to report on the specific month that they first used the drug in the year they had reported previously.

Because the routing logic to the different versions of the MFU and YFU questions was mutually exclusive, a single, composite set of MFU and YFU variables was created from the individual unedited variables. In addition, if respondents indicated a specific year that they first used a drug, the final YFU variables for 2016 were recoded to replace unedited codes with values for the specific years (i.e., 2014 through 2016). If respondents confirmed that they first used a drug at their current age and were interviewed subsequent to their birthday in the current year, a code of "2016" was assigned to the YFU; this was done even if respondents did not know what month they first used in the current year, or if they refused to report what month they first used in the current year. If the MFU and YFU questions had been skipped because respondents first used the drug more than 1 year younger than their current ages, legitimate skip codes were assigned to the final MFU and YFU variables.

Since 2002, consistency checks have been included in the instrument if the values for the MFU and YFU were inconsistent with the AFU. Specifically, for recent initiates of a given drug, the CAI program calculated a second AFU use based on the MFU and YFU by comparing these data with the respondent's date of birth. This comparison was not done if the respondent reported first use of the drug in the same month that he or she was born; a unique AFU could not be determined from the MFU and YFU in these situations because it was not known whether the drug use occurred before or after the respondent's birthday. Similarly, a consistency check was not triggered if the respondent had missing data in either of the month or year questions, such as if the respondent knew the year when he or she first used a drug but did not know the MFU.

In remaining situations in which respondents provided complete data for the MFU and YFU, a consistency check was triggered if the MFU and YFU suggested that respondents initiated use of the drug at an earlier or a later age than what they had previously reported. For example, a consistency check was triggered if a 16-year-old respondent reported first using a drug at age 16 but then reported first using the drug in a month and year that would have meant the respondent was 15 years old when he or she first used the drug. No editing needed to be done if respondents indicated twice in a row that the AFU that was calculated from the MFU and YFU

was correct. The CAI program updated the value for the AFU (e.g., AGE1STCG for cigarettes) to agree with the values for the MFU and YFU.

If respondents indicated at some point in the consistency check sequence that the value they had reported for their AFU (e.g., question CG04 for cigarettes) was correct, they had an opportunity to revise the values for their YFU and their MFU. If a consistency check was triggered between the AFU and data in the MFU and YFU, the MFU and YFU were updated with any year and month data that the respondent entered in the consistency checks (e.g., CGCC21 and CG221a for any cigarette use). These data were used in subsequent editing steps. Otherwise, the MFU and YFU data were picked up from the original source variables (e.g., CG04A through CG04D for any cigarette use) for use in editing, such as for the "flag and impute" edits that were described in Section 6.2.2.4 and [Table 6.2](#).

[Table B.6](#) in Appendix B lists edits pertaining to the initiation variables when consistency checks had been triggered because of inconsistent data between the AFU and the YFU and MFU. The default when a respondent did not resolve an inconsistency between the AFU and the MFU and YFU was to favor the AFU in subsequent editing decisions. [Table B.7](#) presents information on additional edits involving the initiation variables other than for parent/child relationships, such as situations where the AFU was inconsistent with the respondent's current age. [Table B.8](#) presents information on edits involving parent/child relationships in the substance use initiation data. In addition to parent/child relationships that have been described previously for cocaine and hallucinogens, parent/child relationships for initiation included first use of any cigarette and initiation of daily cigarette use.

6.2.6 Editing of Lifetime Daily Cigarette Use Status

Lifetime users of cigarettes were asked in question CG15 whether they ever had a period when they smoked cigarettes every day for at least 30 days. Respondents were not asked question CG15 if they had already reported that they smoked cigarettes on all 30 days in the past 30 days.

Edits associated with determining whether respondents had ever smoked cigarettes daily for at least 30 days (edited variable CIGDLYMO) are presented in [Table B.9](#) in Appendix B. In particular, if question CG15 had been skipped because respondents smoked cigarettes on all 30 days in the past 30 days, then CIGDLYMO was assigned a code of 5, where 5 = Yes LOGICALLY ASSIGNED (from skip pattern). A code of 5 was assigned rather than a legitimate skip code (Section 2.4.2) to indicate that it could be logically inferred from the skip logic that the respondent had a lifetime period of having smoked cigarettes daily for at least 30 days.

6.2.7 Edits for Tobacco Variables that Do Not Undergo Imputation

The following tobacco variables were edited but did not undergo further imputation:

- for adolescents aged 12 to 17 who had never smoked a cigarette: their likelihood of smoking a cigarette if their friends offered them one or of smoking a cigarette in the next 12 months;
- lifetime smoking of 100 or more cigarettes;

- usual brands of cigarettes or cigars that respondents used in the past 30 days;
- among respondents who smoked cigarettes in the past 30 days:
 - the average number of cigarettes that they smoked per day;
 - the type of cigarette they smoked (light, ultralight, medium, or full flavor);
 - whether the brand of cigarettes that was smoked most often was a menthol cigarette;
 - for respondents who usually smoked Marlboro cigarettes in the past 30 days, the length of the Marlboro cigarettes that they smoked (shorts, regulars or king-sized, or 100s); and
 - whether respondents smoked a "roll-your-own" cigarette in the past 30 days.

6.2.7.1 Editing of Likelihood Variables for Adolescent Nonsmokers

If respondents were aged 12 to 17 and had never smoked a cigarette, they were routed to questions CG02 and CG03. Question CG02 (edited variable CIGOFRSM) asked adolescents if they would smoke a cigarette if one of their friends offered them a cigarette. Question CG03 (edited variable CIGWILYR) asked adolescents how likely they thought they would be to smoke a cigarette in the next 12 months.

No editing was done to CIGOFRSM and CIGWILYR if adolescents were routed to the corresponding questions because they never smoked cigarettes. Otherwise, a legitimate skip code of 99 (Section 2.4.2) was assigned if the questions had been skipped because (1) respondents were aged 18 or older or (2) they were aged 12 to 17 but they had already smoked a cigarette in their lifetime.

6.2.7.2 Editing of Lifetime Smoking of 100 or More Cigarettes

Question CG16A (edited variable CIG100LF) asked cigarette smokers whether they had ever smoked 100 or more cigarettes in their lifetime. Minimal editing was done to CIG100LF. As indicated in [Table B.9](#), respondents were not asked this question if their answers to previous questions indicated that they had smoked 100 or more cigarettes. For example, respondents who smoked cigarettes in the past 30 days were asked in question CG07 to report the number of days that they smoked cigarettes in the past 30 days. Respondents who smoked cigarettes on more than 1 day in the past 30 days were asked in question CG08 to report the usual number of cigarettes that they smoked on those days when they smoked cigarettes. Question CG08 was a categorical variable that gave ranges of numbers of cigarettes that were smoked per day and, where relevant, the equivalent number of packs of cigarettes (e.g., 16 to 25 cigarettes per day, or about 1 pack).

If the product of the number of days that respondents smoked cigarettes and the lower bound of the range was 100 or greater, then question CG16A was skipped. For example, if a respondent smoked cigarettes on all 30 days in the past 30 days and usually smoked about a pack a day (i.e., 16 to 25 cigarettes per day), then a conservative estimate of the number of cigarettes that he or she smoked in the past 30 days (i.e., based on the lower bound) would be $30 \times 16 = 480$.

In this situation, the respondent would not be asked question CG16A because the number of cigarettes that the respondent logically smoked in the past 30 days was greater than 100. As indicated in Table B.9, if question CG16A had been skipped because had been skipped because respondents the quantity and frequency of cigarette use in the past 30 days indicated that respondents had smoked 100 or more cigarettes, then CIG100LF was assigned a code of 5, where 5 = Yes LOGICALLY ASSIGNED (from skip pattern).

6.2.7.3 Editing of Tobacco Brand Variables

As noted in Section 2.3.1.2, the 2016 CAI instrument included questions to identify the specific brands of cigarettes or cigars that were used most commonly by respondents who reported use of these tobacco products in the past 30 days.⁹⁰ Respondents could choose from a list of brands or they could indicate use of "a brand not on this list." Procedures for coding the "OTHER, Specify" data when respondents reported use of a brand not on the list were described in Section 2.3.1.2.

The coding approach that was described in Section 2.3.1.2 also applied to situations where respondents chose tobacco brands from the lists of brands that they were shown in the questions. For cigarettes, for example, codes of 101 through 126 were used for Basic through Winston, corresponding to the brands that had been listed in question CG11 prior to the 2005 survey.⁹¹ Cigarette brand codes of 127 through 160 were used for Alpine through True, corresponding to the brands listed in question CG11A prior to the 2005 survey.⁹² Thus, the edited variable CIG30BRN for the brand of cigarettes that respondents smoked most often in the past 30 days could be assigned a code of 101 if respondents reported in question CG11 that they usually smoked Basic cigarettes or if they specified Basic cigarettes as "a brand not on this list."

If respondents entered a brand from an available listing, they were asked to confirm their answer. If they confirmed their answer, they were asked no further questions about the brand they used for that particular type of tobacco. However, if respondents indicated that their previous answer was not correct, they were routed back through the series of 30-day brand questions for that type of tobacco. Thus, respondents had the opportunity to make corrections in situations where they may have miskeyed a number, such as if they keyed the number immediately above or below the number of the brand they meant to choose. For each type of tobacco that respondents reported using in the past 30 days, they were allowed to make corrections up to a total of three times. Respondents exited the loop once they confirmed an answer or specified use of a brand not on the list. Respondents also exited the loop if they answered "don't know" or "refused" when asked to confirm their answer.

⁹⁰ Questions for brands of snuff or chewing tobacco were dropped as part of the partial questionnaire redesign for 2015 (Section 2.3.1.2).

⁹¹ Reference is made to response choices in 2004 or earlier because new response choices were added in 2005. Prior to 2005, for example, the first category in question CG11 was Basic. American Spirit was added to question CG11 in 2005, and this became category 1. To preserve continuity with the codes prior to 2005, these codes were not renumbered in 2005 to reflect any changes to the tobacco brand questions.

⁹² The numbering of codes corresponding to responses in question CG11A started with 127 instead of 128 because prior to 2005, the response option of 27 in question CG11 meant "a brand not on this list." This response option was simply a toggle to question CG11A. Because the coding of brands resumed at 127 for brands that were listed in question CG11A, there was no break in the codes.

Because of this routing logic, the brands that respondents confirmed that they used most often in the past 30 days were assigned to the edited variables CIG30BRN (for cigarettes) and CGR30BRN (for cigars). If respondents were rerouted through the series of questions and confirmed their answer on their second or third pass through the questions, the final tobacco brand coding procedures retained the final answer that respondents confirmed and disregarded whatever previous answers the respondent had given but did not confirm. Respondents who answered "don't know" or "refused" when asked to confirm what brand they used were assigned that corresponding code to the final brand variable for that type of tobacco. If respondents did not confirm what brand of a given tobacco type they used most often in the past 30 days despite three passes through the series of questions, a final code of 9000 to the edited variable for that tobacco type's brand. The code of 9000 indicated that these respondents did not confirm their brand despite three opportunities to do so.

As noted in Section 2.3.1.2, respondents sometimes specified that the cigarette brand that they smoked most often was actually a brand of little cigars. No editing was done to the codes for tobacco brands that applied to a different type of tobacco. For example, if respondents specified a little cigar brand as the brand of cigarettes that they smoked most often in the past 30 days, the edited variable CIG30BRN retained a code in the 400 or 4000 series for cigars. Continuing this example, if respondents did not report use of cigars in the past 30 days, no editing was done to the data for cigars to indicate use of cigars in the past 30 days in the edited recency variable CIGARREC or to assign the cigar brand code from cigarettes to CGR30BRN.

In addition, the precoded response options in the questions for cigarette and cigar brands since 2005 have included more prevalent brands that were mentioned in "OTHER, Specify" data from prior years. Since 2005, response options also are no longer present for less prevalent brands from 1999 to 2004. Documentation of these changes since 2005 is provided in the report on general principles and procedures for editing drug use data in the 2011 NSDUH methodological resource book (Krouil, Handley, & Bradshaw, 2013).

These changes in 2005 did not affect the creation of the edited tobacco brand variables. Nevertheless, analysts are advised that this change could affect analyses comparing trends for certain brands prior to 2005 and from 2005 onward. For example, significant differences in prevalence between some tobacco brand estimates prior to 2005 and after 2005 could occur if respondents prior to 2005 needed to type in the brand as "a brand not on this list" but they could choose it from a list in 2005 and beyond.

6.2.7.4 Editing of Miscellaneous Cigarette Use Variables for the Past 30 Days

As noted previously, respondents who reported that they smoked cigarettes in the 30 days prior to the interview were asked a series of additional questions about their cigarette use in that period (i.e., other than the brand of cigarettes that they smoked most often). These questions covered the following topics:

- the average number of cigarettes that they smoked per day (Section 6.2.7.4.1);
- the type of cigarette they smoked (light, ultralight, medium, or full flavor) and whether the brand of cigarettes that respondents smoked most often was a menthol cigarette (Section 6.2.7.4.2);

- for respondents who usually smoked Marlboro cigarettes in the past 30 days, the length of the Marlboro cigarettes that they smoked (shorts, regulars or king-sized, or 100s) (Section 6.2.7.4.3); and
- whether respondents smoked a "roll-your-own" cigarette in the past 30 days (Section 6.2.7.4.4).

6.2.7.4.1 Average Number of Cigarettes Smoked Per Day

Respondents who smoked cigarettes in the past 30 days were asked one of two possible questions regarding the number of cigarettes they smoked per day. Respondents who smoked on only 1 day in the past 30 days were asked to report the number of cigarettes they smoked on that 1 day. Respondents who smoked on more than 1 day (or who estimated the number of days they smoked in the past 30 days) were asked to report the average number of cigarettes they smoked per day. A single, composite variable (CIG30AV) was created from these two questions using data from whatever question the respondents were asked. No further editing was done to the data from these two questions.

6.2.7.4.2 Type of Cigarettes Including Menthol Cigarettes

Past month cigarette users who reported that they smoked one of the specific brands in questions CG11, CG11A, RCG11, RCG11A, RRCG11, or RRCG11A (i.e., not counting "a brand not on this list") were routed to one of two possible questions about the type of cigarette they usually smoked in the past 30 days (CGTAR1 or CGTAR2). Since 2005, these questions have included a new response option for respondents who smoked "mediums" most often in the past 30 days. Addition of this new level for mediums in 2005 affected the distribution of responses in CGTAR1 and CGTAR2 relative to the distribution in 2004, such that the data were not comparable between 2004 and 2005. Therefore, the name of the edited variable corresponding to CGTAR1 and CGTAR2 has been CIG30TPE since 2005; prior to 2005, this variable was called CIG30TYP.

Past month cigarette users also were routed to one of two possible questions regarding whether the brand of cigarette they usually smoked was menthol (CGMENTH1 or CGMENTH2). Routing to CGTAR1 or CGTAR2 and to CGMENTH1 or CGMENTH2 was mutually exclusive for respondents who smoked cigarettes in the past 30 days (i.e., respondents were routed to one or the other question in a set but not both). Therefore, composite variables were created for the cigarette type (CIG30TPE) and whether the cigarette brand that respondents smoked most often was menthol (CIG30MEN).

No attempt was made to edit CIG30TPE or CIG30MEN for consistency with the cigarette brand from the variable CIG30BRN. In developing these items, instrument development staff consulted with tobacco research experts regarding which brands offered or did not offer menthol, light, or ultralight varieties. No conclusive information was obtained. As was discussed in Section 2.3.1.2, the "OTHER, Specify" data for tobacco brands did not capture information for particulars such as regular or menthol forms, or light, ultralight, or full-flavor varieties. Therefore, in situations where respondents specified that level of detail regarding the brand of cigarette that they smoked most often in the past 30 days, that information was not used to edit CIG30TPE or CIG30MEN.

6.2.7.4.3 *Type of Marlboro Cigarettes*

Respondents who reported in questions CG11, RCG11, or RRCG11 that they usually smoked Marlboro cigarettes in the past 30 days (and who confirmed this report of smoking Marlboro cigarettes) have been asked in question CGLNTH since 2005 about the length of the Marlboro cigarettes they smoked most often in the past 30 days: (1) shorts, (2) regulars or king-sized, or (3) 100s. The edited variable corresponding to CGLNTH was CIG30MLN. CIG30MLN was assigned a legitimate skip code of 99 (Section 2.4.2) if the cigarette brand variable CIG30BRN did not have a missing value and did not indicate that respondents smoked Marlboro cigarettes most often in the past 30 days; this included situations in which respondents reported that they smoked "a brand not on this list" and did not specify that they smoked Marlboro cigarettes.⁹³ Because respondents who reported that they smoked some other brand of cigarettes in the past 30 days were not asked CGLNTH, CIG30MLN retained values of 98 (blank) for any respondents who specified that the "other" brand of cigarettes was Marlboro; however, CIG30MLN was blank for this reason for fewer than 10 of approximately 5,400 respondents in 2016 who reported that they smoked Marlboro cigarettes most often in the past 30 days. In addition, CIG30MLN had missing values if respondents who smoked cigarettes in the past 30 days did not know or refused to report at the outset what cigarette brand they smoked most often in the past 30 days, or if they failed to confirm the brand that they smoked most often.

6.2.7.4.4 *"Roll-Your-Own" Cigarettes*

The cigarette section also included a question (CG14) about whether respondents smoked part or all of a "roll-your-own" cigarette in the past 30 days. The edited variable CIG30ROL corresponded to this question. The cigarette brand question CG11A⁹⁴ included response categories for two roll-your-own brands of cigarette tobacco. Respondents who chose either of these roll-your-own brands were skipped out of question CG14; by choosing a roll-your-own brand from the list of cigarette brands, these respondents already had indicated that they had smoked a roll-your-own cigarette in the past 30 days. Therefore, if question CG14 had been skipped and the cigarette brand was one of the roll-your-own brands from CG11a, a code of 5 (Yes LOGICALLY ASSIGNED [from skip pattern]) was assigned to the edited variable CIG30ROL.

However, respondents could specify a cigarette brand that was not on the list in questions CG11 and CG11A and then specify a roll-your-own brand in question CG12. In this situation, respondents were routed to question CG14. If respondents specified a roll-your-own brand and CG14 was already answered as "yes," no further editing needed to be done. If CG14 was not answered as "yes" when a respondent had specified a roll-your-own cigarette brand, the respondent's original answer was replaced with a code of 3 (Yes LOGICALLY ASSIGNED). This

⁹³ This assignment of legitimate skip codes included situations in which the "other" cigarette brand was for a different type of tobacco (e.g., if respondents specified that the brand of "cigarettes" they smoked most often was a cigarillo or little cigar) or if respondents reported smoking a brand of cigarettes "not on this list" and they had missing values in their "OTHER, Specify" data. In this latter situation, CIG30BRN continued to have a code of 1999 (Cigarette; brand otherwise unspecified), and it was inferred in the editing of CIG30MLN that this otherwise unspecified cigarette brand was not Marlboro.

⁹⁴ Respondents could be routed to questions RCG11, RRCG11, RCG11A, RRCG11A, RCG12, or RRCG12 if they cycled through the cigarette brand questions more than once. For brevity, however, reference here is limited to the first set of cigarette brand questions: CG11, CG11A, or CG12.

code of 3 signified to analysts that the respondent's original answer in question CG14 had been overwritten to make the roll-your-own data consistent with the cigarette brand information that was recorded in the variable CIG30BRN. However, this edit was implemented in 2016 for fewer than 10 respondents out of approximately 12,700 who reported cigarette use in the past 30 days.

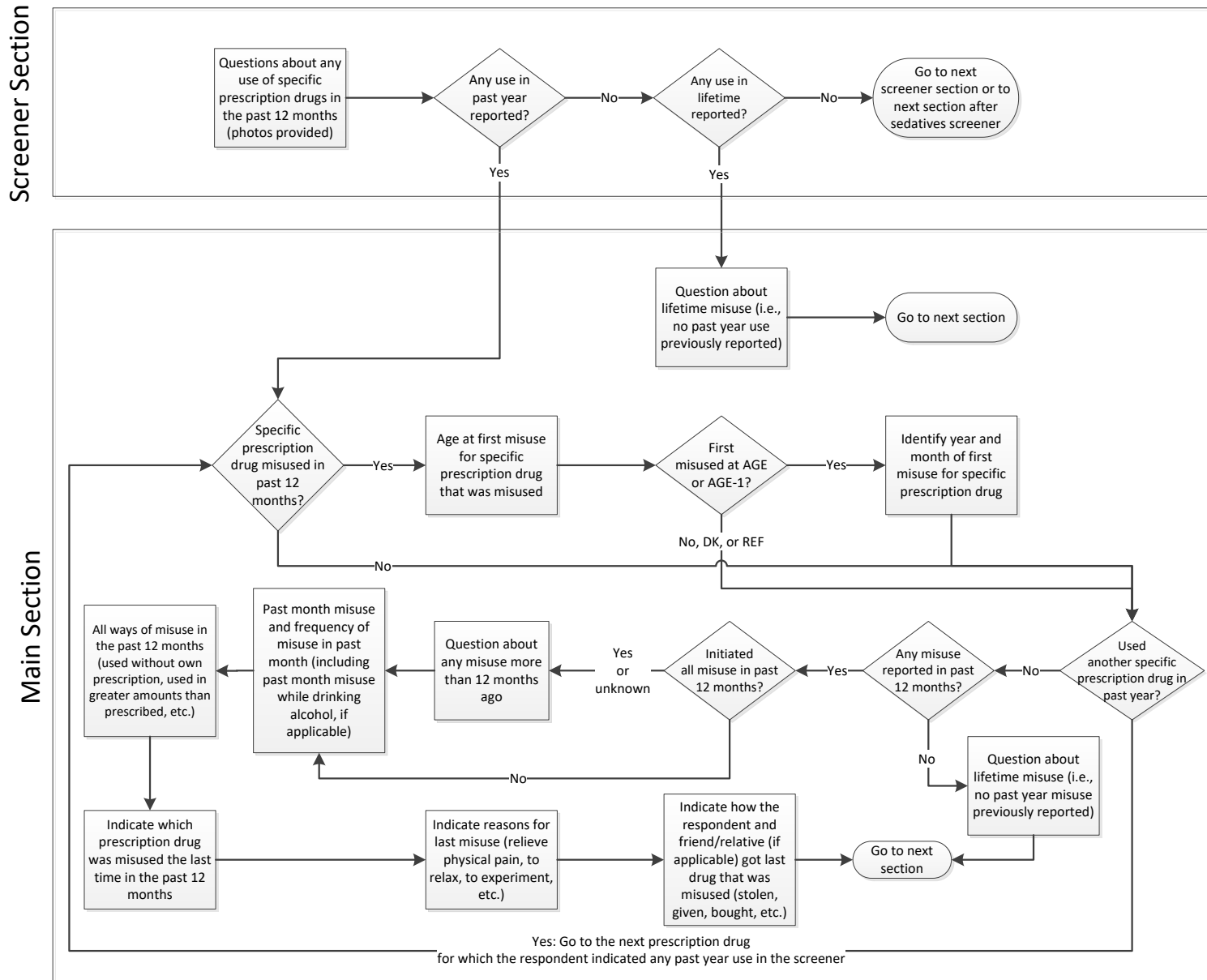
6.3 Editing Procedures for the Prescription Drug Variables

Figure 6.2 provides an overview of the general routing logic for the prescription psychotherapeutic drug sections based on the redesigned questions beginning in 2015. For each category of psychotherapeutic drugs (i.e., pain relievers, tranquilizers, stimulants, and sedatives), respondents first were asked "screener" questions about any use of specific prescription drugs in the past 12 months or lifetime use of any prescription drug in that category if they did not report use in the past 12 months. Respondents who reported any use of prescription drugs in the past 12 months or lifetime periods were asked more detailed questions about their misuse of prescription drugs in that category.⁹⁵ The questions for misuse included the following: misuse of specific prescription drugs in the past 12 months, if respondents reported any use of a given specific prescription drug in that period; initiation of misuse of specific prescription drugs that were misused in the past 12 months (i.e., age at first misuse [AFM], and for recent initiates, the month of first misuse [MFM] and year of first misuse [YFM]); lifetime misuse of any prescription drug in a given category (if misuse in the past 12 months was not reported, or if respondents reported any use of prescription drugs in that category in their lifetime but not in the past 12 months); misuse of any prescription drugs in a given category more than 12 months prior to the interview, if respondents reported initiation of misuse in the past 12 months for all prescription drugs that they misused in that period (i.e., to identify respondents who initiated misuse of any prescription drug in the category more than 12 months prior to the interview); and frequency of misuse in the past 30 days.

In addition, respondents who reported misuse of any prescription psychotherapeutic drug in a given category in the past 30 days and who also drank alcohol in the past 30 days were asked whether they misused prescription drugs at the same time or within a couple of hours of when they drank alcohol in that period. Respondents who reported misuse of any prescription psychotherapeutic drug in a given category in the past 12 months were asked to report (1) the ways in which they misused prescription drugs within that category in the past 12 months (e.g., without a prescription of their own, in greater amounts than directed), (2) the last prescription drug in that category that they misused in the past 12 months (if they reported misuse of more than one prescription drug), (3) their reasons for misusing that last prescription drug they misused in the past 12 months, and (4) how they (or their friends or relatives, if applicable) obtained the last prescription drug they misused in that period. In addition, respondents who reported the misuse of prescription stimulants in the past 12 months were asked whether they used a needle to inject stimulants in the past 12 months or (if applicable) the past 30 days.

⁹⁵ The definition for misuse of prescription psychotherapeutic drugs was changed in 2015. In 2014, misuse (also known as nonmedical use) was defined as use of prescription drugs that were not prescribed for respondents or that respondents took only for the experience or feeling that the drugs caused. Beginning in 2015, misuse was defined as use of prescription drugs in any way that a doctor did not direct respondents to use them, including (1) use without a prescription of the respondent's own; (2) use in greater amounts, more often, or longer than told to take a drug; or (3) use in any other way not directed by a doctor.

Figure 6.2 Routing Logic for Respondents in the Prescription Psychotherapeutic Drug Sections



DK = don't know; REF = refused.

6.3.1 Mentions of Over-the-Counter Drugs

To make the questions about the use of specific prescription drugs exhaustive for a given prescription psychotherapeutic drug category, respondents were asked whether they used "any other" prescription drug in that category in the past 12 months. Respondents who reported use in the past 12 months of other prescription drugs in a given category were asked whether they misused these other drugs in that period, and if so, they were asked to specify the names of up to five prescription drugs in that category that they misused in the past 12 months (Section 2.3.1.1).

One way that individuals can misuse prescription drugs is by taking them without having their own prescription. Because over-the-counter (OTC) drugs by definition are available without a prescription, respondents were instructed not to include OTC drugs when answering the questions in screener sections for pain relievers, stimulants, and sedatives; these drug categories include products that are available in the United States without a prescription (e.g., the pain reliever aspirin, the stimulant Dexatrim®, and the sedative Nytol®). In the "OTHER, Specify" data for other prescription drugs that respondents misused in the past 12 months, however, respondents could report the misuse of OTC drugs.

A code of 4 (No LOGICALLY ASSIGNED) was assigned to the lead question for the misuse of any other drug in that category in the past 12 months (e.g., PNR0TNMYR, corresponding to question PRYOTH for pain relievers) when the only drugs⁹⁶ that respondents specified for misuse in the past 12 months were OTC drugs, and (1) they reported the past year misuse of at least one other prescription drug in that section or (2) they answered questions for the misuse of other specific prescription drugs in that category in the past 12 months as "don't know" or "refused." That is, it could be inferred that the respondents had not misused other prescription drugs in that category in the past 12 months, even if they had misused some specific prescription drugs in that period (or had potentially misused prescription drugs for which they answered the questions for misuse in the past 12 months as "don't know" or "refused"). Codes in the "OTHER, Specify" variables were overwritten with a code of 9989 (LEGITIMATE SKIP Logically assigned; see Section 2.4.2) to indicate that respondents logically should have skipped these questions.

If it could be determined unambiguously that "other" prescription drugs in a given psychotherapeutic category (e.g., pain relievers) were the only prescription drugs that respondents misused in the past 12 months and the misuse of only OTC drugs was reported, then respondents were inferred not to have misused any prescription drug in that category in the past 12 months. Respondents were misusers of only the "other" prescription drugs in a category in the past 12 months if they reported misuse of "any other" prescription drug in the category and

- they reported use in the past 12 months of specific prescription drugs, and they answered all of the follow-up questions for misuse of these specific drugs as "no," or

⁹⁶ Respondents were considered to have specified the misuse of only OTC drugs if the "OTHER, Specify" data contained reports of only OTC drugs or the data included only OTC drugs and codes for "bad data." This condition was not considered to have been satisfied in situations where the "OTHER, Specify" data included responses of "don't know" or "refused" in addition to reports of OTC drugs (i.e., but no reports of the misuse of prescription drugs). These responses of "don't know" or "refused" were interpreted to mean that the respondents had potentially misused other drugs that require a prescription.

- the only prescription drugs in a given category that they used in the past 12 months for any reason were other prescription drugs, such that the lead question (e.g., PRYOTH for pain relievers) asked whether respondents had misused *any* prescription drug in that category in the past 12 months.

Editing of other variables within a prescription drug section is discussed in further detail in subsequent sections for situations in which respondents were determined not to have misused any prescription drug in a given category because of their reports of misuse of only OTC drugs.

However, no editing was done to the corresponding screener question about any use of other drugs in that category in the past 12 months (e.g., PRANYOTH for pain relievers), which resulted in respondents being routed to the question about misuse of any other drug in the category. Even if the only drugs in a given category for which respondents reported misuse in the past 12 months were OTC drugs, respondents could have used other *prescription* drugs in the past 12 months that they did not misuse.

6.3.2 Any Past Year or Lifetime Use of Specific Prescription Drugs

Questions for the use of specific prescription drugs in the past 12 months were "enter all that apply" questions, in which respondents could report the use of more than one drug in the list. Respondents were instructed to key "95" if they had never used any of the specific drugs in a given list in the past 12 months (e.g., the pain relievers Vicodin[®], Lortab[®], Norco[®], Zohydro[®] ER, and generic hydrocodone in question PR01). Respondents could choose one or more drugs from a list or they could key "95," but they could not do both. Data for the use or nonuse of specific prescription drugs were captured in separate variables (e.g., the variables VICOANYR, LTABANYR, NRCOANYR, ZOHYANYR, and HYCDANYR, corresponding to the specific prescription pain relievers in PR01).

Questions for the use of any other prescription drug in a given category in the past 12 months (e.g., question PRANYOTH for pain relievers) were "yes/no" questions, where 1 = Yes, and 2 = No. Respondents could choose only one response in these questions for the use of other prescription drugs in that category (i.e., besides the specific prescription drugs that respondents were asked about).

6.3.2.1 Editing of Screener Variables for Past Year Users

This section describes editing procedures for the prescription drug screener variables if respondents reported the past year use of at least one prescription drug in that section (e.g., pain relievers), including the use of any other drug in the category. Minimal editing was done to the variables for the use of other prescription drugs in a given category when respondents reported the use of any prescription drug in a given screener section, other than to assign mnemonic names to the edited variables (e.g., edited variable PNROTANYR, corresponding to question PRANYOTH). Consequently, the edited variables for the past year use of other prescription drugs in a category retained values of 1 (Yes), 2 (No), or codes for missing data (i.e., 94 for "don't know," 97 for "refused," or 85 for "bad data").

Consistent with the procedures described in Section 2.4.4, if respondents reported using one or more specific prescription drugs in the past 12 months from a given list of "enter all that apply" questions, the variables corresponding to the specific prescription drugs that respondents used were coded as 1 (Yes). Variables corresponding to other drugs in the list that respondents did not report using in the past 12 months were coded as 2, where 2 = No (not entered).

As described in Section 2.4.5, if respondents did not report using a specific prescription drug in the past 12 months but they specified misusing it in that period in the "OTHER, Specify" data for the relevant prescription drug category (e.g., pain relievers), then the corresponding screener variable for that drug was assigned a code of 5 (Yes LOGICALLY ASSIGNED). For example, if a respondent did not report using the pain reliever Norco® in the past 12 months but reported misusing that drug in the same period, then the respondent logically used Norco® for any reason in the past 12 months. The edited variable NRCOANYR for any use of Norco® in that period was assigned a code of 5. The following additional edits were implemented to the other screener variables from a given list when a code of 5 was assigned through these editing procedures:

- If respondents reported the use of other prescription drugs from that list in the past 12 months, then the remaining prescription drug variables from that list were coded as 1 or 2, as per the procedures described above. For example, if NRCOANYR was assigned a code of 5 but the respondent had previously reported using Vicodin® in the past 12 months, then VICOANYR was assigned a code of 1, and the remaining variables LTABANYR (for Lortab®), ZOHYANYR (for Zohydro® ER), and HYCDANYR (for generic hydrocodone) were assigned codes of 2.
- If respondents keyed 95 to indicate that they had not used any of the prescription drugs in that list in the past 12 months, then the edited variables for the remaining prescription drugs in that list were assigned a code of 2 (No [Not entered]). For example, if NRCOANYR was assigned a code of 5 but the respondent had previously reported not using any of the pain relievers in question PR01, then the remaining variables VICOANYR, LTABANYR, ZOHYANYR, and HYCDANYR were assigned codes of 2.

The additional edits described below were implemented if respondents did not report the use of any specific prescription drugs in the past 12 months from a given list in an "enter all that apply" question (e.g., PR01) but they reported use in the past 12 months elsewhere in that screener section. These edits applied to situations in which no screener variables corresponding to prescription drugs in a given list were assigned a code of 5 based on "OTHER, Specify" data for the misuse of prescription drugs in that category.

- If respondents keyed 95 to indicate that they had not used any prescription drug in a given list in the past 12 months, then all of the edited variables corresponding to the prescription drugs in that list were assigned a code of 95, where 95 = I have not used any of these [DRUGS (e.g., pain relievers)] in the past 12 months. For example, if respondents keyed 95 for question PR01 to indicate that they had not used any of these pain relievers in the past 12 months but they reported in question PR02 that they used OxyContin® in that period, then the edited variables VICOANYR,

LTABANYR, NRCOANYR, ZOHYANYR, and HYCDANYR were assigned a code of 95.

- If respondents reported that they did not know or they refused to report whether they used any prescription drug in a given list in the past 12 months, then all of the edited variables corresponding to the prescription drugs in that list were assigned codes of 94 (DON'T KNOW) or 97 (REFUSED), respectively. For example, if respondents did not know whether they used any of the pain relievers in question PR01 in the past 12 months, then the edited variables VICOANYR, LTABANYR, NRCOANYR, ZOHYANYR, and HYCDANYR were assigned a code of 94.

When respondents reported the use of at least one prescription drug in a given category in the past 12 months, the question for the lifetime use of any prescription drug in that category was skipped (e.g., question PRLANY for pain relievers) because users in the past 12 months logically had used prescription drugs in that category in their lifetime. The corresponding edited variable for lifetime use (e.g., PNRANYLIF, corresponding to PRLANY) was assigned a code of 5, where 5 = Yes LOGICALLY ASSIGNED (from skip pattern).

6.3.2.2 Editing of Screener Variables for Lifetime but Not Past Year Users

Respondents who did not report using any prescription drug in a given category in the past 12 months (including respondents who did not know or refused to report whether they used certain prescription drugs in that period) were asked whether they ever used any prescription drug in that category for any reason (e.g., question PRLANY for pain relievers). When respondents answered the question for any use in their lifetime as "yes," then the edited variable (e.g., PNRANYLIF, corresponding to PRLANY) was assigned a code of 1 (Yes). The edits described below were implemented when respondents reported lifetime use of prescription drugs in a given category for any reason (e.g., PNRANYLIF = 1).

- If respondents keyed 95 for all "enter all that apply" questions in the screener section to indicate that they had not used any of these prescription drugs in the past 12 months, and they also reported that they did not use any other drug in that category in the past 12 months, then all edited screener variables pertaining to the past 12 months were assigned a code of 93, where 93 = DID NOT USE [PRESCRIPTION DRUG CATEGORY] IN THE PAST 12 MONTHS. For example, if PNRANYLIF = 1, responses of 95 were keyed in questions PR01 through PR10, and PRANYOTH was answered as 2 (no), then the edited variables VICOANYR through PNOTANYR (corresponding to PRLANY) were assigned a code of 93 (DID NOT USE PAIN RELIEVERS IN THE PAST 12 MONTHS).
- If respondents had a combination of responses of 95 or missing data (i.e., codes of 85, 94, or 97 for "bad data," "don't know," or "refused," respectively), then the edited variables corresponding to prescription drugs that respondents indicated that they explicitly did not use in the past 12 months were assigned a code of 95 (for nonuse of any of these prescription drugs in the past 12 months). The other variables with missing data retained the codes for missing values because the respondents could have used some of these prescription drugs in the past 12 months. For example, if respondents keyed 95 in question PR01 but they answered question PR02 as "don't

know," then the edited variables VICOANYR, LTABANYR, NRCOANYR, ZOHYANYR, and HYCDANYR from PR01 were assigned a code of 95. The edited variables OXCNANYR (for OxyContin[®]), PRCTANYR (for Percocet[®]), and PRDNANYR (for Percodan[®]) were assigned a code of 94 for "don't know."

6.3.2.3 Editing of Screener Variables for Lifetime Nonusers

When respondents did not report use in the past 12 months of any prescription drug within a given category, and they answered the question for any use in their lifetime as "no," then the edited variable (e.g., PNRANYLIF, corresponding to PRLANY) was assigned a code of 2 (No). The edits described below were implemented when respondents reported that they never used any prescription drug in a given category for any reason (e.g., PNRANYLIF = 2).

- If respondents keyed 95 for all "enter all that apply" questions in the screener section to indicate that they had not used any of these prescription drugs in the past 12 months, and they also reported that they did not use any other drug in that category in the past 12 months, then all edited screener variables pertaining to the past 12 months were assigned a code of 91, where 91 = NEVER USED [PRESCRIPTION DRUG CATEGORY]. For example, if PNRANYLIF = 2, responses of 95 were keyed in questions PR01 through PR10, and if PRANYOTH was answered as 2 (no), then the edited variables VICOANYR through PNOTANYR (corresponding to PRLANY) were assigned a code of 91 (NEVER USED PAIN RELIEVERS).
- If respondents had a combination of responses of 95 or missing data (i.e., codes of 85, 94, or 97 for "bad data," "don't know," or "refused," respectively), then the edited variables corresponding to prescription drugs that respondents indicated that they explicitly did not use in the past 12 months were assigned a code of 91. If variables had missing data for use in the past 12 months, the missing values were replaced with codes of 81 (see the example below for pain relievers). Thus, the editing procedures gave precedence to reports of lifetime nonuse of any prescription drug in a category over missing data for use in the past 12 months. For example, if respondents keyed 95 in question PR01 but they answered question PR02 as "don't know," then the edited variables VICOANYR, LTABANYR, NRCOANYR, ZOHYANYR, and HYCDANYR from PR01 were assigned a code of 91. The edited variables OXCNANYR (for OxyContin[®]), PRCTANYR (for Percocet[®]), and PRDNANYR (for Percodan[®]) were assigned a code of 81 (NEVER USED PAIN RELIEVERS Logically assigned).

6.3.3 Most Recent Use of Prescription Drugs for Any Reason

Unlike the sections of the interview for tobacco through methamphetamine, respondents were not explicitly asked to report the last time when they used prescription drugs for any reason. Instead, recoded variables were created for the most recent use of prescription drugs in a given category for any reason. Creation of the recoded variable PNRANYREC for the most recent use of prescription pain relievers for any reason is described as an example.

- If use in the past 12 months was reported for at least one prescription pain reliever in the screener section, then PNRANYREC for the most recent use was set to 1, where 1 = Within the past 12 months.
- If no use of specific prescription pain relievers in the past 12 months was reported in the screener section, but lifetime use of any prescription pain reliever was reported:
 - PNRANYREC was set to 2 (More than 12 months ago) if all edited screener variables for the use of pain relievers in the past 12 months had been set to 93 (DID NOT USE PAIN RELIEVERS IN THE PAST 12 MONTHS). When respondents explicitly reported that they had not used any pain relievers in the past 12 months, the most recent use of any pain reliever logically occurred more than 12 months prior to the interview.
 - PNRANYREC was set to 9 (Used at some point in the lifetime) if any of the edited screener variables had missing values (e.g., "don't know"). Respondents in this situation could have used pain relievers in the past 12 months but did not report doing so.
- If question PRLANY for any lifetime use of pain relievers was answered as "no," then PNRANYREC was set to 91 (i.e., never used).
- If PRLANY was answered as "don't know" or "refused," responses of 95 were keyed in questions PR01 through PR10, and PRANYOTH was answered as 2 (no), then PNRANYREC was set to 83 (DID NOT USE IN THE PAST 12 MONTHS [LIFETIME UNKNOWN] Logically assigned). Respondents in this situation definitely did not use pain relievers in the past 12 months, but it was unknown whether they used them more than 12 months ago or they never used pain relievers.
- Otherwise, PNRANYREC was assigned a code of 98 (blank) to indicate that any lifetime use or lifetime nonuse was unknown. In particular, PNRANYREC was coded as 98 when the question for any lifetime use was answered as "don't know" or "refused," or when the corresponding pain reliever screener variables for all questions that respondents answered had been assigned a code for bad data because of patterned responses (Section 2.3.2.1).

6.3.4 Misuse of Specific Prescription Drugs in the Past 12 Months

As shown in [Figure 6.2](#), respondents who reported any use of specific prescription drugs in the past 12 months were asked whether they misused these prescription drugs in that period. Misuse since 2015 has been defined as the use of prescription drugs in any way a doctor did not direct respondents to use them. Examples of behaviors constituting misuse that were presented to respondents included (1) use without a prescription of the respondent's own; (2) use in greater amounts, more often, or longer than told to take a drug; or (3) use in any other way a doctor did not tell respondents to take a drug.

6.3.4.1 Editing of Specific Prescription Drug Misuse Variables for Prescription Drug Misusers in the Past 12 Months

Questions for the misuse of specific prescription drugs in the past 12 months (e.g., question PRY01 for the misuse of Vicodin[®]) were "yes/no" questions, where 1 = Yes, and 2 = No. In the absence of any patterned responses for a given prescription drug section (Section 2.3.2), edited variables for misuse of specific prescription drugs would be coded as 1 ("yes"), 2 ("no"), 94 ("don't know"), or 97 ("refused") for the prescription drugs that respondents used for any reason in the past 12 months. Questions for the misuse of prescription drugs that respondents did not report using in the past 12 months were coded as 93 (or 993, etc.) to indicate that the respondents did not use those particular drugs in the past 12 months. For example, if a respondent reported the past year use of Vicodin[®] in question PR01, did not report using any of the other drugs listed in that question, and there were no reports of the misuse of these other drugs in the past 12 months in the "OTHER, Specify" data, then the edited variables LTABNMYR (for the misuse of Lortab[®]), NRCONMYR (for the misuse of Norco[®]), ZOHYNMYR (for the misuse of Zohydro[®] ER), and HYCDNMYR (for the misuse of generic hydrocodone) were assigned codes of 93.

As described in Section 2.4.5, if respondents did not report that they misused a specific prescription drug in the past 12 months but they specified misusing it in that period in the "OTHER, Specify" data for the relevant prescription drug category (e.g., pain relievers), then the corresponding variable for the misuse of that drug in the past 12 months was edited as follows:

- If the respondent reported any use of that drug in the past 12 months in the corresponding screener section but did not answer the question for misuse as "yes," then the edited misuse variable was assigned a code of 3 (Yes LOGICALLY ASSIGNED). For example, suppose a respondent reported using the pain reliever Norco[®] in the past 12 months. For this respondent, the screener variable NRCOANYR for Norco[®] would be coded as 1 (Yes). If the respondent did not report misusing the drug but specified it as some other pain reliever that he or she misused in the past 12 months, then NRCONMYR (corresponding to question PRY03) would be edited to a value of 3.
- If the screener variable (e.g., NRCOANYR for Norco[®]) had been assigned a code of 5 because the respondent had not reported using the drug in the past 12 months but specified its misuse in that period in the "OTHER, Specify" data (Section 6.3.2.1), then the edited misuse variable for that drug (e.g., NRCONMYR) also was assigned a code of 5 (Yes LOGICALLY ASSIGNED).

Respondents who reported the misuse of at least one prescription drug in a given category in the past 12 months were not asked whether they ever misused any prescription drug in that category (e.g., question PRL02) because these respondents logically had misused prescription drugs in that category in their lifetime. The corresponding edited variable for lifetime misuse (e.g., PNRNMLIF for pain relievers) was assigned a code of 5, where 5 = Yes LOGICALLY ASSIGNED (from skip pattern).

6.3.4.2 Editing of Specific Prescription Drug Misuse Variables for Misusers of Over-the-Counter Drugs in the Past 12 Months

As noted in Section 6.3.1, respondents were logically inferred not to have misused other prescription drugs in a given category when the only drugs that they specified for misuse in the past 12 months were OTC drugs. No further editing was done to the variables for the past year misuse of specific prescription drugs if the lead question for the misuse of any other drug in that category in the past 12 months (e.g., PNROTNMYR, corresponding to question PRYOTH for pain relievers) had been assigned a code of 4 (No LOGICALLY ASSIGNED) but the respondent reported the misuse in the past 12 months of other prescription drugs in that category.

Respondents were inferred not to have misused any prescription drug in the past 12 months in the following situations where the respondents reported the misuse of only OTC drugs:

- All other questions for the misuse of specific prescription drugs in the past 12 months had been answered as "no" or had been legitimately skipped because the respondent did not report any use of that drug in the same period.
- The only screener question for which the respondent reported any use in the past 12 months was for "any other" drug in the category, such that the corresponding question for misuse (e.g., PRYOTH for pain relievers) asked about the misuse of *any* prescription drug in that category (i.e., as opposed to the misuse of any *other* prescription drug in that category).

Assignment of the relevant code to the corresponding recoded variable for the most recent misuse when respondents misused only OTC drugs is described in Section 6.3.5.

No additional editing was done if respondents had been inferred not to have misused other prescription drugs in a given category in the past 12 months because the only drugs they misused in that period were OTC drugs and the respondents answered questions for the misuse of other specific prescription drugs in that period as "don't know" or "refused." In this situation, respondents had reported any use of specific prescription drugs in the past 12 months in addition to reporting any use of "any other" prescription drug in that category. Consequently, these respondents could have misused the specific prescription drugs for which they answered the misuse questions as "don't know" or "refused."

6.3.4.3 Editing of Specific Prescription Drug Misuse Variables for Lifetime but Not Past Year Misusers

Respondents who did not report any past year use of prescription drugs in the relevant screener section but who reported lifetime use for any reason of any prescription drug in that category were asked whether they ever misused any prescription drug in that category (e.g., question PRL01 for pain relievers). Respondents who were asked about misuse of specific prescription drugs in the past 12 months but who did not report misuse in that period also were asked whether they ever misused any prescription drug in that category (e.g., question PRL02 for pain relievers). Thus, these lifetime misuse questions were mutually exclusive. For pain

relievers, for example, respondents could be routed either to PRL01 or PRL02 to indicate whether or not they had ever misused pain relievers in their lifetime, but not to both questions.

If respondents reported lifetime misuse (e.g., PNRNMLIF = 1 for pain relievers), then the following edits were implemented for the variables for the misuse of specific prescription drugs:

- If all questions that respondents were asked for their misuse of specific prescription drugs in the past 12 months were answered as "no," then all variables for the misuse of specific prescription drugs in the past 12 months were assigned a code of 93 (or 993, etc.) to indicate that the respondent misused no prescription drugs within that psychotherapeutic category in the past 12 months.
- If respondents did not report the misuse of prescription drugs within a psychotherapeutic category in the past 12 months but they had missing data for the misuse of other prescription drugs within that category, then the following edits were implemented:
 - Variables corresponding to prescription drug questions that were skipped because respondents did not report any use in the corresponding screener section were assigned a code of 93 (or 993, etc.).
 - If respondents were routed to a question for a specific prescription drug, and that question was answered as "no," then the corresponding edited variable (e.g., NRCONMYR for the misuse of Norco[®] in the past 12 months) retained a value of 2 (i.e., No).
 - Variables for the misuse of specific prescription drugs in the past 12 months that had missing data retained missing values.

6.3.4.4 Editing of Specific Prescription Drug Variables for No Lifetime Misuse

Respondents who answered a question for lifetime misuse as "no" (e.g., question PRL01 or PRL02 for pain relievers) were determined never to have misused any prescription drug in that category in their lifetime. The edits described below were implemented when respondents were routed to the section for the misuse of prescription drugs in a given category (e.g., pain relievers) and they reported that they never misused any prescription drug in that category (e.g., PNRNMLIF = 2).

- If the variables for the misuse in the past 12 months of all specific prescription drugs in that section were coded as 2 (i.e., no) or 93, then all edited variables for the misuse of these specific prescription drugs were coded as 91 (or 991, etc.) to indicate no lifetime misuse.
- If respondents had a combination of responses of 2, 93, or missing data, then the edited variables corresponding to prescription drugs that respondents explicitly did not misuse were assigned a code of 91 (or 991, etc.). If variables had missing data for misuse in the past 12 months, then any answers to these questions that were coded with missing values (e.g., "don't know") were replaced with codes of 81 (or 981, etc.). Thus, the editing procedures gave precedence to reports of no lifetime misuse of any prescription drug in a category over missing data for use in the past 12 months.

6.3.5 Most Recent Misuse of Prescription Drugs

Similar to the variables for the most recent use of prescription drugs in a given category for any reason, recoded variables were created for the most recent misuse. Creation of the recoded variable PNRNMREC for the most recent misuse of prescription pain relievers is described as an example.

- If misuse of any prescription pain reliever in the past 30 days was reported in question PRM01, or if PRM01 had been skipped because respondents reported that they initiated the misuse of one or more pain relievers in the 30 days prior to the interview (Section 6.3.6), then PNRNMREC was set to 1, where 1 = In the past 30 days.
- If respondents misused one or more specific pain relievers in the past 12 months:
 - PNRNMREC was set to 2 (More than 30 days ago but within the past 12 months) if PRM01 was answered as "no." In this situation, it was unambiguous that the respondents who reported misuse in the past 12 months had not misused any prescription pain relievers in the past 30 days.
 - PNRNMREC was set to 8 (Used at some point in the past 12 months LOGICALLY ASSIGNED) if PRM01 was answered as "don't know" or "refused." These respondents definitely misused prescription pain relievers in the past 12 months and also potentially misused pain relievers in the past 30 days.
- If no misuse of specific prescription pain relievers in the past 12 months was reported, but lifetime misuse of any prescription pain reliever was reported:
 - PNRNMREC was set to 3 (Misused more than 12 months ago) if all questions for the misuse of specific pain relievers in the past 12 months were answered as "no" or were coded as 93 because they had been legitimately skipped.
 - PNRNMREC was set to 9 (Misused at some point in the lifetime) if any of the questions for misuse of specific pain relievers in the past 12 months were answered as "don't know" or "refused." Respondents in this situation could have misused pain relievers in the past 12 months but did not report doing so.
- If the only prescription drugs in a given category that respondents misused in the past 12 months were OTC drugs, then PNRNMREC was set to 83 (DID NOT MISUSE IN THE PAST 12 MONTHS (LIFETIME UNKNOWN) Logically assigned). Because the respondents during the interview reported misuse in the past 12 months, the lifetime misuse question (e.g., PRL02 for pain relievers) was skipped.
- If questions PRL01 or PRL02 for any lifetime misuse of pain relievers were answered as "no," then PNRNMREC was set to 91 (i.e., never used). The answer of "no" in the lifetime question overruled any answers of "don't know" or "refused" for the misuse of specific prescription drugs in the past 12 months.
- If the screener data indicated that respondents had never used pain relievers in their lifetime for any reason (PNRANYREC = 91), then PNRNMREC also was set to 91.

- Otherwise, PNRNMREC was assigned a code of 98 (blank) to indicate that lifetime misuse or no lifetime misuse was unknown. In particular, PNRNMREC was coded as 98 when either question for lifetime misuse was answered as "don't know" or "refused" or when all pain reliever screener and misuse data had been assigned a code for bad data because of patterned responses (Section 2.3.2.1).

6.3.6 Initiation of Misuse of Prescription Drugs

The approach for identifying individuals who initiated the misuse of prescription drugs in the past 12 months changed beginning in 2015. The questions for the initiation of misuse since 2015 have focused only on the specific prescription drugs that respondents misused within the past year; respondents were not asked about initiation for prescription drugs that they may have misused in their lifetime but not in the past 12 months. Consequently, since 2015, respondents cannot report initiation in the past 12 months without also reporting misuse in that period. In contrast, prior to 2015, respondents could answer the questions for the first misuse of prescription drugs in a given category (e.g., pain relievers) as though they initiated misuse in the past 12 months (e.g., first misuse at the respondent's current age) and then report that they last misused any prescription drug in that category "more than 12 months ago." When the initiation questions focused on the specific prescription drugs that respondents misused in the past 12 months, cognitive testing of the redesigned questions also suggested that respondents were able to think about specific life events surrounding their first misuse of a prescription drug as a strategy for recalling their first misuse.

6.3.6.1 Identification of the Initiation of Misuse in the Past 12 Months

To establish that a respondent in 2016 had initiated the misuse of *all* prescription drugs in a given category in the past 12 months, any initiation prior to that period must be ruled out. Even if respondents reported initiation in the past 12 months for all of the specific prescription drugs in a given overall category that they misused in that period, this information does not conclusively rule out the misuse of other prescription drugs that may have occurred more than 12 months prior to the interview. For example, suppose a respondent reported the misuse of one pain reliever (e.g., OxyContin[®]) in the past 12 months and that he or she first misused this pain reliever in the same period. This information established that the respondent initiated the misuse of OxyContin[®] in the past 12 months. However, this information is not sufficient for establishing that the respondent initiated the misuse of *all* pain relievers in the past 12 months. For example, the respondent may have misused other pain relievers (e.g., Vicodin[®]) more than 12 months ago but was not asked about the misuse of these other pain relievers.

Therefore, respondents who reported that they initiated misuse in the past 12 months for all of the specific prescription drugs in a given category that they misused in that period were asked a follow-up question (e.g., question PRL03 for pain relievers) to establish whether they had *ever* misused prescription drugs in that category more than 12 months before being interviewed. Respondents also were asked this follow-up question if there were missing data for initiation in the past 12 months (e.g., if the AFM for a specific prescription drug was answered as "don't know") without any other reports of initiation more than 12 months prior to the interview. Respondents who reported that they initiated the misuse of one or more prescription drugs in a category more than 12 months before being interviewed (e.g., if a 20-year-old respondent

reported first misusing OxyContin® at age 18) were skipped out of this follow-up question because their initiation data for misuse indicated that they initiated the misuse of prescription drugs in that category more than 12 months prior to the interview date.

If respondents were asked this follow-up question, and they answered it as "no," meaning that they had never misused any prescription drug in that category more than 12 months prior to the interview date, then it could be concluded that the respondents initiated the misuse of all prescription drugs in that category in the past 12 months. However, if respondents answered the follow-up as "yes," then they would *not* be considered recent initiates for the misuse of any prescription drug in that category (e.g., pain relievers) because this response meant that the respondents initiated the misuse of some other drug in the category more than 12 months ago.

The edited variables for indicating whether respondents initiated the misuse of all prescription drugs in a given category in the past 12 months were PNRNMINIT for pain relievers (corresponding to question PRL03), TRQNMINIT for tranquilizers (corresponding to TRL03), STMNMINIT for stimulants (corresponding to STL03), and SEDNMINIT for sedatives (corresponding to question SVL03). Creation of the edited variable PNRNMINIT for the initiation of misuse of prescription pain relievers in the past 12 months is described as an example.

The following edits were implemented for PNRNMINIT if respondents misused prescription pain relievers in the past 12 months:

- If respondents answered question PRL03 as "yes" or "no," then PNRNMINIT was reverse coded.
 - If PRL03 was answered as "no" (PRL03 = 2), then PNRNMINIT was set to a value of 1, where 1 = Past year initiate.
 - If PRL03 was answered as "yes" (PRL03 = 1), then PNRNMINIT was set to a value of 2, where 2 = Not a past year initiate.
- If respondents answered question PRL03 as "don't know" or "refused," then PNRNMINIT retained the code for the answer from PRL03 (e.g., 94 = DON'T KNOW).
- If question PRL03 was skipped because respondents reported that they first misused any pain reliever more than 12 months prior to the interview, the PNRNMINIT was coded as 4, where 4 = Not a past year initiate LOGICALLY ASSIGNED (from skip pattern).

Additional edits when PRL03 had been skipped because respondents did not report the misuse of prescription pain relievers in the past 12 months took into account data from the recency of misuse variable PNRNMREC (Section 6.3.5).

- If respondents last misused any pain reliever more than 12 months ago (PNRNMREC = 3), then PNRNMINIT was set to a value of 93, where 93 = DID NOT MISUSE PAIN RELIEVERS IN THE PAST 12 MONTHS.

- If respondents did not misuse pain relievers in the past 12 months but it was unknown whether the respondents misused pain relievers more than 12 months ago or that they never misused pain relievers (PNRNMREC = 83; Sections 6.3.4.2 and 6.3.5), then PNRNMINIT was set to a value of 83, where 83 = NOT A PAST YEAR INITIATE (AT-RISK STATUS UNKNOWN).⁹⁷
- If respondents misused pain relievers in their lifetime but their most recent period of misuse was unknown (PNRNMREC = 9), then PNRNMINIT was coded as 98 (i.e., blank). PNRNMINIT retained a missing value for these respondents because they were potential misusers in the past 12 months (and potential initiates in that period).
- If PNRNMREC had been coded as 98, then PNRNMINIT also was coded as 98.

6.3.6.2 Age, Year, and Month of Initiation for Individuals Who Initiated Misuse in the Past 12 Months

For each specific prescription drug that respondents misused in the past 12 months, respondents were asked to report their age when they first misused that drug (AFM). If respondents reported that they first misused a drug (e.g., Vicodin[®]) within 1 year of their current age, they were asked to report the year and month when they first misused that drug (YFM and MFM, respectively). Recoded variables for the age, year, and month of first misuse also were created for the overall prescription drug categories (e.g., pain relievers) for respondents who initiated the misuse of all prescription drugs in that category in the past 12 months.

6.3.6.2.1 Age-, Year-, and Month-of-Initiation Variables for Specific Prescription Drugs

Unlike the logic in the other substance use sections, a "hard error" rather than a consistency check was triggered if a respondent reported an AFM for a specific prescription drug that was greater than his or her current age. A hard error does not allow the respondent to move to the next question until the respondent enters an AFM that is consistent with his or her current age or the respondent changes the AFM answer to "don't know" or "refused." In contrast, the consistency checks in the other substance use sections give the respondent a single opportunity to enter an AFU that is consistent with his or her current age but do not require the respondent to resolve the inconsistency. Because of the use of a hard error when the AFM was inconsistent with the respondent's current age, the AFM variables for specific prescription drugs would have values less than or equal to the current age or values for "don't know" or "refused," unless a different final age value had been assigned (see Section 4.2.2). In these situations, if the AFM was greater than the final age, the edited AFM was coded as bad data.

Similar to the logic in other substance use sections, the CAI program compared the AFM for a given prescription drug with the age that was calculated from the YFM, MFM, and date of

⁹⁷ Respondents who never misused prescription drugs in a given category in their lifetime are considered to be "at risk" for initiating misuse in the past 12 months. However, respondents who initiated the misuse of prescription drugs in a given category more than 12 months prior to being interviewed are no longer considered to be at risk for initiating misuse in the past 12 months because they already misused prescription drugs in that category. Consequently, if it is unknown whether respondents misused any prescription drug in a category in their lifetime, then their status for being at risk for initiation also is unknown.

birth (subsequently referred to as the MYR1ST age, such as MYR1STPR1, for Vicodin®). A consistency check was triggered if the calculated MYR1ST age for that prescription drug differed from the corresponding AFM. Therefore, the edits described in Table B.6 in Appendix B also applied to the AFM, YFM, and MFM variables for specific prescription drugs.

If the AFM for a specific prescription drug was 2 or more years younger than the respondent's current age, then the edited YFM and MFM were assigned legitimate skip codes (Section 2.4.2). Specifically, if the YFM and MFM had been skipped, then they were coded as 9999 and 99, respectively. If the respondent initially reported first misuse within 1 year of his or her current age but changed the AFM in response to a consistency check to being 2 or more years younger than the his or her current age (Table B.6), then any values in the YFM and MFM were overwritten with codes of 9989 and 89, respectively.

For any specific drug, it also was possible for the AFM to be at a younger age than would be expected from the YFM and MFM because the respondent's birthday was on the 1st day of the month and the MFM was in the same month as the birth month. For example, if a respondent was born on June 1, 2001, was interviewed in July 2016, and reported first misuse at age 14 and in June 2016, the respondent would have been 15 on June 1. In this example, the respondent would be expected to have first misused the prescription drug at age 15 rather than at age 14. Therefore, when this pattern occurred, the AFM was set to the age that was consistent with the YFM and MFM (e.g., age 15 for the scenario that was described).

Additional edits to the AFM, YFM, and MFM variables (e.g., VICONMAGE, VICONMYFU, and VICONMMFU for Vicodin®) reflected the values that had been assigned through editing to the variable corresponding to the lead question for misuse (e.g., VICONMYR).

- If the lead question for misuse had a value of 2 (No), 4 (No LOGICALLY ASSIGNED), or 93 (Did not use), then blank values in the AFM, YFM, and MFM were replaced with codes of 993, 9993, and 93, respectively. Nonblank values were overwritten with corresponding codes of 983, 9983, and 83.
- If the lead question for misuse had a value of 91 because respondents had not misused any prescription drugs in that category in their lifetime, then blank values in the AFM, YFM, and MFM were replaced with codes of 991, 9991, and 91, respectively. Nonblank values were overwritten with corresponding codes of 981, 9981, and 81.
- If the lead question for misuse had a value of 81 because of the edits that were described in Section 6.3.4.4, the AFM, YFM, and MFM were globally assigned codes of 981, 9981, and 81, respectively.

6.3.6.2.2 Age-, Year-, and Month-of-Initiation Variables for Initiates in the Past 12 Months

Recoded AFM, YFM, and MFM variables also were created for the misuse of any prescription drug in a given psychotherapeutic category (e.g., pain relievers) for respondents who reported initiating the misuse of all prescription drugs in that category in the past 12 months. If initiates in the past 12 months reported the misuse of only one prescription drug in that category,

then creation of the recoded AFM, YFM, and MFM variables was straightforward; the AFM, YFM, and MFM from the specific prescription drug were transferred to the corresponding recoded variables for initiation of any prescription drug in that category. The general principles listed below were followed for creating these recoded AFM, YFM, and MFM variables when recent initiates reported the misuse of more than one prescription drug in the past 12 months and they had no missing initiation data for the specific prescription drugs.

- The minimum AFM was chosen as the AFM for the overall category. If this AFM was not reported for any other prescription drug in that category, then the YFM and MFM from the specific prescription drug corresponding to that AFM were assigned to the YFM and MFM variables for the first misuse of any prescription drug in that category.
- If the minimum AFM was reported for more than one prescription drug, then the minimum YFM was chosen as the YFM for the overall category. If this YFM was not reported for any other prescription drug in that category, then the MFM from the specific prescription drug corresponding to that YFM was assigned to the corresponding MFM variable for the first misuse of any prescription drug in that category.
- If the minimum AFM and YFM were reported for more than one prescription drug, then the minimum MFM was chosen as the MFM for the overall category.

Focusing on initiates of misuse in the past 12 months facilitated the creation of recoded variables in the following ways:

- By definition, first misuse for initiates in the past 12 months would have occurred at the respondent's current age or at the age that is 1 year earlier than the respondent's current age. Consequently, if recent initiates reported first misusing one prescription drug at the age that was 1 year earlier than their current age but they had missing AFM data for other prescription drugs, it could still be inferred that the AFM for any prescription drug in that category was 1 year earlier than the current age.
- First misuse for initiates in the past 12 months would have occurred in the current year or the year immediately prior to the current year (e.g., 2015 for the 2016 survey). Consequently, if recent initiates in 2016 reported first misusing one prescription drug in 2015 but they had missing YFM data for other prescription drugs, it could still be inferred that the YFM for any prescription drug in that category was in 2015.

Additional edits for the AFM, YFM, and MFM when respondents did not report initiation in the past 12 months took into account data from the corresponding initiation variable (e.g., PNRNMINIT). Edits for PNRNMAGE, PNRNMYFU, and PNRNMMFU when PNRNMINIT did not indicate initiation of the misuse of prescription pain relievers in the past 12 months are described as examples.

- If PNRNMINT had a value of 2 or 4 because respondents misused in the past 12 months but initiated misuse more than 12 months ago, then PNRNMAGE, PNRNMYFU, and PNRNMMFU were assigned codes of 993, 9993, and 93, respectively.

- If PNRNMINT was coded as 93 because respondents did not misuse pain relievers in the past 12 months, then PNRNMAGE, PNRNMYFU, and PNRNMMFU were assigned codes of 993, 9993, and 93, respectively.
- If PNRNMINT was coded as 83 because respondents did not misuse pain relievers in the past 12 months but their lifetime misuse status was unknown, then PNRNMAGE, PNRNMYFU, and PNRNMMFU were assigned codes of 983, 9983, and 83, respectively.
- If PNRNMINT was coded as 91 because respondents did not misuse pain relievers in their lifetime, then PNRNMAGE, PNRNMYFU, and PNRNMMFU were assigned codes of 991, 9991, and 91, respectively.
- Otherwise, if PNRNMINIT had a missing value, then PNRNMAGE, PNRNMYFU, and PNRNMMFU were assigned codes for blank (i.e., 998, 9998, and 98, respectively). For example, if PNRNMINIT = 94 (i.e., don't know) then the AFM, YFM, and MFM variables retained missing values.

6.3.7 Misuse of Prescription Drugs in the Past 30 Days

A key component of editing the variables for misuse of prescription drugs in the past 30 days involved comparing these variables with the corresponding recency variable for misuse (e.g., PNRNMREC for pain relievers). If the question pertaining to misuse of any prescription drug in the category in the past 30 days (e.g., PRM01 for pain relievers; edited variable PNRNM30D) had been answered as "no" (e.g., PNRNM30D = 2), then the code of 2 for "no" was retained in the edited variable to indicate that the respondent did not misuse any prescription drug in that category in the past 30 days.

- If the recency had a value of 2 or 3 to indicate that the most recent misuse definitely occurred more than 30 days prior to the interview, then the skipped variables pertaining to misuse in the past 30 days were assigned a code of 93 (DID NOT MISUSE IN THE PAST 30 DAYS).
- If the recency had been assigned a code of 83 because the only misuse in the past 12 months involved the misuse of OTC drugs (Sections 6.3.4.2 and 6.3.5), then any responses in the 30-day questions were overwritten with a code of 83 in the edited variables (DID NOT MISUSE IN THE PAST 30 DAYS Logically assigned). If questions had been skipped (e.g., if the lead question for any misuse had been answered as "no"), then the edited variables were assigned a code of 93.
- If the recency had a value of 91 to indicate that respondents had never misused (or used) any prescription drug in that category in their lifetime, then the skipped variables for misuse in the past 30 days were assigned a code of 91 (NEVER USED/MISUSED).
- If the recency had a value of 8 or 9, then skipped variables for misuse in the past 30 days typically retained a code of 98 (blank). An exception for the misuse of prescription drugs with alcohol in the past 30 days is discussed in Section 6.3.7.2.
- Otherwise, if the recency had a value of 98 (i.e., blank), then the variables for misuse in the past 30 days were assigned a missing value code to indicate that lifetime

misuse or no lifetime misuse was unknown. If the variables for misuse in the past 30 days had a code of 98 (blank), then these variables retained that code of 98. Otherwise, values were overwritten with a code of 85 (BAD DATA Logically assigned).

If respondents reported that they initiated the misuse of any specific prescription drugs in the 30 days prior to the interview,⁹⁸ then they were skipped out of the question that asked whether they misused any prescription drug in that category in the past 30 days (e.g., PRM01 for pain relievers). Consequently, it was not possible for most respondents in 2016 to report that they initiated the misuse of prescription drugs in a given category in the past 30 days and then to answer the question for misuse in that period as "no."⁹⁹ When the question for misuse of any prescription drug in a category had been skipped because respondents reported that they initiated the misuse of specific prescription drugs in the past 30 days, then the edited variable (e.g., PNRNM30D) was assigned a code of 5, where 5 = Yes LOGICALLY ASSIGNED (from skip pattern).

When respondents reported that they initiated misuse of specific prescription drugs in the past 30 days, they were eligible to be asked the follow-up questions pertaining to misuse in the past 30 days (i.e., frequency of misuse, misuse with alcohol). Consequently, data would exist in the additional edited variables for misuse in the past 30 days when the lead question for any misuse in the past 30 days had been assigned a code of 5; an exception for misuse with alcohol in the past 30 days is discussed in Section 6.3.7.2.

Remaining edits that are described in this section pertain to respondents who misused any prescription drug in a given category within the past 30 days.

6.3.7.1 Frequency of Misuse of Prescription Drugs in the Past 30 Days

If respondents misused prescription drugs in a given category in the past 30 days, they were asked to report the number of days in that period on which they misused any prescription drug (e.g., question PRM02 for pain relievers). Thus, the edited variables for the frequency of misuse in the past 30 days (e.g., PNRNM30FQ, corresponding to question PRM02) were continuous variables, with valid values ranging from 1 to 30.

If respondents did not know or refused to report the exact number of days on which they misused any prescription drug in a category in the past 30 days, they were asked a follow-up question (e.g., PRM02DKRE for pain relievers) in which they were asked to give their best

⁹⁸ Respondents were identified as having initiated the misuse of a prescription drug in the past 30 days if they (1) reported that they first misused that drug in the same month that they were interviewed; or (2) first misused at their current age (i.e., if not in the same month as the interview month), and their birthday was within 30 days of the interview date.

⁹⁹ An exception was identified in 2016 for respondents who had their last birthday in December 2015, who initiated misuse of pain relievers at their current age and in December 2015, and who were interviewed in January 2016 within 30 days of their birthday. The CAI logic in 2016 did not take this pattern into account and allowed some respondents to answer question PRM01 as "no." Because this pattern was identified late in the 2016 data processing and occurred very rarely in 2016, no editing was done in 2016 for respondents with this pattern. In 2017, however, these respondents have been logically inferred to have misused in the past 30 days. The CAI logic also has been corrected for 2018 to identify these respondents as initiates in the past 30 days.

estimate of the number of days they misused prescription drugs in that category in the past 30 days. Similar to the edited variables for the estimated 30-day frequency in other sections of the interview (Section 6.2.4.4), these edited variables (e.g., PNRNM30ES, corresponding to question PRM02DKRE) were categorical, based on the following response options: (1) 1 or 2 days, (2) 3 to 5 days, (3) 6 to 9 days, (4) 10 to 19 days, (5) 20 to 29 days, or (6) all 30 days. If a value between 1 and 30 days was reported in the 30-day frequency question (e.g., PRM02), then the edited variable for the estimated 30-day frequency was assigned a code of 99 (LEGITIMATE SKIP).

Unlike 30-day frequency questions for alcohol through methamphetamine, respondents could not report that they misused prescription drugs on "0 days" in the past 30 days, because respondents had been asked a "yes/no" question for whether they misused prescription drugs in that category in the past 30 days, or else they reported that they had initiated misuse of specific prescription drugs in the past 30 days. Especially for respondents who were directly asked whether they misused any prescription drugs in a category in the past 30 days (e.g., PRM01 for pain relievers), an explicit answer of "yes" would indicate that respondents had misused prescription drugs in that category on at least 1 day in that period. Because the 30-day frequency questions for prescription drugs did not allow respondents to report misuse on 0 days in that period, the edit that was described in Table 6.2 for alcohol through methamphetamine did not apply to the editing of the 30-day frequency variables for prescription drugs.

However, if respondents reported for all specific prescription drugs in a category that they misused in the past 12 months that they initiated misuse within 30 days of the interview date (regardless of whether or not they initiated the misuse of all prescription drugs in that category in the past 12 months), then it was possible for them to report a frequency of misuse that was greater than the number of days that they could have misused based on their initiation data. For example, suppose a respondent was not an initiate in the past 12 months for the misuse of prescription pain relievers and the only prescription drug that he or she misused in the past 12 months was OxyContin[®]. If this respondent reported first misusing OxyContin[®] in July 2016 and was interviewed on July 15, 2016, then the maximum number of days that the respondent could have misused OxyContin[®] in the past 30 days since his or her first misuse would be 15 days. This situation also held if the respondent *was* an initiate in the past 12 months, misused only OxyContin[®] in that period, and initiated misuse within 15 days of his or her interview date. For both scenarios, a report of misuse of prescription pain relievers on more than 15 days in the past 30 days would be inconsistent with the respondent's data for the first misuse of OxyContin[®]. When these responses occurred, the 30-day frequency (e.g., PNRNM30FQ) was "trimmed" back to the value that would be consistent with the maximum number of days that the respondent could have misused prescription drugs in that category in the past 30 days. In the example cited in this paragraph, PNRNM30FQ would be trimmed back to a value of 15.

6.3.7.2 Misuse of Prescription Drugs with Alcohol in the Past 30 Days

Respondents who reported that they misused any prescription drug in a given category in the past 30 days and who also reported previously in the alcohol section that they last used alcohol in that period were asked whether they misused prescription drugs in that category in the past 30 days while they were drinking alcohol or within a couple of hours of drinking (e.g., question PRM03 for pain relievers; edited variable PNRNM30AL). In addition to

respondents not being asked this question if they did not misuse a particular category of prescription drugs in the past 30 days, respondents were skipped out of this question if they did not report alcohol use in the past 30 days. If respondents reported misuse in the past 30 days for a given category of prescription drugs but the alcohol recency variable ALCREC indicated unambiguously that the most recent use of alcohol occurred more than 30 days ago (i.e., ALCREC = 2 or 3), then the edited variable for misuse with alcohol in the past 30 days (e.g., PNRNM30AL) was assigned a code of 99 (LEGITIMATE SKIP). Similarly, if respondents never used alcohol but they misused prescription drugs in a given category in the past 30 days, then the edited variable for misuse with alcohol in the past 30 days was assigned a code of 99.

However, if respondents indicated misuse in the past 30 days, the 30-day frequency was skipped, and it was unknown whether respondents misused alcohol in the past 30 days, then the 30-day variable for misuse with alcohol retained a code of 98 (blank). For example, if ALCREC had been coded as 9 because respondents last used alcohol at some point in their lifetime (and potentially in the past 30 days) and respondents misused pain relievers in the past 30 days (i.e., PNRNMREC = 1), then PNRNM30AL retained a code of 98 if the corresponding question (e.g., PRM03) had been skipped.

In the assignment of codes when respondents legitimately skipped out of a question for the misuse of prescription drugs with alcohol in the past 30 days, the edits described at the beginning of Section 6.3.7 when respondents did not misuse prescription drugs in the past 30 days took precedence over situations in which respondents did not use alcohol in the past 30 days. Thus, for example, if ALCREC had been coded as 9 (last used alcohol at some point in the lifetime) but respondents last misused prescription drugs in a given category more than 30 days ago but within the past 12 months (e.g., PNRNMREC = 2 for pain relievers), then the edited variable for misuse with alcohol (e.g., PNRNM30AL) could be assigned a code of 93 despite it not being known whether the respondents used alcohol in the past 30 days.

Also, respondents could originally report that they last used alcohol in the past 30 days but report that they used alcohol on 0 days in that period. As noted in [Table 6.2](#), ALCREC was coded as 8 (Used at some point in the past 12 months) when this occurred. However, when ALCREC had been coded as 8 in this situation, the CAI logic for 2016 used respondents' original answer of alcohol use in the past 30 days (e.g., ALLAST3 = 1) to determine that respondents who reported misuse of prescription drugs in the past 30 days were eligible to be asked about their misuse with alcohol in that period. An answer of "yes" to a question about misuse with alcohol in the past 30 days (e.g., PRM03 = 1) would logically mean that the most recent use of alcohol occurred in the past 30 days. When ALCREC had been coded as 8, however, the final imputed alcohol recency IRALCRC could indicate that the most recent use occurred more than 30 days ago but within the past 12 months (IRALCRC = 2). In this situation, a report of misuse of prescription drugs with alcohol in the past 30 days would be inconsistent with the final value in IRALCRC.

Therefore, if ALCREC had been set to 8 and the question for a given prescription drug category for misuse with alcohol in the past 30 days had been answered as "yes," then a value of 10 was added to the code of 1 for the report of misuse with alcohol in the past 30 days

(i.e., "bump" by 10).¹⁰⁰ Thus, for example, PNRNM30AL would be coded as 11 in this situation, where 11 = Yes (ALCREC = 8). This edit would allow analysts to decide whether to take into account or disregard the reports of misuse with alcohol in the past 30 days.¹⁰¹

No editing was done to variables for misuse with alcohol in the past 30 days if ALCREC had been set to 8 but the corresponding questions for misuse with alcohol has been answered as "no." In this situation, respondents could answer the question as "no" because they did not use alcohol in the past 30 days.

6.3.8 Ways of Misuse in the Past 12 Months

Respondents who misused one or more prescription drugs in a given category in the past 12 months were asked to report all of the ways in which they misused these prescription drugs in the past 12 months (e.g., question PRYWAYS for pain relievers). These questions applied to all prescription drugs that respondents misused in the past 12 months and at any point in that period. Response options in PRYWAYS for pain relievers as an example were (1) I used prescription pain relievers without a prescription of my own, (2) I used prescription pain relievers in greater amounts than they were prescribed, (3) I used prescription pain relievers more often than they were prescribed, (4) I used prescription pain relievers for longer than they were prescribed, and (5) I used prescription pain relievers in some other way a doctor did not direct me to use them. The examples of misuse that had been given to respondents earlier in the interview referred to use in greater amounts, more often, or longer than told to take a drug as a single component of misuse. However, cognitive testing of a single response option for use in any of these ways indicated that some respondents thought that they needed to have overused prescription drugs in all three of these ways in order to choose this option. Therefore, separate response options were created for each way in which respondents could overuse medications to encourage more complete and accurate reporting.

These questions for the ways in which respondents misused prescription drugs in the past 12 months were "enter all that apply" questions. Therefore, separate variables were created that corresponded to each response option. The edited variables for the response options in PRYWAYS were PNRWYNORX (use without a prescription), PNRWYGAMT (use in greater amounts than directed), PNRWYOFTN (use more often than directed), PNRWYLNGR (use longer than directed), and PNRWYOTWY (use in some other way not directed). As noted in Section 2.4.4, these variables were coded as 1 (Yes) if a particular way of misuse was reported from the list or as 2 (No [not entered]) if a particular way of misuse was not chosen but at least one of the other ways of misuse was chosen from the list.

¹⁰⁰ This edit will continue to be in place for the 2016 NSDUH. Because only a small number of respondents in 2015 had this pattern (i.e., fewer than 10), respondents in 2016 who reported misuse in the past 30 days and previously confirmed the use of alcohol on 0 days in that period continued to be eligible for receiving the questions about misuse with alcohol. However, this logic has been changed for the 2017 NSDUH to make respondents ineligible for the questions about misuse of prescription drugs with alcohol in the past 30 days if they confirmed use of alcohol on 0 days in that period.

¹⁰¹ The CAI logic has been changed for 2017 to skip respondents out of the questions about misuse of prescription drugs with alcohol in the past 30 days if respondents reported alcohol use on 0 days in the past 30 days. Therefore, the edit described in this sentence is not expected to be implemented in 2017.

A key component of editing the variables for the ways in which respondents misused prescription drugs in the past 12 months involved comparing these variables with the corresponding recency variable for misuse (e.g., PNRNMREC for pain relievers).

- If the recency had a value of 3 to indicate that the most recent misuse definitely occurred more than 12 months prior to the interview, then the skipped variables pertaining to the ways of misuse in the past 12 months were assigned a code of 93 (DID NOT MISUSE IN THE PAST 12 MONTHS).
- If the recency had been assigned a code of 83 because the only misuse in the past 12 months involved the misuse of OTC drugs (Sections 6.3.4.2 and 6.3.5), then any responses in the questions for the ways of misuse in the past 12 months were overwritten with a code of 83 in the edited variables (DID NOT MISUSE IN THE PAST 12 MONTHS Logically assigned). If certain ways of misuse had not been chosen from the list, then the corresponding edited variables were assigned a code of 93.
- If the recency had a value of 91 to indicate that respondents had never misused (or used) any prescription drug in that category in their lifetime, then the skipped variables for the ways of misuse in the past 12 months were assigned a code of 91 (NEVER USED/MISUSED).
- If the recency had a value of 9, then skipped variables for misuse in the past 12 months retained a code of 98 (blank).
- Otherwise, if the recency had a value of 98 (i.e., blank), then the variables for the ways of misuse in the past 12 months were assigned a missing value code to indicate that lifetime misuse or no lifetime misuse was unknown. If the variables for the ways of misuse had a code of 98 (blank), then these variables retained that code of 98. Otherwise, values were overwritten with a code of 85 (BAD DATA Logically assigned).

In addition, respondents could report that the only way that they misused prescription drugs in a given category in the past 12 months was without a prescription of their own. However, respondents also could report that they got their last prescription drug through a prescription, either from one doctor or from more than one doctor. Although it would be inconsistent for respondents to have misused prescription drugs only without a prescription of their own in the past 12 months and to have gotten the last prescription drug that they misused through a prescription, no editing was done to the data for the ways of misuse. Instead, a flag variable was set (e.g., PNRNORXFG) to alert analysts when this pattern occurred.

6.3.9 Injection of Prescription Stimulants in the Past 12 Months or Past 30 Days

Beginning in 2015, respondents who reported that they misused prescription stimulants at any point in the past 12 months were asked whether they used a needle to inject prescription stimulants at any time in this period (question STYNDL). Respondents who answered STYNDL as "yes" and who reported that they misused stimulants in the past 30 days were asked in question STYNDL30 to report when they last used a needle to inject stimulants (i.e., within the past 30 days or more than 30 days ago but within the past 12 months); logically, respondents

who misused stimulants more than 30 days ago but within the past 12 months would not have used a needle to inject stimulants in the past 30 days.

The edited variables corresponding to these questions were STMNDLYR (for question STYNDL) and STMNDLREC (for STNDL30). STMNDLYR and STMNDLREC were edited according to the recency of misuse of prescription stimulants (STMNMREC).

- The following edits were implemented if respondents misused stimulants in the past 30 days (STMNMREC = 1):
 - If STYNDL was answered as "yes" and STNDL30 indicated misuse with a needle in the past 30 days (STNDL30 = 1), then both STMNDLYR and STMNDLREC were coded as 1. A code of 1 meant "Yes" for STMNDLYR and "Within the past 30 days" for STMNDLREC.
 - If STYNDL was answered as "yes" and STNDL30 indicated misuse with a needle more than 30 days ago but within the past 12 months (STNDL30 = 2), then STMNDLYR was coded as 1 (Yes). STMNDLREC was assigned a code of 2 (More than 30 days ago but within the past 12 months).
 - If STYNDL was answered as "no," then STMNDLYR was coded as 2 (No), and STMNDLREC was assigned a code of 99 (LEGITIMATE SKIP).
- The following edits were implemented if respondents last misused stimulants more than 30 days ago but within the past 12 months (STMNMREC = 2):
 - If STYNDL was answered as "yes," then STMNDLYR was coded as 1. Because STYNDL30 had been skipped, STMNDLREC was assigned a code of 12. This code of 12 had the following meaning: More than 30 days ago but in the past 12 months (STMNMREC=2).
 - If STYNDL was answered as "no," then STMNDLYR was coded as 2, and STMNDLREC was assigned a code of 99.
 - If STYNDL was answered as "don't know" or "refused," then STMNDLYR retained the relevant missing data code (e.g., 97 for "refused"), and STMNDLREC retained a code of 98 (blank).
- Most respondents who were logically assigned to have last misused stimulants at some point in the past 12 months (STMNMREC = 8) answered STYNDL as "no." Consequently, STMNDLYR was coded as 2, and STMNDLREC was assigned a code of 99. However, some respondents in 2016 with STMNMREC = 8 answered STYNDL as "refused." In this situation, the preceding edit applied when STYNDL was refused. No respondents in 2016 reported use of stimulants with a needle in the past 30 days when STMNMREC had a value of 8.
- If respondents last misused stimulants more than 12 months ago (STMNMREC = 3), then STMNDLYR and STMNDLREC were both assigned a code of 93 (DID NOT MISUSE IN THE PAST 12 MONTHS).
- If STMNMREC had been assigned a code of 83 because the only misuse in the past 12 months involved the misuse of OTC drugs (Sections 6.3.4.2 and 6.3.5), then any

responses STYN DL and STYN DL30 were overwritten with a code of 83 in the edited variables STMNDLYR and STMNDLREC, respectively (DID NOT MISUSE IN THE PAST 12 MONTHS Logically assigned). If STYN DL and STYN DL30 had been skipped, then the edited variables were assigned a code of 93.

- If STMNMREC had a value of 91 to indicate that respondents had never misused (or used) prescription stimulants in their lifetime, then STMNDLYR and STMNDLREC were assigned a code of 91 (NEVER USED/MISUSED).
- If STMNMREC had a value of 9 and the stimulant needle use variables had been skipped, then STMNDLYR and STMNDLREC retained a code of 98 (blank).
- Otherwise, STMNDLYR and STMNDLREC were assigned a code of 98 (blank) to indicate that lifetime misuse or no lifetime misuse was unknown.

6.3.10 Identification of the Last Prescription Drug Misused in the Past 12 Months

Questions about respondents' reasons for misuse and how they obtained prescription drugs pertained to the last prescription drug in a given category that respondents misused in the past 12 months. For example, if a respondent who misused prescription pain relievers in the past 12 months reported that he or she misused Vicodin® the last time, then question PRYMOTIV asked the respondent to think about the last time that he or she misused Vicodin® in the past 12 months and to report the reasons he or she misused Vicodin® the last time.

If respondents reported that they misused more than one prescription drug in a category in the past 12 months, they were asked to report which drug they misused the last time (e.g., question PRYLAST for pain relievers). In these questions, the response options for the prescription drugs that respondents misused in the past 12 months were filled in sequentially according to the drugs that respondents reported misusing. Consequently, the response options could have different meanings according to the specific prescription drugs that respondents misused. For example, if a respondent misused Vicodin®, generic hydrocodone, and OxyContin® in the past 12 months, then response option 1 in PRYLAST was filled with "Vicodin," response option 2 was filled with "Hydrocodone (generic)," and response option 3 was filled with "OxyContin." If a second respondent misused generic hydrocodone, OxyContin®, and generic morphine in the past 12 months, then response option 1 in PRYLAST was filled with "Hydrocodone (generic)," response option 2 was filled with "OxyContin®," and response option 3 was filled with "Morphine (generic)."

In order for variables pertaining to the last prescription drug that was misused in the past 12 months to have one, and only one, meaning for all respondents (e.g., category 1 for pain relievers always pertaining to Vicodin® as the last pain reliever that was misused), corresponding recoded variables for the last prescription drug were created as part of the CAI program (e.g., RPRYLAST, corresponding to PRYLAST). For example, if "Hydrocodone (generic)" had been reported in PRYLAST as the last prescription pain reliever that a respondent misused (regardless of the drug's order in the list), then RPRYLAST was coded as 5 because generic hydrocodone was the fifth pain reliever in the screener for pain relievers.

The recoding in the CAI program also took into account situations in which respondents reported the misuse of only one prescription drug in that category in the past 12 months. For

these respondents, the CAI program automatically filled in the name of that prescription drug in the questions about the reasons for misuse and how respondents obtained the last prescription drug that they misused. For example, if a respondent reported that the only pain reliever he or she misused in the past 12 months was generic hydrocodone, then RPRYLAST was coded as 5 (Hydrocodone [generic]).

These recoded variables that were created from the CAI program were used as the starting point for creating the edited variables for the last prescription drug that was misused in the past 12 months. As noted in Section 1.4.1.8, Roxicet® and Actiq® were removed in 2016 as examples of prescription pain relievers, and buprenorphine plus naloxone was added as a pain reliever example. The following tranquilizers also were removed from the 2016 questionnaire: buspirone, hydroxyzine, and meprobamate. For stimulants, Ritalin® SR was dropped from the 2016 questionnaire, such that questions in 2016 referred to Ritalin® LA (i.e., rather than "Ritalin SR or Ritalin LA"). Therefore, the levels in the edited variables in 2016 for the last pain reliever and the last tranquilizer did not correspond to the levels in 2015. With the deletion of Roxicet® in 2016, for example, Roxicodone® was shown as level 10 in the edited variable in 2015 and as level 9 in 2016. For tranquilizers, the last response for "Some other tranquilizer" was shown as level 16 in 2015 and as level 13 in 2016, with the deletion of buspirone, hydroxyzine, and meprobamate. For the variable pertaining to the last stimulant that was misused, documentation for level 8 changed from "Ritalin SR or Ritalin LA" in 2015 to "Ritalin LA" in 2016. For these reasons, the edited variables for the last pain reliever, tranquilizer, and stimulant that were misused underwent the changes shown below.

- The name for the variable for the last pain reliever that was misused changed from PNRNMLAST in 2015 to PNRNMLAS1 in 2016.
- The name for the variable for the last tranquilizer that was misused changed from TRQNMLAST in 2015 to TRQNMLAS1 in 2016.
- The name for the variable for the last stimulant that was misused changed from STMNMLAST in 2015 to STMNMLAS1 in 2016.

The name for the last sedative that was misused continued to be SEDNMLAST in 2016 because no changes were made to the specific sedatives in the 2016 questionnaire.

Minimal additional editing was done to these variables. However, if respondents did not misuse prescription drugs in a given category in the past 12 months, then the edits that were described previously in Section 6.3.8 were implemented. For example, if respondents last misused prescription pain relievers more than 12 months ago (PNRNMREC = 3), then PNRNMLAS1 was assigned a code of 93.

6.3.11 Reasons for the Last Misuse of Prescription Drugs in the Past 12 Months

As noted in Section 6.3.10, respondents were asked to report the reasons they misused the last prescription drug in a category that they misused in the past 12 months (e.g., question PRYMOTIV for pain relievers). Respondents who reported "some other reason" as a reason for their last misuse were asked to specify this reason (e.g. question PRYMOTOT for pain

relievers). If respondents reported more than one reason, they were asked to report which one of these reasons was the main reason for their misuse (e.g., PRYMOT1 for pain relievers).

The response options for the reasons for misuse varied by prescription psychotherapeutic category. Response options for pain relievers were (1) to relieve physical pain, (2) to relax or relieve tension, (3) to experiment or to see what the drug is like, (4) to feel good or get high, (5) to help with my sleep, (6) to help me with my feelings or emotions, (7) to increase or decrease the effect(s) of some other drug, (8) because the respondent is "hooked" or has to have the drug, or (9) for some other reason. The same response options were presented for tranquilizers and sedatives, except that "to relieve physical pain" was not presented as an option; the first response option for both of these psychotherapeutic categories was "to relax or relieve tension." Response options for stimulants were (1) to help lose weight, (2) to help concentrate, (3) to help be alert or stay awake, (4) to help study, (5) to experiment or to see what the drug is like, (6) to feel good or get high, (7) to increase or decrease the effect(s) of some other drug, (8) because the respondent is "hooked" or has to have the drug, or (9) for some other reason.

Although the response options varied by psychotherapeutic category, a common set of codes was used for coding the "OTHER, Specify" data for other reasons for misuse. [Table 6.4](#) lists the codes for "OTHER, Specify" data that corresponded to the response options that were presented to respondents for reasons for misuse. The use of a common set of codes for the reasons for misuse facilitated the use of the computer-assisted procedures described in Section 2.3.1 for the coding of "OTHER, Specify" data. For example, if a respondent specified relief of physical pain as a reason for misuse, this response was coded as 1 regardless of whether the respondent specified it as a reason for the misuse of pain relievers or as some other reason for the misuse of tranquilizers.

Table 6.4 "OTHER, Specify" Codes for Reasons for Misuse that Corresponded to Available Response Options

Code	Description	Comments
1	To relieve physical pain	Response option 1 for pain relievers. Not available as a precoded response option for the other prescription drugs.
2	To relax or relieve tension	Response option 2 for pain relievers and response option 1 for tranquilizers and sedatives. Not available as a precoded response option for stimulants.
3	To experiment or to see what it's/they're like	Response option 3 for pain relievers, response option 2 for tranquilizers and sedatives, and response option 5 for stimulants.
4	To feel good or get high	Response option 4 for pain relievers, response option 3 for tranquilizers and sedatives, and response option 6 for stimulants.
5	To help with my sleep	Response option 5 for pain relievers and response option 4 for tranquilizers and sedatives. Not available as a precoded response option for stimulants.

Table 6.4 "OTHER, Specify" Codes for Reasons for Misuse that Corresponded to Available Response Options (continued)

Code	Description	Comments
6	To help me with my feelings or emotions	Response option 6 for pain relievers and response option 5 for tranquilizers and sedatives. Not available as a precoded response option for stimulants.
7	To increase or decrease the effect(s) of some other drug	Response option 7 for pain relievers and stimulants and response option 6 for tranquilizers and sedatives.
8	Because I am "hooked" or I have to have it/them	Response option 8 for pain relievers and stimulants and response option 7 for tranquilizers and sedatives.
9	To help me lose weight	Response option 1 for stimulants. Not available as a precoded response option for the other prescription drugs.
10	To help me concentrate	Response option 2 for stimulants. Not available as a precoded response option for the other prescription drugs.
11	To help me be alert or stay awake	Response option 3 for stimulants. Not available as a precoded response option for the other prescription drugs.
12	To help me study	Response option 4 for stimulants. Not available as a precoded response option for the other prescription drugs.
13	I used it/them for some other reason	Response option 9 for pain relievers and stimulants and response option 8 for tranquilizers and sedatives. However, this was a "placeholder" code that was not used in the actual coding of "OTHER, Specify" data. Instead, new codes for other reasons started with level 14.

A key component of editing the variables for the reasons for the last misuse of prescription drugs in a given category in the past 12 months involved comparing these variables with the corresponding recency variable for misuse (e.g., PNRNMREC for pain relievers). If respondents did not misuse prescription drugs in a given category in the past 12 months, then the edits that were described previously in Section 6.3.8 were implemented. For example, if respondents last misused prescription pain relievers more than 12 months ago (PNRNMREC = 3), then the edited variables for the reasons for that last misuse of pain relievers were assigned codes of 93 (or 993 for the "OTHER, Specify" variable PNRRSOTRS for other reasons for misuse). Remaining edits that are described in this section applied to respondents who last misused prescription drugs in a category in the past 12 months.

Because these questions for the reasons for misuse of the last prescription drug in a given category in the past 12 months were "enter all that apply" questions, separate variables were created for each psychotherapeutic category that corresponded to each response option. As noted

in Section 2.4.4, these variables were coded as 1 (Yes) if a particular reason for misuse was reported from the list or as 2 (No [not entered]) if a particular reason for misuse was not chosen but at least one of the other reasons for misuse was chosen from the list. If respondents did not choose "some other reason" from the list, then the edited variable corresponding to the "OTHER, Specify" response (e.g., PNRRSOTRS for pain relievers) was assigned a code of 999 (LEGITIMATE SKIP).

No editing of the variables for any reasons for misuse was done if respondents reported multiple reasons for their misuse of a single prescription drug the last time they misused it in the past 12 months. However, "OTHER, Specify" data were used to edit the variables for any reasons for misuse. For example, if respondents reported that they misused pain relievers for some other reason and this other reason was to relieve physical pain (i.e., PNRRSOTRS coded as 1) and this was not chosen as a reason in PRYMOTIV, then the edited variable PNRRSPAIN for the misuse of pain relievers to relieve physical pain was assigned a code of 3 (Yes LOGICALLY ASSIGNED). Similarly, if respondents reported that their other reason for misusing tranquilizers was to relax or relieve tension (i.e., TRQRSOTRS coded as 2 for the edited "OTHER, Specify" variable) and this was not chosen as a reason in TRYMOTIV, then the edited variable TRQRSRELX for the misuse of tranquilizers to relax or relieve tension was assigned a code of 3.

As noted previously, if respondents reported more than one reason for their last misuse, they were asked to report which reason was their main reason (e.g., PRYMOT1 for pain relievers). As was the case with the question for the last prescription drug that was misused in the past 12 months, however, if respondents reported more than one reason for their last misuse, the response options for the reasons for the last misuse were filled in sequentially according to the reasons that respondents gave. Consequently, the response options could have different meanings according to the specific reasons that respondents reported.

In order for variables pertaining to the main reason for the last misuse for a given prescription drug category to have one, and only one, meaning for all respondents (e.g., response option 5 the main reason for the misuse of pain relievers always pertaining to misuse to help with sleep), corresponding recoded variables for the main reasons for the last misuse were created as part of the CAI program (e.g., RPRYMOT1, corresponding to PRYMOT1). For example, if misuse to help with sleep had been reported in PRYMOT1 as the main reason for the last misuse of pain relievers, then RPRYMOT1 was coded as 5, regardless of that reason's order in the list from PRYMOT1.

These recoded variables that were created from the CAI program were used as the starting point for creating the edited variables for the main reason for the last misuse. When respondents reported only one reason for their last misuse, then the edited variable (e.g., PNRRSMAIN for the main reason for the last misuse of pain relievers) was assigned a code of 99 (LEGITIMATE SKIP).

"OTHER, Specify" data for other reasons for the last misuse also were used to edit the data for the main reason for misuse. Editing of the variables for the reason for last misuse took into account the available response options for a given prescription drug category. For example, the first response option for the reasons for the last misuse of tranquilizers was "To relax or

relieve tension." Therefore, if a respondent reported that the "other" reason for misusing the last tranquilizer in the past 12 months was to relax or relieve tension (TRQRSOTRS = 2) and the respondent reported that the main reason for the last misuse was, "The other reason I reported," then TRQRSMAIN was coded as 11 (To relax or relieve tension LOGICALLY ASSIGNED).

In addition, when respondents reported more than one reason for their last misuse and they reported that the other reason they reported was the most important reason for their last misuse, the "OTHER, Specify" data for this other reason could have a missing value (i.e., bad data, "don't know," or "refused"). When this occurred, the edited variable for the main reason for misuse was coded as bad data (e.g., PNRRSMAIN = 85 for the main reason for the last misuse of pain relievers) because the respondent did not substantiate what this "most important" other reason was. For example, if a respondent reported misusing pain relievers the last time for some other reason but then reported that he or she did not know what this other reason was, it would not make logical sense for the respondent to report that this was the most important reason for his or her last misuse.

6.3.12 Sources of Prescription Drugs for the Last Misuse in the Past 12 Months

Questions about how respondents obtained the last prescription drugs they misused in the past 12 months were moved in 2015 from the prior substance use section to the relevant prescription drug section earlier in the interview. For all of the prescription psychotherapeutic drug categories, respondents were given a list of eight potential sources of prescription medications. These sources included prescriptions from one doctor, prescriptions from more than one doctor, thefts from medical facilities, prescription drugs obtained from friends or relatives (getting them from friends or relatives for free, buying them from friends or relatives, or taking them from friends or relatives without asking), or purchases from a drug dealer or other stranger. Respondents also had the option of indicating that they obtained the last prescription drug in a given category that they misused in the past 12 months "in some other way." Respondents who indicated that they obtained their last prescription drug in some other way were asked to specify what that other way was.

Respondents who reported that they obtained their last prescription drug in a given category from a friend or relative for free were asked a follow-up question (e.g., PRYFRL for pain relievers) to report how the friend or relative originally obtained the drug. Similar to their predecessor questions, respondents were given the same list of eight sources for where a friend or relative could have obtained the last prescription drug that the respondent misused. Respondents also had the option of indicating that the friend or relative obtained this prescription drug "in some other way"; in that situation, respondents then were asked to specify the other way that the friend or relative obtained the medication.

A key component of editing the variables for the source of the prescription drug in a given category that respondents last misused in the past 12 months involved comparing these variables with the corresponding recency variable for misuse (e.g., PNRNMREC for pain relievers). If respondents did not misuse prescription drugs in a given category in the past 12 months, then the edits that were described previously in Section 6.3.8 were implemented. For example, if respondents last misused prescription pain relievers more than 12 months ago (PNRNMREC = 3), then the edited variables for how they obtained their last pain reliever and

how a friend or relative obtained the last pain reliever that respondents misused were assigned codes of 93. Remaining edits that are described in this section applied to respondents who last misused prescription drugs in a category in the past 12 months.

Respondents who misused prescription drugs in a given category in the past 12 months were asked to report the one way that they obtained the last prescription drug that they misused. If respondents got the last prescription drug they misused in more than one way, they were asked to report one of the ways as their best answer. Thus, for the edited variables PNRGOTLAS, TRQGOTLAS, STMGOTLAS, and SEDGOTLAS (pertaining to how respondents got the last prescription pain relievers, tranquilizers, stimulants, or sedatives, respectively, the last time they misused them), codes in these variables corresponded to those in the corresponding questions PRYGOT, TRYGOT, STYGOT, and SVYGOT.

If respondents reported that they obtained the last prescription drug from a friend or relative for free in the past 12 months (e.g., PNRGOTLAS for pain relievers), similar edits applied to the variables pertaining to how friends or relatives obtained the medication that respondents last misused. Respondents could indicate only one way that the friend or relative obtained the prescription drug that respondents last misused.

Miscellaneous edits based on skip logic also applied to the source of psychotherapeutics variables. These edits applied to all of the prescription psychotherapeutic drugs, with situations being cited for pain relievers. For example, if respondents who misused prescription pain relievers in the past 12 months reported a way in which they obtained the last pain reliever they misused other than "in some other way," then the corresponding "OTHER, Specify" variable (e.g., PNRGOTOTH, corresponding to question PRYGOTSP) was assigned a code of 99 (LEGITIMATE SKIP). In addition, if respondents reported a way in which they obtained the last pain reliever other than from a friend or relative for free, then the edited variables PNRFFGOT and PNRFFGTSP (corresponding to questions PRYFRL and PRYFRLSP, respectively) were assigned a code of 99.

Levels 1 through 7 in the "OTHER, Specify" variables were used for responses that corresponded to existing response options. For example, a code of 4 was assigned to PNRGOTOTH (corresponding to question PRYGOTSP) if respondents reported in PRYGOTSP that the other way they got the last pain reliever they misused in the past 12 months was from a friend or relative for free. Because it could be inferred that these respondents got the last pain reliever they misused from a friend or relative for free, PNRGOTLAS was assigned a code of 14 (Got from a friend or relative for free LOGICALLY ASSIGNED).

[Table B.10](#) in Appendix B describes additional edits that were relevant to the source of psychotherapeutics variables. For these edits, the source of pain relievers variables are used as examples, although these edits also applied to the other prescription psychotherapeutic drugs. As noted in Section 6.3.8, for example, a flag was set (e.g., PNRNORXFG for pain relievers) when respondents reported that they got their last prescription drug through a prescription (i.e., from one doctor or from more than one doctor) but the only way in which they misused prescription drugs in a given category in the past 12 months was without a prescription of their own. Where relevant, the edits that are described in [Table B.10](#) for how respondents obtained the last

prescription drug they misused in the past 12 months also were applied to the variables for ways that friends or relatives obtained the last prescription drug that respondents misused.

6.3.13 Prescription Drug Edit Rules

The editing and imputation teams have been taking small steps toward better integration of procedures. Efforts have focused mainly on prescription drugs because the prescription drug modules that were redesigned for the 2015 survey are structured very logically. Also, with the redesign of the prescription drug questions in 2015, these efforts would not disrupt trends. Therefore, the editing and imputation teams have worked with SAMHSA to develop "edit rules" that directly identify patterns that are unallowable in the final analytic dataset. The code that ensures that data meet the edit rules was shared between the teams as follows:

- Before the edited variables were finalized, code was run with the edited data to identify any unallowable inconsistencies. The final edited data were corrected if any unallowable inconsistencies were identified.
- The code that applies the edit rules was used to create bounds for 30-day frequency variables for all seven modPMN drugs (hallucinogens, inhalants, methamphetamine, and the four prescription drug categories), as described in Section 6.5.2.4.5.1.
- The code that applies the edit rules was used in the imputation programs that assign the age at first misuse and date of first misuse for prescription drugs (Section 6.5.2.6.2.1) to create upper and lower bounds for the date of first misuse. This step simplified the code in part by making it unnecessary to segregate the past year initiates into missing data patterns before dates of first misuse were randomly assigned.
- Before the imputation-revised variables were finalized, a confirmation was run on the imputation-revised data to confirm that no unallowable inconsistencies remained in the data.

The advantages of these small modifications are twofold. First, a smaller dataset can be carried through the modPMN imputation programs because the estimated 30-day frequency, the birth date, the interview date, and the initiation data for misuse are no longer needed. Second, changes across the editing and imputation processes can be implemented more efficiently because common SAS® programs are shared across both processes. The sharing of common SAS® programs also reduces the potential for errors.

The edit rules were decided upon with SAMHSA using three steps.

1. Conceivable inconsistent patterns were listed in a table, along with the number of cases in the 2015 data that followed this pattern.
2. SAMHSA provided decisions on which inconsistent patterns were allowable and which were not.
3. For each unallowable pattern, an edit was decided upon that resolved the inconsistency.

Table 6.5 lists the inconsistent patterns and their resolutions.

Table 6.5 Inconsistent Patterns in the Prescription Drug Modules and their Resolution

Pattern	Inconsistent Pattern	Section References	Action
1	For respondents (Rs) who were known to have initiated misuse in the past month at the category level (e.g., for pain relievers), the 30-day frequency exceeded the number of days that the R could have misused in that period.	6.3.7.1 (Editing) 6.5.2.4.5.1 (Imputation)	The 30-day frequency was trimmed to the maximum number of days that the R could have misused.
2	The R is not a past year initiate, but for each specific drug that was misused in the past year, the R reported a month of first misuse (MFM) and year of first misuse (YFM) that was within 30 days of the interview date. As for Pattern 1, the 30-day frequency exceeded the number of days that the R could have misused in that period.	6.3.7.1 (Editing) 6.5.2.4.5.1 (Imputation)	The 30-day frequency was trimmed to the maximum number of days that the R could have misused.
3	For any specific prescription drug, the age at first misuse (AFM) was inconsistent with the YFM and MFM (i.e., the R failed to resolve the inconsistency when confronted with the consistency check).	6.3.6.2.1 and Table B.6 (Editing) 6.5.2.6.2.1 (Imputation)	The YFM/MFM or both the YFM/MFM and AFM were set to missing, depending on the answers in the consistency checks.
4	For any specific prescription drug, the AFM was inconsistent with the YFM and MFM only because the R's birthday was on the first day of the month and the MFM was in the same month as the birth month.	6.3.6.2.1 (Editing) 6.5.2.6.2.1 (Imputation)	The AFM reported by the R was replaced with an AFM that was 1 year older than the originally reported AFM.
5	The R reported last use of alcohol in the past 30 days but confirmed use of alcohol on 0 days in that period. The R also was routed to the question about misuse of a prescription drug in the past 30 days in combination with alcohol and answered "yes," ALCREC was set to a value of 8 (used at some point in the last 12 months), and ALCDAYS was set to 85 (bad data) as part of the editing procedures. If IRALCRC was imputed to indicate that the most recent alcohol use occurred more than 30 days ago but within the past 12 months (IRALCRC = 2), then misuse in the past 30 days in combination with alcohol would be inconsistent with IRALCRC.	6.2.2.4, 6.3.7.2, and Table 6.2 (Editing) Not applicable for imputation	The edited variable associated with past month misuse in combination with alcohol was set to a value of 11 to alert analysts to this pattern. If IRALCRC indicated alcohol use in the past 30 days (IRALCRC = 1), then the data were not inconsistent between IRALCRC and misuse in the past 30 days in combination with alcohol. With the setting of the variable for misuse in the past 30 days in combination with alcohol to 11, the inconsistency was allowed to remain between IRALCRC = 2 and the reported misuse with alcohol in the past 30 days.

Table 6.5 Inconsistent Patterns in the Prescription Drug Modules and their Resolution (continued)

Pattern	Inconsistent Pattern	Section References	Action
6	The R reported only use without a prescription in the past 12 months (e.g., in PRYWAYS for pain relievers). However, the R also reported getting the last drug as a prescription from one doctor or from more than one doctor	6.3.8 and Table B.10 (Editing) Not applicable for imputation	No editing was done, but a flag was set to alert analysts to this pattern.
7	The R reported more than one reason for misusing the last prescription drug and reported that the "other reason" was the most important reason for misuse. However, the "other" reason that was specified was coded with a missing value (i.e., bad data, "don't know," or "refused").	6.3.11 (Editing) Not applicable for imputation	The edited main reason variable was set to bad data because the R did not substantiate what this "most important" other reason was.
8	The R had his or her last birthday in December 2015, initiated misuse of pain relievers at his or her current age and in December 2015, and was interviewed in January 2016 within 30 days of the R's birthday. The CAI logic routed the R to the question for misuse in the past 30 days, and the R answered the question as "no." Consequently, the R had a recency of misuse other than in the past month.	6.3.7 (Editing) 6.5.2.6.2.1 (Imputation)	If the R initiated the misuse of all prescription drugs in that category in the past 12 months, then the date of first misuse for 2016 was imputed to be consistent with the initiation data but was inconsistent with the recency. In 2017, however, these respondents have been logically inferred to have misused in the past 30 days. The CAI logic also has been corrected for 2018 to identify these respondents as initiates in the past 30 days.
9	The R was interviewed on the last day of the month and was a past year initiate but reported first misuse in the same month as the interview in the previous calendar year.	Not applicable for editing 6.5.2.6.2.1 (Imputation)	The date of first misuse was imputed to be consistent with the initiation data, even though it is technically more than 1 year prior to the interview date. This handling is consistent with the 12-month date range presented to respondents in the interview.

6.4 Additional Edits Applied During Imputation Processing

In addition to the logical edits applied as described previously, edits to selected variables were applied during the imputation process and are discussed in this section. In general, these edits affected only a few records. They were implemented mostly to resolve residual inconsistencies that prevented the determination of a valid interval for the assignment of the date of first use (Sections 6.5.1.4.4 and 6.5.2.6.1).

6.4.1 "Other" Hallucinogens

For respondents who were known to have never used "other" hallucinogens certain logical deductions could be made regarding the relationship between the parent drug data and the child drug data if all the necessary conditions that also are described were met.¹⁰² Note that these edits also could have been applied to respondents who were imputed to lifetime nonuse of the "other" variable.

1. If the respondent was known never to have used "other" hallucinogens, the overall hallucinogens recency was past month, one of the child recencies was missing, and no other child recency was past month or missing, then the child recency that was missing was edited to past month. Similar deductions were made when the overall hallucinogens recency was past year not past month and lifetime not past year.
2. If the respondent was known never to have used "other" hallucinogens including ketamine, the tryptamine hallucinogens DMT/AMT/"Foxy," and *Salvia divinorum*, only one child AFU value was missing,¹⁰³ and the minimum of the nonmissing child AFU values was greater than the parent AFU value, then the missing child AFU value was edited to the parent AFU value.
3. If the respondent was known never to have used "other" hallucinogens including ketamine, the tryptamine hallucinogens DMT/AMT/"Foxy," and *Salvia divinorum*, only one child AFU value was missing,¹⁰⁴ the minimum of the nonmissing child AFU values was equal to the parent AFU value, and the earliest of the nonmissing child months and years of first use was later than the parent MFU and YFU, then the missing child AFU value was edited to the parent AFU value.¹⁰⁵

6.4.2 Respondents Imputed to Lifetime Use for Child Drug Variables

As discussed in Section 6.5, the first substance use variables to undergo imputation consisted of the lifetime drug use measures. The results of these imputations could restrict the range of plausible values for other drug use measures, and, therefore, based on this additional information obtained from the imputations, certain editing rules that were applied to the unedited recency and frequency data had to be reapplied. The list of these edits follows:

1. If the parent drug recency of use was known to be lifetime but not past year, and the respondent was imputed to lifetime use of the child drug(s), then the child drug recency was set to lifetime but not past year. This was done because the respondent could not have used the child drug more recently than the parent drug.
2. If the respondent used the parent drug on exactly 1 day in the past 12 months, the parent drug initiation data indicated a date of first use in the past year, and the respondent was imputed to lifetime use of the child drug, then the child drug recency

¹⁰² The creation of these "other" indicators is described in Section 6.1.2.

¹⁰³ "Only one child AFU value was missing" refers to only LSD, PCP, and Ecstasy. AFU data are not available for other hallucinogen child drugs.

¹⁰⁴ See footnote 32.

¹⁰⁵ These cases occur rarely. Procedures flag cases like these for further examination to ensure the automated edits are applied correctly.

of use was set equal to the parent drug recency of use, and the child drug 12-month frequency of use was set to 1 day. This was done because the respondent could not have used the child drug on any days when the parent drug was not used, so the recencies and frequencies cannot differ. This edit only applied to crack, which is the only child drug with a 12-month frequency.

3. If the parent drug initiation data indicated a date of first use in the past year, the parent drug recency of use was past year but not past month, and the respondent was imputed to lifetime use of the child drug(s), then the recency of use for the child drug was set to past year but not past month. This was done because the respondent could not have used the child drug more recently than the parent drug (eliminating the possibility of past month recency), and the respondent also could not have started using the child drug before the parent drug (eliminating the possibility of lifetime but not past year recency). This edit only applied to crack.
4. Similarly, if the parent drug initiation data indicated a date of first use in the past year, the parent drug recency of use was past month, and the respondent was imputed to lifetime use of the child drug(s), then the recency of use for the child drug was set to past year (whether the respondent had used in the past month could not be determined). This was done because the respondent could not have started using the child drug before the parent drug (eliminating the possibility of lifetime but not past year recency). Again, this edit only applied to crack.

6.4.3 Age-at-First-Use Variables

The edits that are described in this section were implemented in preparation for the imputation process. These edits are designed to reduce processing time and to set the imputation indicators properly (i.e., to "logically assigned" instead of to "statistically imputed"). As with the other edits that were described previously, these edits affected only a small number of records.

1. If the parent AFU value was missing and the minimum of the child AFU values was 3 years, then the parent AFU value was edited to 3 years. This was done because respondents with AFU values of less than 3 years were ineligible to be donors (Sections 6.5.1.4.1 and 6.5.2.4.5.2).¹⁰⁶ This edit applied to all parent AFU variables: cigarettes, overall hallucinogens, and cocaine.
2. If the parent AFU was equal to the respondent's current age, all missing child AFU values were edited to the same age. This edit applied to the AFU variables for daily cigarettes, and crack.¹⁰⁷
3. If the parent AFU was equal to 1 less than the respondent's current age, the child recency¹⁰⁸ was lifetime but not past year (or, for cigarettes, past 3 years but not past

¹⁰⁶ This could be considered to be an imputation rule rather than an editing rule. Nevertheless, once the decision is made to prevent those with age-at-first-use values of less than 3 years from being donors, it is clear that cases like these do not require any sort of stochastic imputation.

¹⁰⁷ For LSD, PCP, and Ecstasy, this edit was accomplished through restrictions on the donor pool during the imputation step.

¹⁰⁸ Because there was no recency question associated with daily cigarettes, the overall cigarette recency was used instead.

year), and the child AFU value was missing, then the child AFU value was assigned to 1 less than the respondent's current age. This was done because the child AFU cannot be less than AGE – 1, because the parent AFU is AGE – 1, and the respondent could not have begun using a child drug before using the parent drug. It also cannot be greater than AGE – 1, because the child drug recency of lifetime but not past year indicates that the respondent did not use the drug while at his or her current age (because he or she did not use the drug at all in the past year). If the respondent did not use the drug at all in the past year, then he or she could not have begun using the drug in the past year. Because the child AFU cannot be less than AGE – 1 or greater than AGE – 1, it must be equal to AGE – 1. This edit also applied to daily cigarettes and crack.

4. If the age at first cigarette use was equal to AGE – 3, cigarette recency was lifetime but not past 3 years, and age at first daily cigarette use was missing, then age at first daily cigarette use was assigned to AGE – 3. The logic is similar to the above edit: the age at first cigarette use precludes the possibility that the age at first daily cigarette use was less than AGE – 3, and the cigarette recency precludes the possibility that the age at first daily cigarette use was greater than AGE – 3.

6.5 Imputation of Substance Use Variables

Two imputation methodologies were applied to the 2016 drug data as described in Chapter 3. The first method, predictive mean neighborhood (PMN), used in the imputation of drug variables beginning in 1999, was applied in a similar manner to previous years for a subset of the drug variables. Drugs imputed using PMN include cigarettes, smokeless tobacco, cigars, pipes, alcohol, marijuana, cocaine and crack, and heroin.

The second imputation method, modified predictive mean neighborhood (modPMN), used in the imputation of some drug variables since 2015 was applied to the remaining drugs including inhalants, hallucinogens, methamphetamine, pain relievers, tranquilizers, sedatives, and stimulants. Note that hallucinogen child drug lifetime measures were imputed using PMN methodology with the exception of the provisional imputation step, which used collocated stochastic imputation (CSI) procedures instead of hot-deck procedures. While these measures appear as set 20 in [Table 6.6](#), these are not imputed until after overall hallucinogen lifetime has been imputed (set 21 in [Table 6.10](#)).

The drug use measures collected in the 2016 NSDUH included lifetime use, recency of any use, recency of use/misuse, frequency of use in the past 12 months, frequency of use/misuse in the past 30 days, past year initiation status, and AFU/AFM, MFU/MFM, and YFU/YFM. However, depending on the drug in question, only a subset of these measures was collected and imputed.

6.5.1 Variables Imputed Using PMN Methodology

[Table 6.6](#) summarizes the drugs and drug use measures that were imputed using the PMN methodology. This table also indicates how these measures were segregated into units referred to as imputation sets. See Section 3.4 for more information on imputation sets.

Because there are numerous sets, some of the set-specific descriptions are explained as deviations from the procedures applied to an earlier set. For example, imputation set 2 is described in detail in Section 6.5.1.3, and the rest of the recency and frequency sets are described in Section 6.5.1.6 as deviations from the procedures described in Section 6.5.1.3.

Because drug use was highly correlated with age, and to facilitate more timely implementation of the imputation procedures, the model building and final assignment of imputed values for tobacco, alcohol, marijuana, cocaine and crack, and heroin use variables were performed separately within three distinct age groups: 12 to 17, 18 to 25, and 26 or older.

Table 6.6 Drugs and Drug Use Measures, PMN Imputation Sets

Drug	Drug Use Measures					
	Lifetime Usage	Recency of Use	12-Month Frequency of Use	30-Day Frequency of Use	Age at First Use	Age at First Daily Use
Cigarettes	NM	Set 2 (12-Month Frequency N/A)			Set 3	Sets 4, 5*
Smokeless Tobacco	Set 1	Set 6 (12-Month Frequency N/A)			Set 7	N/A
Cigars		Set 8 (12-Month Frequency N/A)			Set 9	
Pipes		Set 10 (12-Month and 30-Day Frequency N/A)			N/A	
Alcohol		Set 11			Set 12	
Marijuana		Set 13			Set 14	
Cocaine and Crack		Set 15			Set 16	
Heroin		Set 17			Set 18	
Binge Alcohol Use	N/A	N/A	N/A	Set 19	N/A	
Hallucinogen Child Drugs ¹	Set 20	DM	N/A	N/A	DM	N/A

modPMN = modified predictive mean neighborhood; PMN = predictive mean neighborhood; DM = different method; N/A = not applicable; NM = never missing. Lifetime cigarette use is used to define a unit respondent and is therefore never missing.

*Prior to imputing age at first daily cigarette use, lifetime daily cigarette use must first be imputed.

¹ Hallucinogen child drugs lifetime use variables are imputed using PMN methodology with the exception of the provisional imputation step which uses collocated stochastic imputation procedures. While listed as set 20, imputation is not started until after the overall hallucinogen lifetime use has been imputed (set 21). Other measures for hallucinogen child drugs are imputed using modPMN methodology (indicated with DM). See [Table 6.10](#) for modPMN imputation sets.

6.5.1.1 Lifetime Drug Use (Imputation Set 1)

The lifetime drug use variables were imputed using the single response propensity (RP)/multiple prediction (PRD) type of PMN, as outlined in Section 3.4.3. In general, the response rates for lifetime drug use variables were very high with less than 1 percent of cases requiring imputation. These high response rates were observed, in part, because of the usable case rule that requires that a respondent answer "yes" or "no" to the question on lifetime use of cigarettes and at least nine (9) of the following additional drugs had to contain information about use or nonuse: (1) smokeless tobacco, (2) cigars, (3) alcohol, (4) marijuana, (5) cocaine (in any form), (6) heroin, (7) hallucinogens, (8) inhalants, (9) methamphetamine, (10) pain relievers, (11) tranquilizers, (12) stimulants (i.e., independent of methamphetamine), and (13) sedatives. See Section 2.2.3 for more information about the usable case rule.

Because the single RP/multiple PRD type of PMN was used for the lifetime usage imputations, decisions had to be made on the order in which to fit the PRD models. Drugs later in the sequence would have more covariates in their models, because drugs earlier in the sequence were used as covariates after provisional imputation. The order in which the lifetime indicators of use in set 1 were imputed is shown in [Table 6.6](#), with the exception of lifetime cigarette use.¹⁰⁹ The lifetime use or nonuse of cigarettes was used to define a unit respondent for the NSDUH and, therefore, did not contain any missing values.

6.5.1.1.1 Response Propensity Step

The input to the weight adjustment model in the RP step for imputation set 1 was the preliminary analysis weight, PANALWT. As with the 1999-2015 surveys, the 2016 survey implemented automatic routing of the respondent through the questionnaire based on the respondent's answers, thereby skipping over (i.e., not asking the respondents) specific questions. Within each drug section included in imputation set 1, one question was asked in order to establish whether the respondent had ever used the drug in question during his or her lifetime. For more information on gate questions, see Section 6.2.1.

For an individual to be considered a lifetime use item respondent, he or she must have complete data for all of the following drug section gate questions: cigarettes, cigars, smokeless tobacco, pipes, alcohol, marijuana, cocaine, crack, and heroin. See [Table D.3](#) in Appendix D for details of the covariates used in the RP models for these variables.

6.5.1.1.2 First Prediction Step (Lifetime Smokeless Tobacco Use)

The probability of lifetime smokeless tobacco use was modeled for item respondents within each age group, using the nonresponse-adjusted weights. SUDAAN[®]'s RLOGIST procedure was used to perform dichotomous logistic regression¹¹⁰ to determine the parameter estimates and probability of use for both respondents and nonrespondents.

6.5.1.1.3 First Provisional Hot-Deck Step (Lifetime Smokeless Tobacco Use)

In order to use lifetime usage of a given drug as a covariate for a drug later in the sequence, it was necessary to create temporary imputed values in cases where the original lifetime usage indicator was missing. The lifetime indicator for smokeless tobacco was used as a covariate for later models, so it was necessary to create this provisional value. In the first provisional hot-deck step, matching was done on a single predicted mean from the PRD step, and missing values were replaced with the value from a donor within this neighborhood.

¹⁰⁹ See Section 3.4.2 for a brief discussion of how order is determined for imputation sets that use the multiple RP/multiple PRD or single RP/multiple PRD type of PMN.

¹¹⁰ SAS[®]-callable SUDAAN[®] was used to fit all dichotomous and polytomous logistic regression models. Details about the logistic regression model and additional references can be found in RTI International (2013). SAS software is a registered trademark of SAS Institute Inc. SUDAAN is a registered trademark of Research Triangle Institute.

If possible, donors and recipients were required to be from states with the same level of smokeless tobacco usage (state rank¹¹¹), where the level of usage was defined in terms of the weighted proportion of a given state's residents who were lifetime users of the drug.¹¹² If insufficient donors were available within these constraints, they were loosened in the following order: (1) the delta constraint was removed, and (2) the state-rank constraint was removed, and the delta constraint was reapplied.

No logical constraints were placed on the neighborhoods for any of the lifetime usage indicators.

6.5.1.1.4 Analogous Prediction and Provisional Hot-Deck Steps for Remaining Drugs

PRD models and provisional hot-deck steps were completed in a manner similar to that described above for cigars, pipes, alcohol, marijuana, cocaine and crack, and heroin, with the following deviations:

- Because cocaine and crack were in two separate back-to-back sections in the 2016 NSDUH questionnaire, separate models were fit for the two substances. However, crack is a type of cocaine, so donors for the two substances were obtained using a single neighborhood with multivariate matching.¹¹³ This was true regardless of whether the item nonrespondent was missing only crack, only cocaine, or both crack and cocaine. Once the neighborhood was defined, missing values for crack and/or cocaine were replaced with the values from one donor within this neighborhood.
- For respondents missing cocaine or crack but not both cocaine and crack, donors had to match respondents on the nonmissing value. If the respondent was a lifetime user of cocaine but missing crack, the donor also had to be a lifetime user of cocaine. If the respondent was a lifetime nonuser of crack, then the donor also had to be a lifetime nonuser of crack. It was not possible for a respondent to be a lifetime nonuser of cocaine and have a missing value for lifetime use of crack, because crack is a type of cocaine. Similarly, it was not possible for a respondent to be a lifetime user of crack and have a missing value for lifetime use of cocaine. This constraint was applied during the first two attempts to find a donor but was dropped in subsequent attempts.
- Heroin did not undergo a provisional imputation step, because it was the last variable in the imputation set.

¹¹¹ See Section 3.8 for a general discussion of state-rank variables.

¹¹² Those with a missing lifetime use indicator for the drug were treated as lifetime nonusers.

¹¹³ This provisional hot-deck program actually uses a multivariate delta constraint, but the distance from donor to recipient is based only on the predicted probability of lifetime use of cocaine. This was done to avoid the complexity of the Mahalanobis distance calculation. Strictly, this is neither univariate nor multivariate matching.

6.5.1.1.5 Final Hot-Deck Step

Tables E.20 through E.22 in Appendix E provide details on the final hot-deck step for the lifetime use indicators for imputation set 1. Although the predictive mean vector could be large if several indicators were missing, the hot-deck step included no logical constraints.

No final imputation-revised variables indicating lifetime usage alone were created, because this information was recorded in the final imputation-revised recency-of-use variables. Imputation indicators also were not created, though temporary variables indicating that lifetime usage was imputed were maintained to inform the creation of the recency-of-use imputation indicators.

6.5.1.2 Hallucinogen Child Drug Lifetime Use (Imputation Set 20)

Imputation for the hallucinogen child drug lifetime use variables for imputation set 20 was done in a manner similar to that described above for set 1. The following deviations from the process described for set 1 applied to set 20.

- **Imputation Order:** The seven hallucinogen child drugs were imputed in the following order: LSD, PCP, Ecstasy, ketamine, the tryptamine hallucinogens DMT/AMT/"Foxy," *Salvia divinorum*, and "other" hallucinogens (includes hallucinogens other than those listed).
- **Response Propensity Step:** The input to the weight adjustment model in the RP step was the preliminary analysis weight, PANALWT. The eligible population for the models included lifetime hallucinogen users based on the imputation-revised lifetime variable (imputation set 21). For an individual to be considered an item respondent, he or she must have complete data for lifetime use of LSD, PCP, Ecstasy, ketamine, the tryptamine hallucinogens DMT/AMT/"Foxy," *Salvia divinorum*, and "other" hallucinogens.
- **First Prediction Step:** This step was identical to the process described for set 1.
- **First Provisional CSI Step:** Instead of undergoing a provisional hot-deck step to determine provisional values for use in covariates for later models, the lifetime hallucinogen child drugs underwent a provisional CSI step. See Section 3.6.3 for a general description of the CSI step. Briefly, the PRD model is followed by an imputation step where the predicted probability of lifetime use from the PRD step is used to stochastically impute a value of "Yes" or "No" for each item nonrespondent.
- **Analogous Prediction and Provisional CSI Steps for Remaining Drugs:** PRD models and provisional CSI steps were completed in the manner described above for PCP, Ecstasy, ketamine, the tryptamine hallucinogens DMT/AMT/"Foxy," *Salvia divinorum*, and "other" hallucinogens with one deviation: "Other" hallucinogens did not undergo a provisional imputation step, because it was the last variable in the imputation set.
- **Final Hot-Deck Step:** Tables E.58 through E.61 in Appendix E provide details on the final hot-deck step. A single logical constraint was applied requiring donors to be lifetime users of at least one hallucinogen child drug that the recipient is missing if

the recipient is a nonuser of all nonmissing hallucinogen child drugs or is missing all child drugs. As with set 1, no final imputation-revised variables indicating lifetime usage alone were created.

6.5.1.3 Imputation-Revised Cigarette Recency and Frequency of Use (Imputation Set 2)

As indicated in [Table 6.6](#), the second set of substance use variables to undergo imputation were the cigarette recency and 30-day frequency variables. The multiple RP/multiple PRD type of PMN was used to process these variables.

6.5.1.3.1 Sequence of Imputation

Because recency-of-use and frequency-of-use variables for a given drug were in the same imputation set, the calculation of predicted means for the frequency-of-use variables required the item nonrespondents to be identified as provisional past month and/or past year users. For this reason, cigarette recency was modeled prior to 30-day frequency, and provisional imputations were performed to allow for the identification of provisional past month users of cigarettes.

6.5.1.3.2 First Response Propensity Step (Cigarette Recency of Use)

The input to the weight adjustment model in the RP step for imputation set 2 was the preliminary analysis weight, PANALWT. To impute for missing recency-of-use values for cigarettes, it was first necessary to define the domain within each of the three age groups. Using the imputation-revised lifetime indication of use, the file was reduced to lifetime cigarette users.¹¹⁴ Among these lifetime users, item respondents and nonrespondents were identified across recency-of-use and 30-day frequency-of-use measures. If a valid response was provided for each drug use measure, the individual was deemed an item respondent for cigarettes. Otherwise, he or she was an item nonrespondent. See [Tables D.4](#), [D.5](#), and [D.6](#) in Appendix D for details of the covariates used in the RP models for this variable.

6.5.1.3.3 First Prediction Step (Cigarette Recency of Use)

Using the adjusted weights, the probability of selecting each cigarette recency-of-use category was modeled within each age group using polytomous logistic regression. SUDAAN's MULTILOG procedure was used to estimate the parameters from the appropriate logistic model from which predicted probabilities for each of the cigarette recency categories were calculated for both item respondents and item nonrespondents. The four recency categories were the following:

1. past month;
2. past year, not past month;

¹¹⁴ No imputation was required for lifetime use of cigarettes due to the usable case rule. See Sections 2.2.3 and 6.5.1.1 for more information on the usable case rule.

3. past 3 years, not past year; and
4. lifetime, not past 3 years.

6.5.1.3.4 First Provisional Hot-Deck Step (Cigarette Recency of Use)

In order to define the domain for the cigarette 30-day frequency-of-use variable, it was necessary to create temporary imputed values in cases where the original cigarette recency value was missing. In order to save time and resources and because the imputation was only provisional, a univariate matching procedure was implemented. The only predicted mean used was the predicted probability of past month use, because past month use was the most critical measure of recency of cigarette use.

If possible, donors and recipients were required to be from states with the same level of usage of a given drug (state rank; see Section 3.8), where the level of usage was defined in terms of the weighted proportion of a given state's residents who had used cigarettes in the past month.¹¹⁵ If insufficient donors were available within these constraints, they were loosened in the following order: (1) the delta constraint was removed, and (2) donors and recipients were no longer required to be from states with similar usage levels.

The only logical constraints placed on the neighborhoods involved cases where a general recency category was available for a respondent and imputation was required to determine the specific recency categories. The general recency categories that appeared are discussed in Section 6.2.2.3 and shown in Table 6.7. Logical constraints ensured that only donors with allowable specific recency categories were included in the neighborhood.

Table 6.7 General Incomplete Recency Categories for Cigarettes

General Incomplete Recency Category	Allowable Specific Recency Categories
Lifetime	1. Past month 2. Past year but not past month 3. Past 3 years but not past year 4. Lifetime but not past 3 years
Past Year	1. Past month 2. Past year but not past month
Lifetime, Not Past Year	1. Past 3 years but not past year 2. Lifetime but not past 3 years
Lifetime, Not Past Month	1. Past year but not past month 2. Past 3 years but not past year 3. Lifetime but not past 3 years
Lifetime, Not Past Month but within Past 3 Years	1. Past year but not past month 2. Past 3 years but not past year
Past 3 Years	1. Past month 2. Past year but not past month 3. Past 3 years but not past year

¹¹⁵ Those individuals whose past month use status was unknown were treated as if they were not past month users.

6.5.1.3.5 Second Response Propensity Step (Cigarette 30-Day Frequency)

The input to the weight adjustment model in the RP step for imputation set 2 was the preliminary analysis weight, PANALWT. The modeling of cigarette 30-day frequency followed that of recency. The file was first reduced to the domain, which was past month cigarette users, as defined by the provisional recency variable. Next, item respondents and nonrespondents were defined according to the same criterion used for the cigarette recency imputations. To be an item respondent, the individual had to have provided valid responses to both the cigarette recency and 30-day frequency measures. The item RP adjustment was then computed so that the respondents' weights accurately represented all past month users of cigarettes. See [Tables D.4, D.5, and D.6](#) in Appendix D for details of the covariates used in the RP models for this variable.

6.5.1.3.6 Second Prediction Step (Cigarette 30-Day Frequency)

As stated in the previous section, only past month users of cigarettes were used to build the 30-day frequency-of-use model. The response variable of interest in the 30-day frequency-of-use models, prior to a normalizing transformation, was the proportion of the days in a month (30 days) on which a respondent used cigarettes. The range of values for the proportion was from (greater than) 0 to 1. Hence, to model 30-day frequency of use, the following empirical logit transformation was computed for all respondents:

$$\log[(Y_i + 0.5)/(N - Y_i + 0.5)],$$

where Y_i was the observed 30-day frequency for respondent i and N was 30, the total number of days in the month that the respondent could have used the substance. This transformation was nearly equivalent to the standard logit transformation:

$$Y_i^* = \log[P_i/(1 - P_i)],$$

where P_i was defined as the proportion of days in the past month on which respondent i used the drug. The standard logit transformation was not used because it was not defined for daily users.¹¹⁶ Using the adjusted weights, a linear univariate regression model was then fitted using SUDAAN software for the log-transformed variable Y_i within each age group.

For cigarettes, the empirical distribution for 30-day frequency of use was in fact a mixture distribution, with a positively skewed distribution from 1 to 29 and a spike at 30. This substance was modeled using two separate models. One was a logistic model for daily use versus nondaily use among past month users. For the nondaily past month users (i.e., those who had used between 1 and 29 days), the model described above was used. In this case, the response variable in a linear regression model was a logit of the proportion of the period (30 days) during which a respondent used the substance. Both the predicted probability of daily use and the logit of the proportion of the month used (assuming nondaily use) were used as predicted means in the

¹¹⁶ If the respondent was a daily user of the substance, then $\log[(Y + 0.5)/(N - Y + 0.5)] \approx \log[(N + 0.5)/0.5]$ with $N = 30$ so that it was defined for all respondents. See Cox and Snell (1989) for a discussion of the empirical logit transformation.

subsequent hot-deck step. The logit was back-transformed into a proportion before use in the hot-deck step.

6.5.1.3.7 Final Hot-Deck Step (Cigarette Recency and 30-Day Frequency)

The full predictive mean vector for cigarettes contained probabilities associated with several of the recency-of-use categories, a probability of daily use, and a predicted probability of use on a given day in the past month. Each element in the full vector of predicted means was adjusted so that all elements were conditioned on the same usage status whenever possible. The elements in the predictive mean vector that could have potentially resulted are shown in [Table 6.8](#), with the assumption that only the lifetime usage is known. If other information about the recency of use is known (e.g., past year user), the predictive mean vector is adjusted accordingly. The portion of the full predictive mean vector used to determine the neighborhood for a particular item nonrespondent was dependent on the pattern of missingness for that item nonrespondent. If partial information was available regarding recency of use, then that information was used to adjust the recency-of-use probabilities. The portions of the full predictive mean vector used for each missingness pattern, with accompanying adjustments, are provided in [Table E.27](#) in Appendix E. The Mahalanobis distance was then calculated using only the portion of the predictive mean vector that was associated with the given missingness pattern, with elements appropriately adjusted. The likeness and logical constraints applied to each missingness pattern are also available in [Table E.27](#).

Table 6.8 Elements of Full Predictive Mean Vector: Cigarettes

Drug Use Measure and Category of Interest	Predicted Mean
Recency of Use, Past Month Use ¹	$P(\text{past month user} \mid \text{lifetime user})$
Recency of Use, Past Year but Not Past Month Use ¹	$P(\text{past year but not past month user} \mid \text{lifetime user})$
Recency of Use, Past 3 Years but Not Past Year Use ¹	$P(\text{past 3 years but not past year user} \mid \text{lifetime user})$
30-Day Frequency of Use for Nondaily Users over Past 30 Days	$P(\text{use on a given day in the month} \mid \text{past month user, not a daily user}) \times P(\text{not a daily user} \mid \text{lifetime user}) \times P(\text{past month user} \mid \text{lifetime user})^2$
Daily User over Past 30 Days	$P(\text{daily user} \mid \text{past month user}) \times P(\text{past month user} \mid \text{lifetime user})^2$

¹ The final category for recency (lifetime but not past year or lifetime but not past 3 years) was not needed in the predictive mean vector, because the multinomial probabilities summed to 1, and this probability was determined by the other probabilities.

² Interpreting the proportion of the month used as a probability of use on a given day in the month assumed that the probability of use on each day in the month was equal, which was not true.

6.5.1.3.8 Final Variables (Cigarette Recency and 30-Day Frequency)

The final imputation-revised recency-of-use and 30-day frequency variables were identified with the prefix IR, followed by a five-letter identifier, where a three-letter code identified the drug (CIG) and the final two letters identified the measure (RC = recency of use, FM = 30-day frequency). Each IR variable was accompanied by an imputation indicator with a prefix II instead of IR. The levels for the imputation indicators were the standard levels used for

most imputation-revised variables: 1 = questionnaire data, 2 = logically assigned, 3 = statistically imputed, and 9 = legitimate skip (not a lifetime user).

6.5.1.4 Imputation-Revised Cigarette Age at First Use (Imputation Set 3)

As indicated in Table 6.6, the third imputation set consisted of the cigarette AFU variable. Unlike the recency and 12-month frequency-of-use variables, age at first drug use was not statistically imputed in the surveys prior to 1999. Instead, missing values were excluded from subsequent analyses. However, as with the 30-day frequency, missing AFU values have been replaced using imputation since the 1999 survey. Also, recent drug initiates (i.e., those whose current age was equal to or 1 year greater than the reported AFU) were asked the year and month of their first use. To have this information for all users, both missing MFU and missing YFU for less recent initiates (and recent initiates who did not report MFU and YFU) were replaced by assigning values consistent with the respondent's current age, interview date, imputation-revised AFU, and imputation-revised recency and frequency variables. To have complete date-of-first-use information, day of first use (DFU) was randomly assigned for all users. The combined data gave the respondent's AFU along with the date of first use.

6.5.1.4.1 Response Propensity Step (Cigarette Age at First Use)

The input to the weight adjustment model in the RP step for imputation set 3 was the preliminary analysis weight, PANALWT. To impute for missing cigarette AFU, it was necessary to define the eligible population. Using the imputed recency of use, the files were reduced to lifetime users of cigarettes. If a valid response was provided for the AFU measure,¹¹⁷ the individual was deemed an item respondent. See Tables D.4, D.5, and D.6 in Appendix D for details of the covariates used in the RP models for this variable.

6.5.1.4.2 Prediction Step (Cigarette Age at First Use)

The response variable in the model for AFU, before a normalizing transformation, was the AFU as a proportion of the current age. The numerator in this proportion was an integer representing AFU. However, because this integer was in fact a truncated version of the real AFU, the value was made continuous by adding a random component between 0 and 1. Hence, expressing the proportion as $P_i = Y_i/N_i$, the numerator was given as

$$Y_i = \text{Age at First Use}_i + \text{Uniform}(0,1) \text{ random number}.$$
¹¹⁸

The denominator in the proportion was the total age. The true age was known, based on the interview date and birth date. Expressing it in years rather than days required dividing by the number of days in the year:

¹¹⁷ Respondents who reported an age at first use of 1 or 2 were treated as item nonrespondents in the RP and PRD steps, because of the implausibility of such a young age at first use. In the hot-deck step, their response was left unchanged, but they were not allowed to be donors.

¹¹⁸ In the event that the age at first use was equal to the age, Y_i was constrained so that it was equally likely to be anywhere on the interval $[\text{Age at First Use}_i, N_i]$. Thus, Y_i was prevented from being greater than N_i .

$$N_i = (\text{Interview Date} - \text{Birth Date} + 1)/365.25.$$

After a weight adjustment, the empirical logit transformation was used as the response variable in a weighted linear univariate regression:

$$\log \left[(Y_i + 0.5) / (N_i - Y_i + 0.5) \right].$$

This transformation was nearly equivalent to the standard logit transformation:

$$Y_i^* = \log \left[P_i / (1 - P_i) \right],$$

which was not used, because it might be unstable for respondents who started using at their current age.

One unusual covariate used in the PRD model for cigarette AFU was a modified 30-day frequency variable for cigarettes. It was defined as follows:

$$\begin{aligned} \text{new30}_i &= 0 && \text{if respondent } i \text{ did not use cigarettes in the past month} \\ &= 30\text{-day frequency} && \text{if respondent } i \text{ used cigarettes in the past month} \end{aligned}$$

Naturally, the full model for AFU did not include the lifetime indicator for the drug in question, because the model was built on cigarette users. A summary of the starting and final models can be found in [Tables D.4](#), [D.5](#), and [D.6](#) in Appendix D.

From the final model, a predicted value (based on the Y variable) was calculated for each cigarette user, which was then back-transformed to produce a predicted cigarette AFU.

6.5.1.4.3 Hot-Deck Step

The imputation-revised cigarette AFU assignment was conducted using a single predicted mean: the predicted AFU. [Tables E.43](#) through [E.45](#) provide a complete list of likeness and logical constraints applied to the cigarette AFU imputations. The likeness constraints for AFU were more stringent than those for the other drug use measures. Therefore, it was often necessary to loosen the constraints. Once these likeness constraints were removed, some complex logical constraints remained, based on the interview date, the birth date, and imputation-revised recency and frequency values.

6.5.1.4.4 Date-of-First-Use Assignments

After the AFU imputations, all lifetime users of cigarettes had nonmissing AFU values. Using this AFU, users were assigned values for MFU, YFU, and DFU. Recent initiates, or those respondents whose AFU was within 1 year of his or her age, were asked for their MFU and YFU. The DFU was not collected in the questionnaire and was missing for all respondents. The MFU, YFU, and DFU data contained four patterns of missingness:

Pattern 1: Recent initiates: missing day of first use only;

Pattern 2: Recent initiates: missing month/day of first use;

Pattern 3: Recent initiates: missing year/month/day of first use; and

Pattern 4: Less recent initiates: missing year/month/day of first use.

For each missingness pattern, upper and lower bounds on the date of first use (i.e., the earliest possible date of first use and the latest possible date of first use) were determined. Once the earliest and latest possible dates of first use were determined, a day was randomly selected from this interval. The imputation-revised month/day/year values were then extracted from this date of first use.

6.5.1.4.4.1 Missingness Pattern 1

In this missingness pattern, a recent initiate provided all the information asked by the questionnaire (i.e., both the MFU and YFU). However, to obtain a complete date of first use, a DFU also was needed. Thus, a DFU was randomly assigned, given the respondent's MFU and YFU, in a way that was consistent with both the 30-day frequency/recency and AFU. Below is a brief description of the process used to obtain a date of first use in such cases. The imputed YFU, MFU, and DFU were extracted from the date, as defined below:

Final date of first use = Earliest possible date + [(Days between earliest and latest date) × (a random number generated from a Uniform (0,1) distribution)],

where

Days between earliest and latest date = Latest possible date – Earliest possible date + 1;

Earliest possible date = maximum [(AFUth birthday), (first day of the month indicated by MFU/YFU)]; and

Latest possible date =

- *minimum [(Interview date – 30-day frequency + 1), (1 day before the (AFU + 1)th birthday), (last day of the month indicated by MFU/YFU)], if recency = 1;*
- *minimum [(Interview date – 30), (1 day before the (AFU + 1)th birthday), (last day of the month indicated by MFU/YFU)], if recency = 2; and*
- *minimum [(Interview date – 1 year), (1 day before the (AFU + 1)th birthday), (last day of the month indicated by MFU/YFU)], if recency = 3.*

Note that it is impossible for recent initiates to have recency = 4 (lifetime but not past 3 years). Recent initiates had to have begun using the drug no earlier than their (AFU)th birthday. Because AFU = current age, or AFU = current age – 1, their (AFU)th birthday was within the past 2 years. Respondents who had begun using the drug within the past 2 years must logically have last used the drug within the past 2 years, and therefore could not have had recency = 4.

In rare cases, the *earliest possible date* was set to 29 days before the interview. This occurred for respondents meeting all of the following conditions:

1. The *latest possible date* was within 29 days of the interview.
2. The *earliest possible date* determined by the above rule was within a year of the interview.
3. The recency = 1.
4. The 12-month frequency = 30-day frequency (if applicable), or the 12-month frequency = 1.

Logically, all the lifetime usage of the drug for these respondents occurred in the past 30 days (including the interview date). The first condition ensures that the application of this rule does not cause an inconsistency. The second condition implies that the drug was not used by these respondents more than 1 year ago. The third and fourth conditions imply that the drug was not used by these respondents in the interval (1 year before the interview, 1 month before the interview). Therefore, these respondents did not use the drug more than 1 month ago. All their lifetime use must have occurred in the past month.

6.5.1.4.4.2 *Missingness Pattern 2*

The second missingness pattern occurred when a recent initiate provided his or her YFU but did not provide an MFU. In such cases, a month and day were randomly assigned that were consistent with both the respondent's frequency/recency and with the AFU range. The imputed MFU and DFU were derived in the same manner as the date of first use in Missingness Pattern 1, except with the following changes:

- For the *earliest possible date*, replace "first day of the month indicated by MFU/YFU" with "January 1st of the YFU."
- For the *latest possible date*, replace "last day of the month indicated by MFU/YFU" with "December 31st of the YFU."

6.5.1.4.4.3 *Missingness Pattern 3*

Similar to Missingness Pattern 2, the third missingness pattern occurred when recent initiates provided neither an MFU nor a YFU value. In these cases, the MFU, YFU, and DFU were randomly assigned from a uniform distribution in a way that was consistent with both the cigarette 30-day frequency/recency and the AFU. Again, the imputed MFU, YFU, and DFU were derived in the same manner as described in Missingness Pattern 1.

6.5.1.4.4.4 *Missingness Pattern 4*

The fourth missingness pattern occurred when the respondent reported, or was imputed to, an AFU at least 2 years less than his or her age. This case is analogous to data prior to the 1999 survey, where MFU and YFU were not asked in the questionnaire. In this missingness pattern, the 30-day frequency was immaterial to the final date of first use because the respondent could not have begun using in the past month:

Earliest possible date = AFUth birthday; and

Latest possible date =

- 1 day before the (AFU + 1)th birthday, *if recency* < 4; or
- minimum [(Interview date – 3 years), (1 day before the (AFU + 1)th birthday)], *if recency* = 4.

6.5.1.4.5 Final Age and Date-of-First-Use Variables

The final imputation-revised date-of-first-use variables were identified with the prefix IR, followed by a six-letter identifier, where a three-letter code identified the drug¹¹⁹ and the final three letters identified the measure (AGE = age at first use, MFU = month of first use, YFU = year of first use, DFU = day of first use). Each IR variable was accompanied by an imputation indicator with the requisite II prefix. The levels for the imputation indicators were the standard levels used for most imputation-revised variables: 1 = questionnaire data, 2 = logically assigned, 3 = statistically imputed, and 9 = legitimate skip (not a lifetime user). Because survey respondents are not asked for the specific day on which they first used the drug of interest, all respondents in the domain receive IIxxxDFU = 3. Also, as indicated above, only recent initiates are asked for the year and month of first drug use. Subsequently, these questions have high rates of nonresponse because of the skip logic embedded in the questionnaire, as all other individuals in the domain require imputation for their MFU and YFU.

6.5.1.5 Imputation-Revised Age at First Daily Cigarette Use (Imputation Sets 4 and 5)

In addition to AFU, the cigarettes section also included a question asking for the respondent's age at first daily cigarette use, where a daily user was defined as someone who reported having at some time smoked cigarettes every day for a period of at least 30 days. Imputation procedures for age at first daily cigarette use were similar to AFU, with some key exceptions as discussed below.

One such exception involved the domain of the AFU variable. Whereas the AFU question was asked of all cigarette users, the age-at-first-daily-use question was asked of only daily users. The "daily use" indication came from two sources. If a respondent answered either the 30-day frequency or estimated 30-day frequency with a "30," or if the respondent had a "yes" value for the edited variable associated with the "ever daily used" question (CIGDLYMO), then he or she was considered a daily user. For more information about CIGDLYMO, see Section 6.2.6. The "ever daily used" question (CIGDLYMO) can be thought of as a lifetime "child" drug to the "parent" lifetime cigarette use question (CIGEVER). However, anyone who answers the 30-day frequency or estimated 30-day frequency with a "30" is automatically skipped out of this question. Therefore, it is necessary to obtain the imputation-revised cigarette 30-day frequency (IRCIGFM) prior to imputing the lifetime-daily-cigarette-use variable (IRCDULF) so that it is

¹¹⁹ Exceptions to this rule occurred with marijuana and cigarette daily use. For historical reasons, marijuana contained a two-letter code (MJ). Marijuana variables therefore ended with a five-letter identifier rather than a six-letter identifier. The code for cigarette daily use was CDU, which differed from the general cigarette code of CIG. Details about cigarette daily use are provided in Section 6.5.1.5.

not included with the other lifetime drug indicators as part of imputation set 1. Instead, as indicated in [Table 6.6](#), the age at first daily cigarette use actually contains two separate imputation sets. Imputation set 4 includes the lifetime indicator of daily cigarette use and imputation set 5 includes the age-at-first-daily-cigarette-use variables. At this stage in the process, there should be no missing responses to the 30-day frequency question, which were imputed as part of imputation set 2 as discussed above. Daily users, based on 30-day frequency, should be either known (based on a response in the survey) or imputed. However, responses for the ever-daily-used question (CIGDLYMO) could still be missing, and, therefore, it was first necessary to impute these values to define the domain for the age-at-first-daily-use variable.

6.5.1.5.1 Response Propensity Step (Ever-Daily-Used Cigarettes)

The input to the weight adjustment model in the RP step for imputation sets 4 and 5 was the preliminary analysis weight, PANALWT. To impute for missing values in the ever-daily-used variable, it was necessary to define the domain: lifetime users of cigarettes who had an imputation-revised 30-day frequency¹²⁰ of fewer than 30 days (includes legitimate skip codes for lifetime but not past month users). If a valid response was provided in the ever-daily-used variable, the individual was deemed an item respondent. See [Tables D.4](#), [D.5](#), and [D.6](#) in Appendix D for details of the covariates used in the RP models for this variable.

6.5.1.5.2 Prediction Step (Ever-Daily-Used Cigarettes)

After the weights were adjusted, the ever-daily-used variable was modeled using weighted logistic regression in SUDAAN. The predicted mean from this model was the predicted probability of ever smoking cigarettes daily.

6.5.1.5.3 Hot-Deck Step (Ever-Daily-Used Cigarettes)

The predicted mean from the prior step was used in a straightforward hot-deck step, which is summarized in [Tables E.23](#) and [E.24](#).

6.5.1.5.4 Hot-Deck Step (Age at First Daily Cigarette Use)

Instead of separately modeling age at first daily cigarette use, the predicted means from the age-at-first-cigarette-use models were used to determine neighborhoods. The imputation-revised age-at-first-daily-use assignment was conducted using univariate matching and univariate assignment.

All the logical constraints applied to age at first cigarette use were also applied to age at first daily cigarette use. Besides those logical constraints, an additional logical constraint was applied specifically to age at first daily cigarette use. If the cigarette AFU was not missing for a recipient with a missing age at first daily use, the donors were prevented from having an age at first daily use earlier than the preexisting cigarette AFU. This constraint was applied as daily cigarette users constitute a subset of all cigarette users. Therefore, daily use of cigarettes can be thought of as a child drug for cigarettes, with a lifetime indicator and an AFU measure but no

¹²⁰ The imputation-revised 30-day frequency included responses from the 30-day frequency question (CG07), as well as the estimated 30-day frequency question (CG07DKRE).

recency or frequency. This association required that these variables remain internally consistent for each respondent.

The likeness constraints were nearly identical to those used for cigarette AFU. There was only one difference: an additional step was employed if no donor was found after loosening all of the likeness constraints. In particular, if the AFU and age at first daily use were both initially missing, the imputed AFU was set back to missing and reimputed simultaneously with the age at first daily use so that they were mutually consistent.¹²¹

Subject to these constraints, the age-at-first-daily-use value of the randomly selected donor was then assigned to the recipient.

6.5.1.5.5 Assignment of Date of First Daily Cigarette Use

After the imputation-revised age at first daily cigarette use was created, all daily cigarette users had a valid age at first daily cigarette use. From this age, a year/month/day of first daily use was assigned. The date assignment procedure was identical to the procedure described in Section 6.5.1.4.4 with the following exception. In the setting of the *earliest possible date* for daily cigarette use, the overall cigarette date of first use was used as an additional bound. This was done for cigarettes and other substances to ensure that the child drug's date of first use was never earlier than the parent drug's date of first use.

6.5.1.5.6 Final Variables

The final imputation-revised date-of-first-daily-cigarette-use variables were named in the same manner as described in Section 6.5.1.4.5. However, the three-letter identifier for cigarette daily use was CD2. As with the general cigarette use variables, each IR variable was accompanied by an imputation indicator with a prefix II instead of IR.

6.5.1.6 "Other" Drugs Recency and Frequency

Imputation for the parent and child recency and frequency variables for imputation sets 6, 8, 10, 11, 13, 15, and 17 in [Table 6.6](#) (smokeless tobacco, cigars, pipes, alcohol, marijuana, cocaine and crack, and heroin, respectively) was done in a manner similar to that described above for cigarettes. The following deviations from the process described for cigarettes applied to these "other" drugs.

The order of imputation for smokeless tobacco and cigars was identical to that for cigarettes, with recency of use being modeled first, followed by 30-day frequency. However, not all imputation sets included the same variables. Alcohol, marijuana, cocaine and crack, and heroin also included a measure for 12-month frequency of use. For these drugs, imputation proceeded in the following order: recency of use, 12-month frequency of use, and 30-day frequency of use. For a given drug, this ordering allowed recency of use to be included in the model for 12-month frequency of use and allowed 12-month frequency of use to be included in the model for 30-day frequency. Further, this ordering allowed the provisional recency of use to

¹²¹ Though it has occurred in prior years, the situation where no donors were available, even after loosening all constraints, did not occur in the 2016 NSDUH.

define the domains for the frequency measures. For pipes, the respondents were only asked about their most recent use, and no information was collected regarding frequency of use in the past year or month. Therefore, only the recency-of-use variable required modeling and imputation.

6.5.1.6.1 Recency of Use

6.5.1.6.1.1 Response Propensity Step

The input to the weight adjustment model in the RP step for the parent and child recency-of-use variables was the preliminary analysis weight, PANALWT. Similar to cigarettes, the eligible population for the recency-of-use models included all lifetime users of the drug of interest as identified by the imputation-revised lifetime drug use variables. However, the identification of respondents and nonrespondents for the purpose of imputation differed from drug to drug depending on the information collected in the questionnaire. In general, an individual had to have provided a valid response to all variables included in the imputation set to be classified as a respondent. If the imputation set included both "parent" and "child" drugs ([Table 6.1](#)), then this requirement extended across all observed measures of drug use. For example, to be classified as a respondent for the cocaine and crack imputation set, a valid response must be provided for recency of use, 12-month frequency of use, and 30-day frequency of use for both the cocaine and crack questions. See [Tables D.7](#) through [D.27](#) in Appendix D for details of the covariates used in the RP models for these variables.

6.5.1.6.1.2 Prediction Step

Only cigarettes, cigars, and smokeless tobacco included a recency category for past 3 years but not past year. For all other drugs except pipes, the outcome variable was a three-level categorical variable with the following levels:

1. past month
2. past year, not past month
3. lifetime, not past year

For pipes, the outcome variable had only two levels:

1. past month
2. lifetime, not past month

Because cigarettes were the first recency/frequency imputation set, it was not possible to include the recency information for other drugs as covariates in the PRD model. However, for drugs other than cigarettes, recency-of-use covariates for cigarettes, alcohol, and marijuana replaced the lifetime indicators where applicable. For example, the PRD model for alcohol included recency indicators for cigarettes but only included the lifetime usage indicator for marijuana because this drug comes later in the sequence.

6.5.1.6.1.3 Provisional Hot-Deck Step

For certain cases, a general recency category was assigned during the editing process, and the specific recency was then determined during imputation. However, the categories available for both the general recency and the specific recency varied from drug to drug depending on the number of levels included in the recency-of-use measure. The allowable general and specific recency categories for cigarettes are shown in Table 6.7. The same categories apply to cigars and smokeless tobacco. For all other drugs except pipes, the only general incomplete recency categories that were applicable were lifetime and past year (the first two rows). For pipes, only the lifetime category was applicable. Logical constraints ensured that only donors with allowable specific recency categories were included in the neighborhood of potential donors.

Occasionally, more than one substance was associated with a single predicted mean, leading to a multivariate assignment of imputed values. No provisional imputed values were determined for substances that were a subset of the substance associated with the predicted mean (parent/child drugs). This occurred with cocaine (parent) and crack (child).

6.5.1.6.1.4 Hierarchical Modeling

For heroin, the proportion of users who were past year users was quite small when compared with the total number of lifetime users. The lopsided distributions¹²² for heroin caused convergence problems when fitting polytomous logistic models. To alleviate this problem, the single polytomous logistic model was replaced with two dichotomous logistic models¹²³ that were fit hierarchically.

As with the polytomous logistic model, the first dichotomous logistic model was fit among lifetime users, but the past month and past year but not past month categories in the response variable were collapsed into a single level. In a similar manner to other recency-of-use models, respondents' weights were adjusted so that they represented all lifetime users. The predicted probability of past year use given lifetime use was obtained from this model.

The second model was limited to past year users, where the response variable had two levels: past month and past year but not past month users. For the second model, respondents' weights were adjusted so that they represented all past year users. In order to do this, it was necessary to completely define the domain of past year users. Missing values were provisionally imputed to past year or not past year use by randomly allocating the response using the predicted means from the first model.

From the two dichotomous logistic models, both the probability of past month use and the probability of past year but not past month use were obtained and used in the provisional hot-deck program for recency. Once the predicted means were determined from the two models,

¹²² A "lopsided distribution" in the context of recency of use is where, among the categories of past month use, past year but not past month use, and lifetime not past year use, only a small minority of respondents gave a response of "past month use."

¹²³ The set of covariates used for these dichotomous logistic models were the same as those for logistic modeling given earlier in this section.

a single vector of predicted means conditional on lifetime usage, as with the polytomous logistic models, was determined as follows:

$$P(\text{past month use} \mid \text{lifetime use}) = P(\text{past month use} \mid \text{past year use}) \times P(\text{past year use} \mid \text{lifetime use}), \text{ and}$$

$$P(\text{past year, not past month use} \mid \text{lifetime use}) = P(\text{past year, not past month use} \mid \text{past year use}) \times P(\text{past year use} \mid \text{lifetime use}).$$

6.5.1.6.2 12-Month Frequency of Use

The modeling of 12-month frequency sequentially followed that of recency of use for alcohol, marijuana, cocaine and crack, and heroin.

6.5.1.6.2.1 Response Propensity Step

The input to the weight adjustment model in the RP step for the 12-month drug frequency-of-use variables was the preliminary analysis weight, PANALWT. The eligible population for the imputation of 12-month frequency of use was past year users of the drug in question (as defined by the provisional recency of use). The item response indicator and the RP adjustment were defined among the past year users of each drug. Item respondents were defined using the same criterion as was used in the recency-of-use imputations. Namely, the respondent had to have a valid response to all of the applicable measures for the drug of interest. The item RP adjustment was then computed so that the respondents' weights accurately represented all past year users of the drug. See Appendix D for details of the covariates used in the RP models for these variables.

6.5.1.6.2.2 Prediction Step

As indicated in the previous section, only past year users of the drug of interest were used to build the 12-month frequency-of-use model. The response variable of interest in the 12-month frequency-of-use models for most respondents, prior to a normalizing transformation, was the proportion of the days in a full year (365.25) on which a respondent used a particular drug. For example, if a respondent entered a 12-month frequency of 100, the (untransformed) response variable of interest would be 100/365.25. Some respondents, however, started using the drug within the past year. If they responded to the MFU question, the difference between the MFU and the date of the interview indicated the total time period during which they could have been using drugs.¹²⁴ If the date of the interview was July 10, for example, and the MFU was March of the same year, the maximum period during which the respondent could have used is the number of days between March 1 and July 10 (inclusive), or 101. Thus, if a respondent entered a 12-month frequency of 100, the (untransformed) response variable of interest would be 100/101 instead of 100/365.25. The range of values for the proportion was from (greater than) 0 to 1.

¹²⁴ If a respondent initiated use in the past year (according to his or her age-at-first-use or year-of-first-use response), but did not answer the month-at-first-use question, then the maximum period the respondent could have been using drugs was the number of days between the interview date and the date of the respondent's last birthday or January 1 of the survey year (if the respondent reported first using in the survey year), whichever yielded the smallest number of days.

Hence, in order to model 12-month frequency of use, the following empirical logit transformation was computed for all respondents:

$$\log\left[(Y_i + 0.5)/(N_i - Y_i + 0.5)\right],$$

where Y_i is the observed 12-month frequency for respondent i and N_i is the total number of days in the year that respondent i could have used the substance. This transformation is nearly equivalent to the standard logit transformation:

$$Y_i^* = \log\left[P_i/(1 - P_i)\right],$$

where P_i is defined as the proportion of days in the past year in which respondent i used the drug. The standard logit transformation was not used because it was not defined for daily users.¹²⁵ Using the adjusted weights, a linear univariate regression model using SUDAAN software was then fitted for the log-transformed variable Y_i within each age group.

- **Covariates:** Because the 12-month frequency models were limited to past year users, only two recency categories could have resulted: past month use and past year but not past month use.¹²⁶ Hence, recency of use for the drug being modeled was represented as a covariate in the 12-month frequency-of-use model by a single indicator variable representing these two categories. Imputation-revised recency of use for other drugs was used if available. If the missing values for a given drug's recency of use had not yet been imputed, a single covariate was used that indicated lifetime usage of that drug. To control for state variations in drug use, the state-rank groups defined for the recency-of-use imputations were included as covariates in the 12-month frequency-of-use models.¹²⁷ Appendix D provides a complete summary of the 12-month frequency-of-use models.
- **Predicted Means:** The predicted mean that resulted from the 12-month frequency-of-use model was a logit of the proportion of the year used. This logit was back-transformed into a proportion for use as the variable from which the neighborhoods were created. This proportion could be treated as a probability, which, in turn, could be multiplied by the probability of past year use to make the predicted mean conditional on lifetime use of the drug in question. When calculating predicted means for some item nonrespondents, sometimes it was not known whether they were past year users. Hence, to make the predicted means conditional on the same recency of use, the variables were transformed to make them conditional on what was known.

¹²⁵ If the respondent was a daily user of the substance, then $\log[(Y + 0.5)/(N - Y + 0.5)] \approx \log[(N + 0.5)/0.5]$ with $N = 365.25$ so that it was defined for all respondents. See Cox and Snell (1989) for a discussion of the empirical logit transformation.

¹²⁶ For item nonrespondents, where parameter estimates were used to determine predicted means, past year use was defined based on a provisional imputation.

¹²⁷ As with the recency-of-use models, for a few cases, the state-rank variable could not be included in the model. Usually, but not always, the age group/drug combination that had problems was the same for recency of use and 12-month frequency of use.

6.5.1.6.2.3 *Provisional Hot-Deck Step*

For imputation sets that included both 12-month frequency and 30-day frequency—alcohol (Set 11), marijuana (13), cocaine and crack (15), and heroin (17)—it was necessary to provisionally impute the 12-month frequency-of-use variable so that it could be used as a covariate in the 30-day frequency-of-use imputations.

The logical constraints involved the interview date, initiation (AFU, MFU, and YFU), birthday, recency of use, and 30-day frequency of use. The likeness constraints used in the assignment of values for 12-month frequency of use were similar to those used for recency of use. State-rank groups were again based on level of past month usage. Recipients and donors were also required to have the same recency of use (past month vs. past year but not past month), whether that recency of use was reported or imputed.¹²⁸ If no donors were available within these constraints, then they were loosened in the following order: (1) the delta constraint was removed, (2) donors and recipients were no longer required to be from states with similar usage levels, and (3) donors and recipients were no longer required to have the same recency of use.

6.5.1.6.2.4 *Assignment of Provisional Imputed Values*

For all drug use measures except 12-month frequency, the observed value of interest was donated directly to the recipient. However, because donors and recipients could potentially have had a different maximum possible number of days in the year that they could have used a substance, the observed proportion of the total period was donated rather than the observed 12-month frequency. In the assignment step, the donor's proportion of the total period was multiplied by the recipient's maximum possible number of days in the year on which he or she could have used the substance in order to arrive at a 12-month frequency-of-use value for the recipient.

Occasionally, more than one substance was associated with a single predicted mean. However, for the provisionally imputed values, only the parent drug was used as a covariate in later models. Therefore, multivariate assignments were not needed in the provisional hot-deck step, but they did occur in the final hot-deck step for recency and frequency. For example, the recency and frequency variables for cocaine and crack formed a single imputation set (15). Although 12-month frequency questions were asked for both cocaine and crack, only the 12-month frequency for cocaine was modeled, and only the 12-month frequency for cocaine was used as a covariate in the subsequent PRD model (30-day cocaine frequency). This means that there was no need to impute a provisional value for 12-month frequency for crack.

¹²⁸ Because all respondents in the 12-month frequency-of-use imputation were past year users by definition, item nonrespondents who were past month users required donors who were past month users, and item nonrespondents who were past year but not past month users required donors who matched that specific recency category.

6.5.1.6.3 30-Day Frequency of Use

6.5.1.6.3.1 Response Propensity Step

The input to the weight adjustment model in the RP step for the 30-day drug frequency-of-use variables was the preliminary analysis weight, PANALWT. As with cigarettes, the file was first reduced to the domain of past month users, as defined by the provisional recency variable. Next, item respondents and nonrespondents were defined according to the same criterion used for the recency and 12-month frequency imputations. To be an item respondent, the individual had to have provided valid responses to all applicable measures for the drug of interest. The item RP adjustment was then computed so that the respondents' weights accurately represented all past month users of the drug. In contrast with the RP model for cigarettes, the provisional 12-month frequency was included as a covariate for those drug sections that asked the respondent to report this measure (Table 6.6). In addition, recencies of use for cigarettes, smokeless tobacco, cigars, pipes, alcohol, marijuana, cocaine, crack, and heroin were included if available.¹²⁹ See Appendix D for details of the covariates used in the RP model for these variables.

6.5.1.6.3.2 Prediction Step

As described for cigarettes, only past month users of the drug of interest were used to build the 30-day frequency-of-use model. The response variable of interest in the 30-day frequency-of-use models, prior to a normalizing transformation, was the proportion of the days in a month (30 days) on which a respondent used the drug of interest. In contrast with cigarettes, all other drugs that included a measure for 30-day frequency used a single model for all past month users rather than two separate models for daily use and nondaily use.

6.5.1.6.3.3 Provisional Hot-Deck Step

Because there were no substances within the multivariate set of recency-of-use and frequency-of-use measures lower in the sequence than 30-day frequency, there was no need to provisionally impute 30-day frequencies.

6.5.1.6.4 Final Hot-Deck Step

The same principles that applied to the final hot-deck step for cigarettes applied to other drugs. However, for substances with child drugs and substances with both 12-month and 30-day frequencies, the logical constraints were considerably more complicated, the predictive mean vectors were larger, and the number of missingness patterns was greater. Appendix E provides detailed information on these hot-deck steps, and Table 6.9 provides a listing of the full predictive mean vector as applied to all final hot-deck programs for recency and frequency.

¹²⁹ If the recency of use for a particular drug was not yet defined, the lifetime indication of use was used instead. The recency of use of the drug being modeled was not used, because all respondents in the model were past month users.

Table 6.9 Elements of Full Predictive Mean Vector

Drug Use Measure and Category of Interest	Predicted Mean	Substance
Recency of Use, Past Month Use ¹	$P(\text{past month user} \mid \text{lifetime user})$	All substances
Recency of Use, Past Year but Not Past Month Use ¹	$P(\text{past year but not past month user} \mid \text{lifetime user})$	All substances except pipes
Recency of Use, Past 3 Years but Not Past Year Use ¹	$P(\text{past 3 years but not past year user} \mid \text{lifetime user})$	Tobacco products ² only
12-Month Frequency of Use	$P(\text{use on a given day in the year} \mid \text{past year user}) \times P(\text{past year user} \mid \text{lifetime user})^3$	All substances except tobacco
30-Day Frequency of Use for Alcohol and Substances with Few Daily Users ⁴	$P(\text{use on a given day in the month} \mid \text{past month user}) \times P(\text{past month user} \mid \text{lifetime user})^5$	All substances except cigarettes and pipes
30-Day Frequency of Use for Substances with Many Daily Users (excluding Alcohol and Smokeless Tobacco)	$P(\text{use on a given day in the month} \mid \text{past month user, not a daily user}) \times P(\text{not a daily user} \mid \text{lifetime user}) \times P(\text{past month user} \mid \text{lifetime user})^5$	Cigarettes
Daily User over Past 30 Days	$P(\text{daily user} \mid \text{past month user}) \times P(\text{past month user} \mid \text{lifetime user})^5$	Cigarettes

¹ The final category for recency (lifetime but not past year or lifetime but not past 3 years) was not needed in the predictive mean vector, because the multinomial probabilities summed to 1, and this probability was determined by the other probabilities.

² "Tobacco products" includes cigarettes, cigars, and smokeless tobacco.

³ Interpreting the proportion of the year used as a probability of use on a given day in the year assumed that the probability of use on each day in the year was equal. However, this was not true. The violation of this assumption did not seriously affect the ability to find a reasonable variable to use for finding a neighborhood, and it did allow the predicted mean to be made conditional on what was known.

⁴ Alcohol, with many daily users, was included in this group because the distribution did not show a severe drop-off from 30 days a month to 29 days a month, as was apparent with cigarettes.

⁵ Interpreting the proportion of the month used as a probability of use on a given day in the month assumed that the probability of use on each day in the month was equal, which was not true, in the same manner as the 12-month frequency of use (see note #3 for this table).

The construction of the predictive mean vectors for certain drugs was often complex. The main reason for the complexity is that recency and frequency models were not fit for all child drugs. In fact, the predicted means from the models for the parent drug were often used as surrogates for the child drug predicted means to reduce the number of models that needed to be fit and to avoid convergence problems with small sample sizes for some of the rarer child drugs. For example, if the individual requiring imputation is a past year user of cocaine but he or she has a missing crack recency, then the predictive mean vector includes the probability of past month cocaine use, given that the individual is a past year user of cocaine. When constructing the predictive mean vectors, the following general principles were followed:

1. If both the parent drug recency and the child drug recency were missing, condition on the general recency category of the parent drug.
2. Condition all elements of the predictive mean vector on the same general recency level.

6.5.1.6.5 Final Variables

Similar to the final imputation-revised recency-of-use and 30-day frequency variables for cigarettes (IRCIGRC and IRCIGFM), the final imputation-revised recency and frequency variables for other drugs with the exception of smokeless tobacco were identified with the prefix IR, followed by a five-letter identifier, where a three-letter code identified the drug and the final two letters identified the measure. In addition to the RC and FM identifiers used for cigarettes, the identifier FY was used for the 12-month frequency variable. For smokeless tobacco, the final imputation-revised recency and frequency variables are IRSMKLSSREC and IRSMKLSS30N. Again, each IR variable was accompanied by an imputation indicator with the requisite II prefix.

6.5.1.7 "Other" Drugs Age at First Use

The AFU variables for imputation sets 7, 9, 12, 14, 16, and 18 in [Table 6.6](#) (smokeless tobacco, cigars, alcohol, marijuana, cocaine and crack, and heroin, respectively) were imputed in a manner similar to that described for cigarettes. However, some deviations from the process described for cigarettes applied to the AFU for these "other" drugs as described below.

6.5.1.7.1 Response Propensity Step

The input to the weight adjustment model in the RP step for the AFU variables for other drugs was the preliminary analysis weight, PANALWT. The RP step for AFU for other drugs was very similar to the RP step for cigarette AFU (Section 6.5.1.4.1). For substances that included child drugs, a response had to be provided for each drug for an individual to be considered a respondent for imputation purposes. Appendix D provides a complete list of covariates used in each model to properly adjust the weights.

6.5.1.7.2 Prediction Step

The PRD step for AFU for other drugs was also very similar to the analogous step for cigarettes (Section 6.5.1.4.2). For substances with child drugs, only the parent drug was modeled. Modified versions of the 12-month frequency of use (where applicable) and AFU of previously imputed drugs were used as covariates and were defined as follows:

new12_i	= 0	if respondent <i>i</i> did not use the drug of interest in the past 12 months
	= 12-month frequency	if respondent <i>i</i> used the drug of interest in the past 12 months
AFU_i	= 0	if respondent <i>i</i> is not a lifetime drug user of the drug of interest
	= age at first use	if respondent <i>i</i> is a lifetime drug user of the drug of interest

6.5.1.7.3 Hot-Deck Step

For cocaine (crack), more than one AFU variable was associated with a single predicted AFU. This led to a multivariate assignment of the imputed values. Respondents were asked for their AFU of cocaine and were asked for their AFU of crack. This often led to complex and numerous logical constraints. These constraints used not only cocaine and crack ages at first use, when available, but also imputation-revised recencies and frequencies.

6.5.1.7.4 Month-of-First-Use, Year-of-First-Use, and Day-of-First-Use Assignments

The general principles described in Section 6.5.1.4.4 applied to the remaining drugs with the following exceptions.

- For crack, the corresponding parent drug's date of first use was assigned first. Then, in the setting of the *earliest possible date* for the child drug, the parent drug's date of first use was used as an additional bound. This was done to ensure that the child drug's date of first use was never earlier than the parent drug's date of first use.
- For cocaine, crack recency and frequency information was used to bound the *latest possible date* for the date of first use. For example, respondents with crack recency = 3 (i.e., lifetime but not past year user of crack) could not have first used cocaine within the past year, regardless of the cocaine recency value. The bound created using crack recency and frequency was calculated in exactly the same way as for cocaine recency and frequency information.

6.5.1.7.5 Final Age and Date-of-First-Use Variables

The final imputation-revised date-of-first-use variables for "other" variables, with the exception of smokeless tobacco, were named in the same manner as those for cigarettes: with the prefix IR, followed by a three-letter code identifying the drug and the final three letters identifying the measure (AGE = age at first use, MFU = month of first use, YFU = year of first use, DFU = day of first use). For smokeless tobacco, the six-letter code, SMKLSS, was used to identify the drug rather than a three-letter code. Additionally, TRY was used to identify the age-at-first-use measure. Again, each IR variable was accompanied by an imputation indicator with the requisite II prefix.

6.5.1.8 30-Day Binge Drinking Frequency (Imputation Set 19)

For alcohol, an additional variable was defined that measured level of usage in the past month. In particular, the variable ALCBNG30D measured the binge drinking frequency or the number of days in the past month on which the respondent had five or more drinks for males or four or more drinks for females on the same occasion. Since 30-day binge drinking is defined differently for males and females, the model building and final assignment of imputed values was performed separately within gender (male and female) and age group (12 to 17, 18 to 25, and 26 or older).

6.5.1.8.1 Response Propensity Step

The input to the weight adjustment model in the RP step for the 30-day binge drinking frequency variable was the preliminary analysis weight, PANALWT. The RP model was built using all past month users of alcohol whether they were binge drinkers or not. Item respondents for 30-day binge drinking were defined as those past month alcohol users with a nonmissing value for ALCBNG30D. See [Tables D.16](#), [D.17](#), and [D.18](#) in Appendix D for details of the covariates used in the RP model for this variable.

6.5.1.8.2 Prediction Step

The response variable of interest in the 30-day binge drinking frequency model, prior to a normalizing transformation, was the proportion of the days in a month (30) on which a respondent had five or more drinks for males or four or more drinks for females on the same occasion. The range of values for the proportion was from 0 to 1, inclusive. Hence, to model 30-day binge drinking frequency of use, the following empirical logit transformation was computed for all respondents:

$$\log \left[(Y_i + 0.5) / (N - Y_i + 0.5) \right],$$

where Y_i was the observed 30-day binge drinking frequency for respondent i and N was 30, the total number of days in the month that the respondent could have binge drunk. This transformation was nearly equivalent to the standard logit transformation:

$$Y_i^* = \log \left[P_i / (1 - P_i) \right],$$

where P_i was defined as the proportion of days in the past month during which respondent i had five or more drinks for males or four or more drinks for females. The standard logit transformation was not used, because it was not defined for daily binge drinkers nor was it defined for nonbinge drinkers among past month users.¹³⁰ Using the adjusted weights, a linear univariate regression model was then fitted for the log-transformed variable Y_i within each gender and age group combination.

6.5.1.8.3 Hot-Deck Step

The hot-deck step for 30-day binge drinking frequency was straightforward and implemented separately for each of the age group and gender combinations. Two logical constraints were utilized. One required the donor's 30-day binge drinking frequency to be less than or equal to the recipient's 30-day alcohol use frequency. The other constraint utilized the recipient's response to the question on the usual number of drinks of alcoholic beverages that the respondent consumed on days when he or she used alcohol in the past 30 days. This constraint

¹³⁰ If the respondent was a daily binge drinker of alcohol, then $\log[(Y + 0.5)/(N - Y + 0.5)] \approx \log[(N + 0.5)/0.5]$, where Y was the observed 30-day binge drinking frequency and N was the total number of days that the respondent could have used (usually 30). If the proportion was zero, then $\log[(Y + 0.5)/(N - Y + 0.5)] \approx \log[0.5/(N + 0.5)]$. See Cox and Snell (1989) for a discussion of the empirical logit transformation.

required the donor's 30-day binge drinking frequency to be strictly greater than zero if the recipient reported usually consuming five or more drinks for males or four or more drinks for females on the days he or she consumed alcohol during the past 30 days. The state rank and delta likeness constraints were also applied. Details on the hot-deck step are available in [Tables E.34](#), [E.35](#), and [E.37](#) in Appendix E.

6.5.2 Variables Imputed Using the modPMN Methodology

Starting in 2015, a new imputation methodology was implemented for a subset of the substance use variables covering inhalants, hallucinogens (excluding hallucinogen child lifetime variables), methamphetamine, and prescription drugs, known as modPMN. For details on modPMN methodology, see Sections 3.5 through 3.7. This change in methodology coincided with the questionnaire redesign of these drug sections. With the exception of the hallucinogen child drugs, LSD and PCP, all variables in this section are not comparable with similar variables from years prior to 2014.

Section 6.5.2.1 describes how these variables were segregated into imputation sets. Section 6.5.2.2 describes how the variables were ordered within the imputation sets. Section 6.5.2.3 gives details on the steps of modPMN for categorical variables while Section 6.5.2.4 gives details on the steps of modPMN for continuous variables. Section 6.5.2.6.1 gives details on the imputation of date at first use of inhalants, hallucinogens, and methamphetamine and Section 6.5.2.6.2 gives details on the age and date at first misuse of prescription drugs.

6.5.2.1 Imputation Sets for modPMN Drug Variables

ModPMN drug variables are segregated into imputation sets by measure rather than by drug as is done under PMN methodology. Generally, all drugs for which a measure exists are included in a single imputation set with the following exceptions. Child drugs (e.g., LSD, PCP, ECS) and specific prescription drugs (e.g., OxyContin[®]) are imputed in separate sets following the imputation of the parent drug (e.g., hallucinogens) or the more general drug category (e.g., pain relievers in the case of OxyContin[®]) so that the domain for these measures can be fully defined. In addition, initiation data for the prescription drugs are imputed separately from inhalants, hallucinogens, and methamphetamine due to the difference in domain and imputation methodology (Section 6.5.2.6). [Table 6.10](#) shows the 19 imputation sets covered in this section. Note that imputation sets 1 through 20 are covered in Section 6.5.1. Sets 31 through 34 and 36 through 38 can be run in any order (or in parallel); however, sets 31 through 34 cannot be run until after set 30 is completed, and sets 36 through 38 cannot be run until after set 35 is completed.

Table 6.10 Drugs and Drug Use Measures, modPMN Imputation Sets

Drug	Drug Use Measures										
	Lifetime Usage	Recency of Any Use	Recency of Use/ Misuse	12-Month Frequency of Use	30-Day Frequency of Use/ Misuse	Past Year Initiation	Age at First Use/ Misuse	Date of First Use/ Misuse			
Inhalants	Set 21	N/A	Set 24	Set 27	Set 28	N/A	Set 30	Set 36			
Hallucinogens								Set 37			
Methamphetamine								Set 38			
Stimulants	N/A	Set 22	Set 24	N/A	Set 28	Set 29	Set 31				
Pain Relievers							Set 32				
Tranquilizers							Set 33				
Sedatives							Set 34				
OxyContin®		Set 23 ¹	Set 25 ¹		N/A	Set 28	N/A	Set 29	N/A	N/A	
LSD, PCP, Ecstasy ²	DM	N/A	Set 26	N/A					N/A	Set 35	Set 39
Ketamine, DMT/AMT/Foxy, <i>Salvia divinorum</i> ²										N/A	N/A

DM = different method; modPMN = modified predictive mean neighborhood; N/A = not applicable; PMN = predictive mean neighborhood.

¹ Recency of Any Use and Recency of Misuse for OxyContin® include two levels because only OxyContin® use in the past 12 months is asked in the survey. Recency of Any Use includes levels for past year any use and did not use in the past year. Recency of Misuse includes levels for past year misuse and did not misuse in the past year.

² Hallucinogen child drugs recency/frequency and age at first use are imputed using modPMN. Lifetime use for hallucinogen child drugs are imputed using PMN methodology. See [Table 6.6](#) for the PMN imputation sets.

6.5.2.2 Imputation Order within Imputation Sets

As described in Section 3.7.1, the sequencing of variables in modPMN within an imputation set is by descending item response rates. For sets that included both prescription drug and nonprescription drug measures (i.e., imputation sets 24 and 28), inhalants, hallucinogens, and methamphetamine preceded all prescription drugs. The ordering for 2016 is shown in [Table 6.11](#).

Table 6.11 Sequencing of Drugs within Imputation Sets, 2016

Imputation Set	Order
Lifetime Drug Use	Methamphetamine, Inhalants, Hallucinogens
Recency of Any Use	Stimulants, Sedatives, Tranquilizers, Pain Relievers
Recency of Use/Misuse	Methamphetamine, Hallucinogens, Inhalants, Pain Relievers, Stimulants, Sedatives, Tranquilizers
Hallucinogen Child Drug Recency of Use	<i>Salvia divinorum</i> , LSD, Ketamine, Ecstasy, PCP, DMT/AMT/Foxy
12-Month Frequency of Use	Methamphetamine, Hallucinogens, Inhalants
30-Day Frequency of Use	Methamphetamine, Hallucinogens, Inhalants, Stimulants, Sedatives, Tranquilizers, Pain Relievers
Age at First Use	Methamphetamine, Hallucinogens, Inhalants
Hallucinogen Child Drugs Age at First Use	Ecstasy, LSD, PCP
Past Year Initiation	Sedatives, Tranquilizers, Stimulants, Pain Relievers

6.5.2.3 Categorical Variables

This section provides details on the three steps of modPMN for categorical variables as described in Sections 3.6.1 through 3.6.3. Due to the number of imputation sets, set 24 is explained in detail in Sections 6.5.2.3.1 through 6.5.2.3.4, and the rest of the categorical variable sets are described in Section 6.5.2.3.5 as deviations from these procedures.

Categorical variables include lifetime usage (set 21), recency of any use (sets 22 and 23), recency of (mis)use (sets 24 through 26), and past year initiation (set 29).

6.5.2.3.1 *Recency of (Mis)Use, Set 24: Response Propensity Adjustment*

This section provides specific information about the RP steps for the recency of (mis)use variables (imputation set 24). Section 3.6.1 provides general information about the RP adjustment step of modPMN. One RP model was fit for each drug in the set for each cycle for a total of 14 RP models for this set. The four key inputs to the RP adjustment step as laid out in Section 3.3.1 are described below.

- **Analysis Weight:** The input to all RP models was the preliminary analysis weight, PANALWT.
- **Domain:** For inhalants, hallucinogens, and methamphetamine, the domain includes lifetime users of the drug based on the imputation-revised lifetime use indicator. For prescription drugs, the domain includes lifetime users of the drugs (IRXXXANYREC=1, 2) based on the imputation-revised any use recency variable. See [Table 6.12](#) for domains for all other sets.
- **Item Response Indicator:** If a valid response was provided for the recency of (mis)use measure for the drug, the individual was deemed an item respondent.
- **Covariates:** [Table D.38](#) in Appendix D lists both the starting and the final covariates used in each of the 14 RP models used. The table includes the covariate lists for both cycles.

6.5.2.3.2 *Recency of (Mis)Use, Set 24: Prediction Modeling*

This section provides specific information about the PRD modeling step for the recency of (mis)use set 24. Section 3.6.2 provides general information about the PRD modeling step of modPMN. Each of the 14 RP models segregated the domain members (including both respondents and nonrespondents) into five bins of roughly equal size, and separate PRD models were fit within each of the bins. This resulted in 70 total PRD models for imputation set 24 (7 drugs $\times$ 5 bins $\times$ 2 cycles).

For each of the 70 models, the probability of selecting each recency of (mis)use category was modeled within each bin using polytomous logistic regression. SUDAAN'S MULTLOG procedure was used to estimate the parameters from the appropriate logistic model from which predicted probabilities for each of the recency categories were calculated for both item respondents and item nonrespondents. For inhalants, hallucinogens, and methamphetamine, the three recency categories were the following:

- past month use;
- past year, not past month use; and
- lifetime, not past year use.

For prescription drugs, the four recency categories were the following:

- past month misuse;
- past year, not past month misuse;
- lifetime, not past year misuse; and
- never misused.

These predicted probabilities were used in the subsequent step to impute a recency category of (mis)use. [Table D.39](#) in Appendix D lists both the starting and final covariates for each of the 70 PRD models.

6.5.2.3.3 Recency of (Mis)Use, Set 24: Collocated Stochastic Imputation

This section provides specific information about the CSI step for set 24. Section 3.6.3 provides general information about the CSI step. Each of the 70 PRD models is followed by an imputation step where the predicted probabilities from the previous step, along with a random number, is used to impute a recency category of (mis)use for the given drug for each item nonrespondent.

In the case of partial information, the predictive mean vector is adjusted to prevent inconsistent imputation. For example, if a respondent's edited recency of inhalant use is past year where past month use versus past year not past month use is unknown, the predicted probability associated with lifetime not past year use, $P(\text{Lifetime but not past year use of inhalants} \mid \text{lifetime use of inhalants})$, is set to 0. The remaining probabilities corresponding to past month use and past year not past month use are then rescaled to sum to 1 prior to imputation. For more information on the restrictions based on lifetime use (for inhalants, hallucinogens, and methamphetamine) and recency of any use (for prescription drugs), see [Tables E.67](#) through [E.69](#) and [E.76](#) through [E.79](#).

6.5.2.3.4 Recency of (Mis)Use, Set 24: Final Variables

The final imputation-revised recency-of-use variables for hallucinogens (HALLUC), inhalants (INHAL), and methamphetamine (METHAM) were identified with the prefix IR, followed by the drug identifier and a final "REC" to identify the measure. For the prescription drugs, the final imputation-revised recency of misuse variables were identified with the prefix IR, followed by an eight-letter identifier, where a three-letter code identified the drug (PNR, SED, STM, TRQ) and the final five letters identified the measure (NMREC). Each IR variable was accompanied by an imputation indicator with a prefix II instead of IR.

6.5.2.3.5 "Other" Categorical ModPMN Variables

Imputation for sets 21, 22, 23, 25, 26, and 29 in [Table 6.10](#) was done in a manner similar to that described above for set 24. The following deviations from the process described for set 24 applied to these "other" categorical modPMN variables.

6.5.2.3.5.1 Response Propensity Step

The overall process for the RP step was similar to the process described for set 24; however, the number of models fit and the inputs to the RP step varied from set to set. The total number of RP models fit for each set can be found in [Table 6.12](#). For each of the four key inputs to the RP adjustment step, details of any deviations are below.

- **Analysis Weight:** All sets used the preliminary analysis weight, PANALWT.
- **Domain:** The domain for each set is detailed in [Table 6.12](#).
- **Item Response Indicator:** If a valid response was provided for the given measure for the drug, the individual was deemed an item respondent.
- **Covariates:** [Table 6.12](#) lists the tables in Appendix D that show both the starting and final covariates used in each of the RP models.

Table 6.12 Summary of Response Propensity Models for Categorical modPMN Variables

Set	Number of Response Propensity Models	Domain	Table Reporting Covariate Lists
21: Lifetime Usage	6	All respondents	D.28
22: Recency of Any Use	8	All respondents	D.34
23: Recency of Any Use, OxyContin®	1	Past year users of pain relievers, based on IRPNRANYREC	D.36
24: Recency of (Mis)Use	14	Nonprescription drugs: Lifetime users, based on imputation-revised lifetime use Prescription drugs: Lifetime users, based on IRXXXANYREC	D.38
25: Recency of Misuse, OxyContin®	1	Past year misusers of pain relievers, based on IRPNRNMREC and past year users of OxyContin® based on IROXCNANYR	D.40
26: Recency of Use, Hallucinogen Child Drugs	12	Lifetime user of hallucinogens, based on IRHALLUCREC	D.42
29: Past Year Initiation	8	Past year misuser, based on IRXXXNMREC	D.52

modPMN = modified predictive mean neighborhood.

6.5.2.3.5.2 Prediction Modeling Step

Section 6.5.2.3.2 described the SUDAAN method used to impute a multilevel variable. Other sets that model multilevel variables include sets 22 and 26. The other categorical variable modPMN sets model dichotomous variables. The PRD model for these sets utilized dichotomous logistic regression through SUDAAN's LOGISTIC procedure (called RLOGIST in SAS-callable SUDAAN). The levels for outcome variables for each set are listed in [Table 6.13](#). [Tables D.29](#), [D.35](#), [D.37](#), [D.39](#), [D.41](#), [D.43](#), and [D.53](#) in Appendix D list both the starting and final covariates for the PRD models.

Table 6.13 Outcomes of Prediction Modeling

Set	Outcome Variable Levels
21: Lifetime Usage	1: Lifetime Use 2: Never Used
22: Recency of Any Use	1: Past Year Use 2: Lifetime, Not Past Year Use 3: Never Used
23: Recency of Any Use, OxyContin [®]	1: Past Year Use 2: No Past Year Use
24: Recency of (Mis)Use	Nonprescription Drugs 1: Past Month Use 2: Past Year, Not Past Month Use 3: Lifetime, Not Past Year Use Prescription Drugs 1: Past Month Misuse 2: Past Year, Not Past Month Misuse 3: Lifetime, Not Past Year Misuse 4: Never Misused
25: Recency of Misuse, OxyContin [®]	1: Past Year Misuse 2: No Past Year Misuse
26: Recency of Use, Hallucinogen Child Drugs	1: Past Month Use 2: Past Year, Not Past Month Use 3: Lifetime, Not Past Year Use
29: Past Year Initiation	1: Past Year Misuse Initiate 2: Not a Past Year Misuse Initiate

6.5.2.3.5.3 CSI Step

There are no deviations from the CSI procedure described in Section 6.5.2.3.3. Partial information is not available for the edited variables in sets 21 and 23; therefore, no adjustments to the predictive mean vectors are made in these sets. In addition to partial information from edited variables, in some sets, previously imputed drug values are also used to adjust the predictive mean vectors. Additionally, in set 26, edited hallucinogen age at first use is used to adjust the hallucinogen child drug recency-of-use predictive mean vectors to ensure consistency between parent and child recency and AFU. For more information on these restrictions, see [Tables E.55](#) through [E.57](#), [E.62](#) through [E.66](#), [E.70](#) through [E.75](#), [E.80](#), and [E.95](#) through [E.98](#).

6.5.2.3.5.4 Final Variables

No final imputation-revised variables indicating lifetime usage (set 21) alone were created because this information was recoded in the final imputation-revised recency of (mis)use variables. Imputation indicators also were not created.

The final imputation-revised variables for the other sets followed a similar pattern as those for set 24: the IR prefix followed by a drug abbreviation and ending with an indicator of the specific measure. Table 6.14 provides the final imputation-revised variable format for each set. Each IR variable was accompanied by an imputation indicator with a prefix II instead of IR.

Table 6.14 Final Imputation-Revised Categorical modPMN Variables

Set	IR Variables ¹
21: Lifetime Usage	N/A
22: Recency of Any Use	IRXXXANYREC
23: Recency of Any Use, OxyContin®	IROXCNANYR
24: Recency of (Mis)Use	Nonprescription Drugs: IRXXXREC Prescription Drugs: IRXXXNMREC
25: Recency of Misuse, OxyContin®	IROXCNNMYR
26: Recency of Use, Hallucinogen Child Drugs	LSD and PCP: IRXXXRC Others: IRXXXREC
29: Past Year Initiation	IRXXXNMINIT

modPMN = modified predictive mean neighborhood; N/A = not applicable.

¹XXX represents the drug abbreviation. In some cases, it may be longer than three characters (e.g., HALLUC for hallucinogens).

6.5.2.4 Continuous Variables

This section provides details on the three steps of modPMN for continuous variables as described in Sections 3.6.1 through 3.6.3. Set 27 is explained in detail in Sections 6.5.2.4.1 through 6.5.2.4.4, and the rest of the continuous variable sets are described in Section 6.5.2.4.5 as deviations from these procedures.

Continuous variables include 12-month frequency of use (set 27), 30-day frequency of (mis)use (set 28), and age at first use (sets 30 and 35).

6.5.2.4.1 12-Month Frequency of Use, Set 27: Response Propensity Adjustment

This section provides specific information about the RP steps for the 12-month frequency-of-use set 27. Section 3.6.1 provides general information about the RP adjustment step of modPMN. One RP model was fit for each drug in the set for each cycle for a total of six RP models for this set. The four key inputs to the RP adjustment step as laid out in Section 3.3.1 are described below.

- **Analysis Weight:** The input to all RP models was the preliminary analysis weight, PANALWT.

- **Domain:** The domain includes past year users of the drug based on the imputation-revised recency of use variable, IRXXXREC.
- **Item Response Indicator:** If a valid response was provided for the 12-month frequency-of-use measure for the drug, the individual was deemed an item respondent.
- **Covariates:** [Table D.44](#) in Appendix D lists both the starting and the final covariates used in each of the six RP models used. The table includes the covariate lists for both cycles.

6.5.2.4.2 12-Month Frequency of Use, Set 27: Prediction Modeling

This section provides specific information about the PRD modeling step for the 12-month frequency-of-use set 27. Section 3.6.2 provides general information about the PRD modeling step of modPMN. Each of the six RP models segregated the domain members (including both respondents and nonrespondents) into five bins of roughly equal size, and separate PRD models were fit within each of the bins. This resulted in 30 total PRD models for imputation set 27 (3 drugs $\times$ 5 bins $\times$ 2 cycles).

The response variable of interest in the 12-month frequency-of-use models for most respondents, prior to a normalizing transformation, was the proportion of the days in a full year (365.25) on which a respondent used a particular drug. Some respondents, however, started using the drug within the past year. In these cases, this additional information was used to reduce the number of days the respondent could have used the drug in the past year. The range of values of this proportion of days used is from (greater than) 0 to 1. Therefore, in order to model 12-month frequency of use, an empirical logit transformation was computed for all respondents. Then, a linear univariate regression model using SUDAAN software was fit. For details on the transformation and reasons for use, see Section 6.5.1.6.2.2.

The 30 PRD models produced predicted means for both respondents and nonrespondents that were a logit of the proportion of the year used. These predicted means were back-transformed into a proportion for use as the variables from which the neighborhoods were created. [Table D.45](#) in Appendix D lists both the starting and final covariates used in each of the 30 PRD models.

6.5.2.4.3 12-Month Frequency of Use, Set 27: Hot-Deck

This section provides specific information about the hot-deck imputation step for set 27. As discussed in Section 3.6.3, the hot-deck step for modPMN always employs univariate matching and univariate assignment.

The logical constraints involved the interview date, initiation data (AFU, MFU, and YFU), birthday, imputation-revised recency of use, and 30-day frequency of use. In addition to the delta constraint, the likeness constraints required recipients and donors to have the same recency of use and be in the same state-rank group based on level of past month usage. Additionally, recipients and donors were required to be in the same RP bin. If no donors were available within these constraints, then they were loosened in the following order: (1) the delta

constraint was removed, (2) donors and recipients were no longer required to be from states with similar usage levels, and (3) donors and recipients were no longer required to have the same recency of use. If donors were still not available, the three constraints listed above were reapplied and the constraint requiring donors and recipients to be in the same RP bin was dropped. If no donors could be found after dropping the bin constraint, the likeness constraints that had been reapplied were once again loosened in the same order as described previously until a nonempty neighborhood of donors was obtained. A complete list of likeness and logical constraints applied to the 12-month frequency-of-use imputations can be found in [Tables E.81](#) and [E.82](#).

6.5.2.4.4 12-Month Frequency of Use, Set 27: Final Variables

The final imputation-revised 12-month frequency-of-use variables for hallucinogens (HALLUC), inhalants (INHAL), and methamphetamine (METHAM) were identified with the prefix IR, followed by the drug identifier and a final "YFQ" to identify the measure. Each IR variable was accompanied by an imputation indicator with a prefix II instead of IR.

6.5.2.4.5 "Other" Continuous ModPMN Variables

Imputation for sets 28, 30, and 35 in [Table 6.10](#) were completed in a manner similar to that described above for set 27. The following deviations applied to these "other" continuous modPMN variables. For set 28, a preliminary setup program was implemented starting in 2016, which is described in Section 6.5.2.4.5.1. Additionally, for the prescription drug variables in set 28, the edit rules described in Section 6.3.13 were run on the imputation-revised data to ensure that no unallowable inconsistencies were present.

6.5.2.4.5.1 30-Day Frequency of Use, Set 28: Preliminary Setup

For 30-day frequency of use, several steps not involving any modeling or stochasticity are handled in a preliminary setup program as follows:

- For cases with missing exact 30-day frequency, bounds are created. For prescription drugs, the lower bound is set to 1, and the upper bound is created using the edit rules as described in Section 6.3.13. For the nonprescription drugs in this set, both the lower and upper bounds are created based on the imputation-revised 12-month frequency of use. If the imputation-revised 12-month frequency of use is less than 30, then it forms the upper bound on the 30-day frequency of use. If the imputation-revised 12-month frequency of use is greater than 335, then the difference between the frequency and 335 is the lower bound on the 30-day frequency of use.
- For cases with missing exact 30-day frequency of use but nonmissing estimated 30-day frequency of use, the final edited 30-day frequency of use was assigned to the midpoint of the interval associated with the response category. For example, respondents selecting "10 to 19 days" as a response to the estimated 30-day frequency-of-use question were assigned a final edited 30-day frequency of use of 14.5. In rare cases, the bounds from the preceding step were used to shorten the interval before assigning the midpoint. For example, if the estimated 30-day

frequency of use was "10 to 19 days," but the upper bound for the 30-day frequency was 15, then the final edited 30-day frequency of use was 12.5 instead of 14.5.

- The imputation indicator variables are defined. See Section 6.5.2.4.5.5 for more information on the final imputation-revised and imputation indicator variables.
- The item response indicator used in the RP and PRD steps along with indicators used in the hot-deck step are defined.

6.5.2.4.5.2 *Response Propensity Step*

The overall process for the RP step was similar for the sets; however, the number of models fit and the inputs to the RP step varied from set to set. The total number of RP models fit for each set can be found in [Table 6.15](#). For each of the four key inputs to the RP adjustment step, details of any deviations are described below.

- **Analysis Weight:** All sets used the preliminary analysis weight, PANALWT.
- **Domain:** The domain for each set is detailed in [Table 6.15](#).
- **Item Response Indicator:** If a valid response was provided for the given measure for the drug of interest (i.e., item response was determined separately for each individual drug rather than across all drugs within a given imputation set), the individual was deemed an item respondent. For sets 30 and 35, respondents who reported an age at first use of 1 or 2 were treated as item nonrespondents in the RP and PRD steps because of the implausibility of such a young age at first use. In the hot-deck step, their response was left unchanged, but they were not allowed to be donors.
- **Covariates:** [Table 6.15](#) lists the tables in Appendix D that show both the starting and final covariates used in each of the RP models.

Table 6.15 Summary of Response Propensity Models for Continuous modPMN Variables

Set	Number of Response Propensity Models	Domain	Table Reporting Covariate Lists
27: 12-Month Frequency of Use	6	Past year users, based on imputation-revised recency of use	D.44
28: 30-Day Frequency of (Mis)Use	14	Past month users, based on imputation-revised recency of (mis)use	D.46
30: Age at First Use	6	Lifetime users, based on imputation-revised recency of use	D.48
35: Age at First Use, Hallucinogen Child Drugs	6	Lifetime hallucinogen users, based on IRHALLUCREC	D.50

modPMN = modified predictive mean neighborhood.

6.5.2.4.5.3 *Prediction Modeling Step*

As detailed in Section 6.5.2.4.2, each of the RP models segregated the domain members (including both respondents and nonrespondents) into five bins of roughly equal size, and

separate PRD models were fit within each of the bins. Descriptions of the variable of interest for each set are detailed below.

- **30-Day Frequency of Use, Set 28:** The response variable of interest in the 30-day frequency-of-use models, prior to a normalizing transformation, was the proportion of the days in the past month (30 days) on which a respondent used the drug of interest. The range of values for the proportion was from (greater than) 0 to 1. Therefore, to model 30-day frequency of use, an empirical logit transformation was computed for all respondents. This is the same transformation detailed in Section 6.5.2.4.2. See Section 6.5.1.3.6 for details on the transformation.
- **Age at First Use, Sets 30 and 35:** The response variable of interest in the AFU model, prior to a normalizing transformation, was the AFU as a proportion of the current age. The numerator in this proportion was an integer representing AFU. However, because this integer was in fact a truncated version of the real AFU, the value was made continuous by adding a random component between 0 and 1. The denominator in the proportion was the total age calculated from the respondent's birth date and interview date. The range of values for the proportion was from (greater than) 0 to 1. Therefore, to model AFU, an empirical logit transformation was computed for all respondents. See Section 6.5.1.4.2 for details on the transformation and calculation of the AFU as a proportion of the current age.

For all sets included in [Table 6.15](#), a linear univariate regression model using SUDAAN software was fit and the predicted means were calculated for both respondents and nonrespondents. These predicted means from the models were then back-transformed for use in the hot-deck step. [Tables D.47](#), [D.49](#), and [D.51](#) list both the starting and final covariates used in each of the PRD models.

6.5.2.4.5.4 Hot-Deck Step

Because all hot-deck steps for modPMN employ univariate matching and univariate assignment, the only deviations are found in the logical and likeness constraints applied. Below are details for sets 28, 30, and 35.

- **30-Day Frequency of Use, Set 28:** Due to small domain size for 30-day frequency of use and to increase the likelihood of keeping the delta constraint, the likeness constraint requiring donors and recipients to be in the same RP bin was not applied. For a complete list of logical and likeness constraints used, see [Tables E.86](#) and [E.87](#).
- **Age at First Use, Sets 30 and 35:** [Tables E.99](#) and [E.100](#) provide a complete list of likeness and logical constraints applied to the AFU imputations. The likeness constraints for AFU were more stringent than those for the other drug use measures, often resulting in the need to loosen the constraints.

6.5.2.4.5.5 Final Variables

The final imputation-revised variables for the other sets followed a similar pattern as those for set 27: the IR prefix followed by a drug abbreviation and ending with an indicator of

the specific measure. Table 6.16 provides the final imputation-revised variable format for each set. Each IR variable was accompanied by an imputation indicator with a prefix II instead of IR.

Table 6.16 Final Imputation-Revised Continuous modPMN Variables

Set	IR Variables ¹
27: 12-Month Frequency of Use	IRXXXYFQ
28: 30-Day Frequency of (Mis)Use	Nonprescription Drugs: IRXXX30N Prescription Drugs: IRXXXNM30FQ
30: Age at First Use	IRXXXAGE
35: Age at First Use, Hallucinogen Child Drugs	PCP and LSD: IRXXXAGE Ecstasy: IRECSTMOAGE

modPMN = modified predictive mean neighborhood.

¹XXX represents the drug abbreviation. In some cases, it may be longer than three characters (e.g., HALLUC for hallucinogens).

6.5.2.5 Sequencing of Steps for ModPMN Sets

This section provides the sequencing of steps for the modPMN sets 21 through 30 and 35. General information about the sequencing of steps for modPMN is available in Sections 3.7.2 and 3.7.3.

Imputation sets 21, 27, and 30 include three variables, so there are six 12-step blocks of programs included. The first of the six is for the first variable within the first cycle; the second is for the second variable within the first cycle; the third is for the third variable within the first cycle; the fourth is for the first variable within the second cycle; the fifth is for the second variable within the second cycle; and the sixth and last is for the third variable within the second cycle. Each 12-step block consists of an RP step where the domain members are segregated into five bins, followed by a PRD modeling step within each bin, followed by a collocated stochastic imputation step, and finally followed by a "wrap-up" program that merges the five datasets into one.

Imputation sets 22 and 29 include four variables, imputation set 26 contains six variables, and imputation sets 24 and 28 include seven variables. The algorithm is identical to the one used for imputation sets 21, 27, and 30, except there are a different number of 12-step blocks per cycle.

Imputation sets 23 and 25 include a single variable that eliminates the need for cycling. Therefore, these sets have a single 12-step block of programs.

6.5.2.6 Random Imputation, Sets 31-34 and 36-39

This section provides specific information about the random imputation of the date of first use for inhalants, hallucinogens, and methamphetamine (sets 36 through 39) along with the age and date of first misuse of the prescription drugs (sets 31 through 34).

6.5.2.6.1 *Date-of-First-Use Assignments*

After the AFU imputations, all lifetime users of the drug of interest had nonmissing AFU values.¹³¹ Using this AFU, users were assigned values for MFU, YFU, and DFU when these values were missing. Recent initiates, or those respondents whose AFU was within 1 year of his or her current age, were asked for their MFU and YFU. The DFU was not collected in the questionnaire and was missing for all respondents. The MFU, YFU, and DFU data contained four patterns of missingness:

Pattern 1: Recent initiates: missing day of first use only;

Pattern 2: Recent initiates: missing month/day of first use;

Pattern 3: Recent initiates: missing year/month/day of first use; and

Pattern 4: Less recent initiates: missing year/month/day of first use.

For each missingness pattern, upper and lower bounds on the date of first use were determined. Once the earliest and latest possible dates of first use were determined, a day was randomly selected from this interval. The imputation-revised month/day/year values were then extracted from this date of first use. For more information on the missingness patterns, see Sections 6.5.1.4.4.1 through 6.5.1.4.4.4.

For hallucinogens, the child drug recency and frequency information was used to bound the *latest possible date* for the date of first use. For example, respondents with LSD recency = 3 (i.e., lifetime but not past year user of LSD) could not have first used hallucinogens within the past year, regardless of the hallucinogens recency value. The bound created using the child drug recency and frequency was calculated in exactly the same way as for the parent recency and frequency information.

For hallucinogens child drugs (set 39) the imputation-revised hallucinogen DFU was used as an additional bound when setting the *earliest possible date* for the date of first use. This was done to ensure that the child drug's date of first use was never earlier than the parent drug's date of first use. Additionally, an indicator of lifetime hallucinogen drug use other than Ecstasy, PCP, and LSD was created. The algorithm for handling set 39 is detailed below:

1. The *earliest possible date*, *latest possible date*, and the final date of first use for each child drug for which the respondent was a lifetime user were assigned.
2. For respondents who were lifetime nonusers of other hallucinogens, it was determined which, if any, child drug could have had the same date of first use as hallucinogens. Specifically, it was determined whether the date of first use for hallucinogens was between the *earliest possible date* and *latest possible date* for each child drug. If none of the child drugs were eligible to receive the hallucinogens date of first use, nothing was

¹³¹ For prescription drugs, a date of first misuse was randomly selected and used to calculate the respondent's age at first misuse. For more details on age and date of first misuse, see Section 6.5.2.6.2.

done. Otherwise, one of the eligible child drugs was chosen at random, and its date of first use was overwritten with the hallucinogens date.

6.5.2.6.1.1 *Final Variables*

The final imputation-revised date-of-first-misuse variables were identified with the prefix IR, followed by a drug abbreviation and ending with an indicator of the measure (MFU = month of first use, YFU = year of first use, DFU = day of first use). Each IR variable was accompanied by an imputation indicator with the requisite II prefix. The levels for the imputation indicators were the standard levels used for most imputation-revised variables: 1 = questionnaire data, 2 = logically assigned, 3 = statistically imputed, and 9 = legitimate skip (not a lifetime user). Because survey respondents are not asked for the specific day on which they first used the drug of interest, all respondents in the domain receive IXXXXFU = 3. Also, as indicated above, only recent initiates are asked for the year and month of first drug use. Subsequently, these questions have high rates of nonresponse because of the skip logic embedded in the questionnaire, as all other individuals in the domain require imputation for their MFU and YFU.

6.5.2.6.2 *Age-at-First-Misuse and Date-of-First-Misuse Assignments for Prescription Drugs*

For age at first misuse and date of first misuse for prescription drugs, only past year initiates (i.e., respondents who misused the drug of interest for the first time in 12 months prior to the date of the interview) are included in the domain. Respondents who were not past year initiates receive an appropriate skip code. As a result, the age at first misuse (AFM) is limited to either the respondent's age or to the respondent's age – 1. For each specific prescription drug a respondent misused in the past 12 months, respondents were asked his or her AFM. Recent initiates were asked for their month of first misuse (MFM) and year of first misuse (YFM). The day of first misuse (DFM) was not collected in the questionnaire and was missing for all respondents.

For each specific prescription drug within a given category, candidate upper and lower bounds on the date of first misuse (i.e., the earliest possible date of first misuse and the latest possible date of first misuse) were determined. Additionally, at the category level (i.e., overall pain relievers), additional candidate upper bounds are determined based on the interview date, recency of misuse, and the 30-day frequency of misuse. The final category-level lower bound is the minimum of all candidate lower bounds, and the final category-level upper bound is the maximum of all candidate upper bounds. For example, suppose a respondent was a past year initiate of two specific pain relievers with the following bounds: a lower bound of January 1 and an upper bound of January 31 for the first drug and a lower bound of March 1 and an upper bound of March 31 for the second drug. The pain reliever lower bound would be January 1 and the upper bound would be January 31. Note that if the maximum of the lower bound was taken, a negative interval would have been produced.

In the absence of any other available information to restrict the possible date of first misuse, the lower bound is set to 1 year before the interview date. Once the earliest and latest possible dates of first misuse for the category were determined, a day was randomly selected

from this interval. The imputation-revised age, month, day, and year were extracted from this date of first misuse.

The process described above assumes equal probability for all days in the final interval; however, not all days in the interval should always have an equal probability. For example, suppose a respondent was interviewed on July 31, 2015, and indicated misusing two pain relievers in the past year, OxyContin® and Vicodin. The respondent indicated first misusing OxyContin® in February 2015 but left the year and month of first use for Vicodin as missing. The final category-level interval for possible days of initiation for pain relievers would be August 1, 2014, through February 28, 2015. Ignoring information provided by recency of misuse and 30-day frequency for simplicity, the respondent could have initiated Vicodin any day between August 1, 2014, and July 31, 2015. Likewise, the respondent could have initiated OxyContin® any day between February 1, 2015, and February 28, 2015. Thus, the days occurring in February should actually have a higher probability of being the overall date of initiation for pain relievers than days that are not in February. Due to the complication in calculating these probabilities as the number of specific drugs misused increases and the possibility of incomplete information at the specific drug level, the decision was made to assume equal probability for all days in the final category-level interval.

6.5.2.6.2.1 *Defining Upper and Lower Bounds*

The upper and lower candidate bounds are defined as follows. Depending on the information provided by the past year initiate, some candidate bounds will be missing.

Candidate lower bounds =

- January 1 of YFM, if YFM provided
- First day of the month indicated by MFM/YFM, if MFM and YFM provided
- AFMth birthday, if AFM provided

Candidate upper bounds =

- December 31 of YFM, if YFM provided
- Last day of the month indicated by MFM/YFM, if MFM and YFM provided
- 1 day before (AFM + 1)th birthday, if AFM provided
- Interview date
- Interview date – category level 30-day frequency + 1, if category level recency is past month
- 30 days before the interview date, if category level recency is past year not past month¹³²

¹³² For 2016, there was one case where this upper bound candidate was dropped due to a negative interval. For more information, see inconsistent pattern 8 in [Table 6.5](#).

If a past year initiate provided no valid information, then the respondent could have initiated on any day in the past year giving a lower bound of interview date – 1 year and an upper bound of the interview date.

If a past year initiate was interviewed on the last day of the month and the respondent reported first misuse of a specific drug in the same month as the interview in the previous calendar year, then the lower bound was set to the upper bound. That is, the date of first misuse was imputed to the last day of the month in the previous calendar year.

6.5.2.6.2.2 *Final Variables*

As with all other imputation-revised variables, the final imputation-revised age-at-first-misuse and date-of-first-misuse variables were identified with the prefix IR, followed by an eight-letter identifier, where a three-letter code identified the drug and the final five letters identified the measure (NMAGE = age at first misuse, NMMFU = month of first misuse, NMYFU = year of first misuse, NMDFU = day of first misuse). Each IR variable was accompanied by an imputation indicator with the requisite II prefix. The levels for the imputation indicators were 1 = questionnaire data, 2 = logically assigned, 3 = statistically imputed, past year initiation imputed, 4 = statistically imputed, past year initiation known, and 9 = legitimate skip (not a past year initiate). Because survey respondents are not asked for the specific day on which they first misused the drug of interest, all respondents in the domain receive IXXXXNMDFU = 3 or 4.

6.5.3 Recodes for Additional Analyses

From the final imputation-revised variables, some additional recodes were created to use in additional analyses. For recency-of-use variables, this included three dichotomous indicator variables to indicate lifetime (mis)use (XXX(NM)FLAG), past year (mis)use (XXX(NM)YR), and past month (mis)use (XXX(NM)MON). For the prescription drugs, two dichotomous indicator variables to indicate lifetime any use (XXXANYFLAG) or past year any use (XXXANYYR) were also created. Several other prevalence recodes, which covered the same three measures, were created to incorporate information from several different drugs. The creation of these variables was also straightforward. If the respondent was a lifetime user of any of the drugs, then the FLAG variable was set to 1; otherwise, it was set to 0. The YR and MON variables were processed in a similar manner. [Tables 6.17](#) and [6.18](#) contain information on all dichotomous indicator recodes created. Because age of first misuse for prescription drugs is only defined among past year initiates, the age-at-first-use variables that were created to incorporate information from several different drugs prior to 2015 are no longer created.

Table 6.17 Prevalence Recodes

Drug	Variable Names
Cigarettes	CIGFLAG, CIGYR, CIGMON
Smokeless Tobacco	SMKLSSFLAG, SMKLSSYR, SMKLSSMON
Cigars	CGRFLAG, CGRYR, CGRMON
Pipes	PIPFLAG, PIPMON
Alcohol	ALCFLAG, ALCYR, ALCMON
Marijuana	MRJFLAG, MRJYR, MRJMON

Table 6.17 Prevalence Recodes (continued)

Drug	Variable Names
Cocaine	COCFLAG, COCYR, COCMON
Crack	CRKFLAG, CRKYR, CRKMON
Hallucinogens	HALLUCFLAG, HALLUCYR, HALLUCMON
LSD	LSDFLAG, LSDYR, LSDMON
PCP	PCPFLAG, PCPYR, PCPMON
Ecstasy	ECSTMOFLAG, ECSTMOYR, ECSTMOMON
Tryptamine Hallucinogens DMT, AMT, or "Foxy"	DAMTFXFLAG, DAMTFXYR, DAMTFXMON
<i>Salvia divinorum</i>	SALVIAFLAG, SALVIAYR, SALVIAMON
Ketamine	KETMINFLAG, KETMINYR, KETMINMON
Inhalants	INHALFLAG, INHALYR, INHALMON
Methamphetamine	METHAMFLAG, METHAMYR, METHAMMON
Pain Relievers	PNRANYFLAG, PNRANYR, PNRNMFLAG, PNRNMYR, PNRNMMON
Sedatives	SEDANYFLAG, SEDANYR, SEDNMFLAG, SEDNMYR, SEDNMMON
Stimulants	STMANYFLAG, STMANYR, STMNMFLAG, STMNMYR, STMNMMON
Tranquilizers	TRQANYFLAG, TRQANYR, TRQNMFLAG, TRQNMYR, TRQNMMON

Table 6.18 Prevalence Recodes Incorporating More than One Recency Variable

General Drug Category	Variable Names	Source Recency Variables
Tobacco	TOBFLAG, TOBYR, TOBMON	Cigarettes, smokeless tobacco, cigars, pipes
Psychotherapeutics, Any Use	PSYANYFLAG, PSYANYR	Any use of pain relievers, tranquilizers, stimulants, sedatives
Psychotherapeutics, Misuse	PSYCHFLAG, PSYCHYR, PSYCHMON	Misuse of pain relievers, tranquilizers, stimulants, sedatives
Illicit Drugs Other than Marijuana	ILLEMFLAG, ILLEMYR, ILLEMMON	Psychotherapeutics (misuse), plus inhalants, hallucinogens, cocaine, and heroin
Illicit Drugs, but Only Marijuana	MJONLYFLAG, MJONLYR, MJONLYMON	Same as MRJFLAG, MRJYR, and MRJMON, except set to 0 if the corresponding ILLEM variable is equal to 1
Illicit Drugs	ILLFLAG, ILLYR, ILLMON	Illicit drugs other than marijuana, plus marijuana

Additionally, for 30-day binge alcohol frequency of use, two recodes were created. The first, BNGDRKMON, is an indicator of any binge alcohol use in the past 30 days. The other, HVDYDRKMON, is an indicator of heavy alcohol use in the past 30 days. Heavy alcohol use is defined as drinking five or more drinks on the same occasion for males or four or more drinks on the same occasion for females on each of 5 or more days in the past 30 days.

Prior to 2015, binge alcohol use was defined as drinking five or more drinks on the same occasion, regardless of gender. Similarly, prior to 2015, heavy alcohol use was defined as drinking five or more drinks on the same occasion on each of 5 or more days in the past 30 days. Therefore, these recodes are not comparable with the version of these recodes created from 2002-2014 overall or among females; they are comparable among males.

7. Editing and Imputation for Variables from the Special Drugs and Substance Dependence and Abuse Sections

7.1 Introduction

This chapter discusses editing procedures for the special drugs and substance dependence and abuse sections. The chapter also discusses imputation procedures for substance dependence and abuse variables. These interview sections are included in this chapter because variables for cocaine, heroin, and methamphetamine from the special drugs section are used in editing variables for substance dependence and abuse for the corresponding drugs.

The structure of this chapter also differs from the structure of Chapters 4 through 6 because only a selected set of variables were imputed for substance dependence and abuse. Section 7.2 describes the editing procedures for the variables in the special drugs section; these variables did not undergo imputation. Section 7.3.1 describes the editing procedures for the variables in the substance dependence and abuse section, and Section 7.3.2 describes the imputation procedures for the variables in the substance dependence and abuse section.

Some substance dependence and abuse variables other than cigarette (nicotine) dependence have undergone formal imputation since 2015. Prior to 2015, all substance dependence and abuse variables other than cigarette (nicotine) dependence underwent a simple imputation method as described in Appendix D of the 2016 analytic codebook (Center for Behavioral Health Statistics and Quality, 2017b), where missing responses were replaced by negative responses (called "zero-fill imputation").

The imputation procedures for substance dependence and abuse were implemented for hallucinogens, inhalants, methamphetamine, prescription pain relievers, prescription tranquilizers, prescription stimulants, and prescription sedatives (subsequently referred to as "modPMN drugs" because dependence and abuse data for these substances were imputed using the modified predictive mean neighborhood [modPMN] procedures described in Chapter 3). These procedures were not used for dependence and abuse data for alcohol, marijuana, cocaine, and heroin (subsequently referred to in this chapter as "zero-fill drugs"). For these latter substances, dependence and abuse data were imputed by treating missing values as being equivalent to negative responses, as described in Appendix D of the 2016 analytic codebook (Center for Behavioral Health Statistics and Quality, 2017b). The overall indicators of substance dependence and abuse for the modPMN drugs were imputed beginning in 2015 because imputation was determined to yield significant improvements in data quality, as discussed in Chapter 10 of the evaluation of imputation methods for the NSDUH report (Center for Behavioral Health Statistics and Quality, 2017e). Since 2015, however, the indicators of dependence or abuse for the zero-fill drugs have continued to undergo zero-fill imputation because changing the imputation method for these measures could disrupt trends.

7.2 Special Drugs

The special drugs section asked about the use of GHB (gamma hydroxybutyrate) and the use of over-the-counter (OTC) cold medicines to get high; questions about GHB and OTC cold medicines have been included in this section since 2006. The special drugs section also asked about smoking and sniffing of heroin; use of heroin, cocaine, methamphetamine, or other drugs with a needle; general needle use behaviors (e.g., needle sharing); and where respondents got the last needle that they used.

7.2.1 GHB and Over-the-Counter Cold Medicines

Question SD01 in 2016 asked if respondents had ever used nonprescription cough or cold medicines (also known as OTC medicines) to get high. Question SD03 asked if respondents had ever used GHB (also called "G," "Georgia Home Boy," "Grievous Bodily Harm," "Liquid G," or gamma hydroxybutyrate). If respondents reported that they ever used OTC cough or cold medicines to get high or that they ever used GHB, they were asked to report when they last engaged in that behavior.¹³³

In addition, respondents who reported that they last used OTC cough or cold medicines to get high in the past 12 months were asked to specify what OTC drugs they used to get high in that period. Respondents could specify up to five drugs that they used to get high. Therefore, the edited variables COLDYR1 through COLDYR5 (corresponding to the "OTHER, Specify" items SD02A through SD02E) captured up to five medications that respondents reported using to get high.

The "OTHER, Specify" variables for OTC cough and cold medicines were coded using a decimal system to preserve relationships between similar responses and to capture additional details that respondents provided. For example, there are multiple formulations of OTC cold medicines under the brand name of Robitussin[®] that vary in terms of the specific ingredients they contain, such as the OTC cough suppressant dextromethorphan (DXM). Therefore, if respondents specified that they used a Robitussin[®] product to get high in the past 12 months but the only detail they provided was that they used "Robitussin," that response was assigned a default code of 829.0 (Robitussin[®], no other information). However, if respondents specifically reported that they used Robitussin[®] DM to get high (a product that contains DXM), that response was coded as 829.1 and was documented as Robitussin[®] DM. Thus, the main code of 829 indicated use of a Robitussin[®] product to get high in the past 12 months, and the decimal value indicated whether respondents provided additional detail about their use of specific Robitussin[®] products to get high.

An important aspect of the processing of these drug variables consisted of assigning a code of 91 (or 9991) to variables that had been skipped because the questions did not apply. For example, if respondents reported in question SD03 (edited variable GHB) that they never used GHB, the edited recency-of-use variable GHBREC (corresponding to question SD04) was assigned a code of 91. Similarly, if respondents reported that they last used an OTC cough or

¹³³ Although the order of the questions for OTC cough and cold medicines changed between 2014 and 2015 (i.e., relatively late in the special drugs section in 2014 vs. the first set of questions beginning in 2015), this change was determined not to have affected the comparability of the data between 2015 and prior years.

cold medicine to get high more than 12 months ago, COLDYR1 through COLDYR5 were assigned a code of 9993.0 to indicate that respondents had used OTC cough or cold medicines to get high, but not in the past 12 months.

In addition, if respondents reported that they had ever used these drugs but they did not know or refused to report when they last used the drug, the edited recency-of-use variable (e.g., GHBREC) was assigned a code of 9 (Used at some point in the lifetime LOGICALLY ASSIGNED). If the most recent use of OTC cough or cold medicines (edited variable COLDREC, corresponding to question SD02) had been assigned a code of 9, COLDYR1 through COLDYR5 were left as blank (code of 9998.0) because some of these respondents may have used OTC cough or cold medicines to get high in the past 12 months.

Data from the "OTHER, Specify" needle use variables OTDGNDLA through OTDGNDLE from questions SD15SP to SD15D (Section 7.2.5.2) and the cough and cold medicine variables COLDYR1 through COLDYR5 were used in editing GHB and GHBREC. However, OTDGNDLA through OTDGNDLE were not used to edit COLDMEDS (corresponding to question SD01) and COLDREC (corresponding to SD02) because there could be too much room for interpretation regarding what might constitute a nonprescription cough or cold medicine. In contrast, the questions for the other additional drugs in the special drugs section pertained to specific drugs.

Thus, for example, if a respondent had not reported lifetime use of GHB but the respondent specified using GHB with a needle in OTDGNDLA through OTDGNDLE, then the edited variable GHB (corresponding to question SD03) would be assigned a code of 3 (Yes LOGICALLY ASSIGNED). In turn, the recency-of-use variable GHBREC (corresponding to question SD04) would be assigned a code of 9 (Used at some point in the lifetime LOGICALLY ASSIGNED). However, there were no situations in 2016 in which data from the "OTHER, Specify" needle use variables OTDGNDLA through OTDGNDLE resulted in any responses being logically inferred in GHB or GHBREC. Nevertheless, this logic is in place for handling these types of patterns for GHB in future survey years.

In addition, if a respondent specified use in the past year of GHB, as a nonprescription cough or cold medicine in COLDYR1 through COLDYR5, that response was used to edit the other drug data. For example, if question SD03 was not answered as "yes" (or had not logically been inferred as "yes" from the above edits based on OTDGNDLA through OTDGNDLE), the corresponding edited variable GHB was assigned a code of 3 (Yes LOGICALLY ASSIGNED). In turn, the corresponding edited recency-of-use variable GHBREC was assigned a code of 8 (Used at some point in the past 12 months LOGICALLY ASSIGNED) because COLDYR1 through COLDYR5 pertain to use in the past 12 months. If a respondent had answered question SD03 for GHB as "yes" and reported last using it more than 12 months ago but also specified using it in the past 12 months in COLDYR1 through COLDYR5, GHBREC also was assigned a code of 8. Although this edit was in place, no respondents in 2016 were inferred to have used GHB in the past 12 months based on data from COLDYR1 through COLDYR5.

Because the questions corresponding to COLDMEDS and COLDREC asked specifically about use of OTC cough or cold medicines, respondents also would be expected to report at least one OTC medicine in COLDYR1 through COLDYR5 or to have some response of "don't know"

or "refused" in the corresponding items (i.e., if respondents did not know the name of the drug or did not want to report it). Therefore, a variable called OTCFLAG was created to indicate when respondents' reports of use of OTC cough or cold medicines in the past year were not substantiated by the "OTHER, Specify" data.¹³⁴ OTCFLAG had the values shown below when these situations occurred.

- OTCFLAG had a value of 1 (i.e., "yes") when respondents specified only the use of drugs that require a prescription in the United States¹²⁸ or drugs that are not available as OTC medicines in the United States, even if they may be available as OTC medicines elsewhere.
- OTCFLAG had a value of 2 (i.e., "no") when COLDYR1 through COLDYR5 had values only of 9985.0 (bad data) or 9998.0 (blank).

No further editing was done to the OTC variables when OTCFLAG indicated that respondents' reports of use of OTC cough or cold medicines to get high in the past year were not substantiated in COLDYR1 through COLDYR5. However, OTCFLAG would alert analysts when this issue occurred.

7.2.2 General Edits of Cocaine, Heroin, and Methamphetamine Variables for Nonusers

An important aspect of the processing of variables in this section pertaining to the use of cocaine, heroin, or methamphetamine with a needle or the smoking or sniffing of heroin consisted of assigning codes of 91, 93, or 99 (Section 2.4.2) to variables that had been skipped because the questions did not apply. For example, respondents who never used cocaine, heroin, or methamphetamine were not asked questions in the special drugs section that pertained to these drugs. [Table B.11](#) in Appendix B describes specific edits that were implemented for these variables when the corresponding items were skipped.

7.2.3 Edits for Heroin Users Who Did Not Report Smoking, Sniffing, or Injection of Heroin

Respondents who reported in the earlier heroin section of the interview that they were lifetime heroin users but who answered "no" to all questions about smoking heroin (question SD07), sniffing heroin (question SD09), or using it with a needle (question SD11) were asked a follow-up question SDHEUSE to determine how these respondents administered the heroin they had reported using. SDHEUSE was an "enter all that apply" type of question that allowed respondents to report multiple ways that they used heroin. SDHEUSE included response options for smoking heroin, sniffing heroin, using heroin with a needle, or use of heroin "some other way." Respondents who reported using heroin "some other way" were asked to specify in question SDHEUSE2 what this "other" mode of heroin administration was. If respondents specified using heroin in a way that corresponded to one of the ways that they had been asked about, it was logically inferred that the response in the edited SDHEUSE question should have been chosen. For example, if respondents did not report in SDHEUSE that they smoked heroin

¹³⁴ If respondents reported the use of drugs that may require a prescription in some locations in the United States but may be available without a prescription elsewhere in the United States (e.g., cough medicines containing low doses of codeine), the medication was assumed to have been obtained without a prescription.

but they specified this as "some other way," the edited variable HEOTSMK was assigned a code of 3 (Response entered LOGICALLY ASSIGNED).

In turn, HEOTSMK, HEOTSNF, and HEOTNDL were used to edit the corresponding lifetime heroin variables HERSMOKE, HERSNIFF, and HERNEEDL, respectively. For example, if HEOTSMK indicated that the respondent had smoked heroin (HEOTSMK = 1 or 3), HERSMOKE was edited to indicate that the respondent was logically inferred to have smoked heroin at least once in his or her lifetime.

7.2.4 Additional Edits for Smoking or Sniffing of Heroin

As noted previously, respondents in 2016 who reported that they had ever used heroin were asked in questions SD07 and SD09 to report whether they had ever smoked or sniffed heroin (i.e., inhaled heroin powder through their nose), respectively. Respondents were routed to follow-up recency questions for smoking or sniffing of heroin if they answered the corresponding lifetime questions as "yes" or if they reported using heroin in either of these ways in the corresponding question SDHEUSE (Section 7.2.3). Information from these questions was used in the creation of the heroin smoking and sniffing recency variables HRSMKREC and HRSNFREC. If respondents did not report lifetime use of heroin in these ways (e.g., SD07 for smoking heroin was answered as "no") and smoking or sniffing of heroin was not reported in SDHEUSE, then HRSMKREC and HRSNFREC were assigned codes of 93 to indicate that the respondents had used heroin but had not smoked it or had not sniffed it, respectively.

Also, if respondents initially did not report using heroin in these ways in SDHEUSE, they were skipped out of the follow-up recency questions for smoking and sniffing of heroin. Therefore, if respondents' only indication of smoking or sniffing of heroin came from the "OTHER, Specify" response associated with SDHEUSE, the corresponding variables HRSMKREC and HRSNFREC were set to 9 (Used at some point in the lifetime LOGICALLY ASSIGNED).

Consistent with the principles described in Section 2.4.1 for editing across sections, no editing was done to the heroin data in an earlier section of the interview if lifetime heroin users reported more recent smoking or sniffing of heroin than they reported for when they last used heroin. Consequently, HRSMKREC or HRSNFREC could indicate more recent use of heroin than was indicated in the edited heroin recency variable HERREC or in the imputed heroin recency variable IRHERRC.

Table B.12 describes specific edits other than those pertaining to skip logic that were implemented in the special drugs section for smoking and sniffing of heroin. For example, lifetime users of heroin could report that they smoked heroin at least once but not indicate when they last smoked it. The general edit was to assign a nonspecific value to the edited recency variable (i.e., HRSMKREC) to indicate that the respondent smoked heroin at some point in his or her lifetime. In some special situations, however, it was possible to infer that respondents could not have smoked heroin in the past 12 months. In these situations, respondents reported last using heroin more than 12 months ago, and there were no responses for other heroin-related questions in the special drugs section to indicate that these respondents had used heroin in the past 12 months.

7.2.5 Injection Drug Use

Respondents in 2016 who reported lifetime use of the relevant substance were asked whether they had ever used a needle to inject cocaine (question SD05), heroin (questions SD11 or SDHEUSE), or methamphetamine (SD13). Respondents also were asked in question SD15 whether they used a needle to inject other drugs (if they reported injecting cocaine, heroin, or methamphetamine) or any drug (otherwise). Respondents who reported lifetime injection of a drug were routed to follow-up recency questions (SD06 for cocaine, SD12 for heroin, SD14 for methamphetamine, and SD16 for other drugs). Information from these questions was used in the creation of the corresponding needle recency variables CONDLREC, HRNDLREC, METHNDLRC, and OTDGNDLRC.

7.2.5.1 Injection of Cocaine, Heroin, or Methamphetamine

Table B.12 describes specific edits other than those pertaining to skip logic that were implemented in the special drugs section for the use of drugs with a needle. For example, lifetime users of cocaine could report that they used cocaine with a needle at least once but not indicate when they last injected it. The general edit was to assign a nonspecific value to the edited recency variable (i.e., CONDLREC) to indicate that the respondent used cocaine with a needle at some point in his or her lifetime. In some special situations, however, it was possible to infer that respondents could not have injected cocaine in the past 12 months. In these situations, respondents reported last using cocaine more than 12 months ago in the cocaine section, and there were no responses for other cocaine-related questions in the special drugs section to indicate that these respondents had used cocaine in the past 12 months.

Consistent with the principles described in Section 2.4.1 for editing across sections, no editing was done to the corresponding data in earlier sections of the interview if respondents reported more recent use of cocaine, heroin, or methamphetamine with a needle than they previously reported for when they last used these drugs. Consequently, CONDLREC, HRNDLREC, or METHNDLRC could indicate more recent use of these drugs than was indicated in the corresponding edited or imputed recency variables (e.g., HERREC and IRHERRC for heroin).

As noted previously, respondents in 2016 also were asked in question SD15 (edited variable OTDGNEDL) whether they ever used a needle to inject "some other drug" (if respondents previously reported lifetime use of cocaine, heroin, or methamphetamine with a needle) or "any drug" (if respondents had not previously indicated use of any of these drugs with a needle). If question SD15 was answered as "yes" (OTDGNEDL = 1), respondents then were asked to specify what (other) drug(s) they used with a needle. Respondents could specify up to five drugs that they had injected (edited variables OTDGNDLA through OTDGNMLE).

Consequently, it was possible for respondents to have reported in a previous substance use section that they never used cocaine, heroin, or methamphetamine but then to specify use of that drug with a needle in OTDGNDLA through OTDGNMLE. For example, respondents could indicate in the earlier heroin section that they never used heroin but then specify lifetime injection of heroin in OTDGNDLA through OTDGNMLE. Consistent with the principles described in Section 2.4.1 for editing across sections, no editing was done to the corresponding

data in earlier sections of the interview if respondents reported that they never used cocaine, heroin, or methamphetamine but they reported using any of these drugs with a needle in the special drugs section of the interview. However, these respondents were logically inferred in the relevant special drugs variables to be users of that particular drug at some point in the lifetime. Consequently, CONDLREC, HRNDLREC, or METHNDLRC could indicate use of cocaine, heroin, or methamphetamine, respectively, even if the corresponding edited or imputed variables from earlier sections of the interview indicated lifetime nonuse. Thus, for example, if respondents reported in the heroin section that they never used heroin, but then they specified heroin as "some other drug" that they used with a needle, then the edited lifetime heroin needle use variable HERNEEDL was assigned a code of 3 (Yes LOGICALLY ASSIGNED) and the heroin needle recency variable HRNDLREC was assigned a code of 9 (Used at some point in the lifetime LOGICALLY ASSIGNED). In this example, the edited heroin recency variable HERREC retained a value of 91 and the imputed heroin recency variable IRHERRC retained a value of 9; these values indicated that the respondent had previously reported that he or she never used heroin.

Respondents also could specify in OTDGNDLA through OTDGNDLE that they injected only cocaine, only heroin, or only methamphetamine in their lifetime and then report more recent use of this "other" drug with a needle in SD16 (edited variable OTDGNDLRC) than they had reported in the corresponding needle recency questions for cocaine, heroin, or methamphetamine. When this occurred, CONDLREC, HRNDLREC, or METHNDLRC were assigned the value of most recent injection of other drugs from OTDGNDLRC, plus a value of 10 (i.e., "bump" by 10). For example, if a respondent reported last injecting cocaine more than 12 months ago (SD06 = 3), specified the lifetime injection of only cocaine, and reported last injecting other drugs more than 30 days ago but within the past 12 months (OTDGNDLRC = 2), then CONDLREC was assigned a code of 12 (More than 30 days ago but within the past 12 months LOGICALLY ASSIGNED).

This edit also was implemented if respondents reported lifetime injection of cocaine, heroin, or methamphetamine but the most recent time when respondents injected the drug was unknown. For example, if CONDLREC = 9 (At some point in the lifetime) and OTDGNDLRC = 3 (More than 12 months ago), then CONDLREC was assigned a code of 13. In this situation, it could be inferred that the respondent's most recent use of cocaine with a needle occurred more than 12 months prior to the interview date.

This edit was not implemented if respondents specified that they injected other drugs besides cocaine, heroin, or methamphetamine in their lifetime or they injected more than one of these specific drugs in their lifetime. In this situation, it was not possible to infer that more recent injection of some other drug applied specifically to cocaine, heroin, or methamphetamine.

7.2.5.2 Injection of Other Drugs

As noted previously, respondents could report in question SD15 that they used a needle to inject other drugs, or that they used a needle to inject any drug, if they had not previously reported using a needle to inject cocaine, heroin, or methamphetamine. Respondents who answered SD15 as "yes" were asked to specify the names of up to five other drugs that they injected in their lifetime.

Respondents could report that they injected other drugs in their lifetime, even though they previously reported that they never used marijuana, cocaine, heroin, hallucinogens, inhalants, or methamphetamine, and that they had never misused prescription psychotherapeutic drugs (i.e., pain relievers, tranquilizers, stimulants, or sedatives). However, the data from OTDGNDLA through OTDGNDLE could be used to identify respondents who corroborated their report of lifetime injection drug use (e.g., if injection of anabolic steroids was reported) despite having previously reported nonuse (or no misuse) of all drugs covered in the previous substance use sections. Similarly, it was possible from "OTHER, Specify" data on other drugs that respondents had injected to identify those whose needle use probably was limited to legitimate, medical uses (e.g., injection of antibiotics). Therefore, it was logically inferred that respondents had never used a needle to inject other drugs if (1) they were lifetime nonusers of all drugs covered in the sections for the substances that were described previously; (2) they indicated that they never engaged in behaviors that would be indicative of the misuse of drugs with a needle, such as needle sharing, use of bleach to clean needles, or reusing of needles; and (3) all of the "other" drugs they reported using with a needle typically were not drugs of abuse. In this situation, the variable OTDGNEDL corresponding to question SD15 was set to a value of 4 (No LOGICALLY ASSIGNED). Edits to the general needle use variables when OTDGNEDL had been assigned a code of 4 are discussed in Section 7.2.5.4.

Just as data from question SD16 for the most recent injection of other drugs could be used to infer more recent injection in the edited variables CONDLREC, HRNDLREC, and METHNDLRC, data from CONDLREC, HRNDLREC, and METHNDLRC could be used to infer more recent injection in OTDGNDLRC. As for the edits described in Section 7.2.5.1, data from CONDLREC, HRNDLREC, and METHNDLRC were used to edit OTDGNDLRC only when respondents had specified the lifetime injection of only cocaine, only heroin, or only methamphetamine in OTDGNDLA through OTDGNDLE. For example, if a respondent had reported the lifetime injection of only cocaine, reported last injecting this as "some other drug" more than 12 months ago (SD16 = 3), but the respondent had reported using a needle to inject cocaine more than 30 days ago but within the past 12 months (CONDLREC = 2), then OTDGNDLRC was assigned a code of 12 (More than 30 days ago but within the past 12 months LOGICALLY ASSIGNED).

7.2.5.3 Recoded Needle Use Variables

The recoded needle use variable CHMNDLREC (for cocaine, heroin, and methamphetamine) was created beginning in 2015 from variables pertaining to the use of cocaine, heroin, or methamphetamine with a needle. CHMNDLREC replaced recoded needle use variables from 2014 and earlier (e.g., STNEEDLR, STNLRECR, NEDLRECR) because of the changes to the special drugs section in 2015 and the movement of questions about the injection of prescription stimulants in the past 12 months or past 30 days to the stimulants section of the interview (Section 1.4.1).

Specifically, the needle recency variable CHMNDLREC was created from the variables CONDLREC (most recent use of cocaine with a needle), HRNDLREC (most recent use of heroin with a needle), and METHNDLRC (most recent use of methamphetamine with a needle). The logic for creating CHMNDLREC is documented below.

- If respondents never used a needle to inject any of these drugs (including situations in which respondents never used cocaine, heroin, or methamphetamine), CHMNDLREC indicated that the respondents had never used cocaine, heroin, or methamphetamine with a needle.
- If respondents reported using one or more of these drugs with a needle, the general principle in assigning a value to CHMNDLREC was to identify the most recent use reported by the respondents. In particular, if respondents reported using one or more of these drugs with a needle in the past 30 days, it could be determined unambiguously that the respondents were past month needle users.
- In other situations, however, if one or more of the cocaine, heroin, or methamphetamine needle recency variables indicated nonspecific use at some point in the respondents' lifetime, CHMNDLREC was assigned a value to indicate nonspecific past year or lifetime use. For example, if respondents reported last using cocaine with a needle more than 30 days ago but within the past 12 months, yet all that was known was that they used heroin with a needle at some point in their lifetime, it still could be inferred that they had used some drug with a needle in the past 12 months. The nonspecific value for past year use was assigned (i.e., a code of 8) because the respondents could have been more recent users of heroin with a needle.
- Similarly, if respondents indicated use of one of these drugs with a needle more than 12 months ago and they did not know or refused to indicate when they last used one of the other drugs with a needle, a nonspecific value of lifetime use (i.e., a code of 9) was assigned to CHMNDLREC because the respondents may have been more recent needle users than what they had reported elsewhere.

The recoded needle use variable ANYNDLREC (most recent use of any drug with a needle) also was created beginning in 2015 to take into account data from CONDLREC, HRNDLREC, METHNDLRC, and OTDGNDLRC. The same general principles that were described above for the creation of CHMNDLREC were used to create ANYNDLREC. For example, if CONDLREC, HRNDLREC, METHNDLRC, or OTDGNDLRC indicated use of drugs with a needle, the general principle in assigning a value to ANYNDLREC was to identify the most recent use reported by the respondents. In particular, if CONDLREC, HRNDLREC, METHNDLRC, or OTDGNDLRC indicated use of a drug with a needle in the past 30 days, then ANYNDLREC was assigned a code of 1 (Within the past 30 days). In this situation, it could be determined unambiguously that the respondents were past month needle users.

Consistent with the general principle of not editing across sections (Section 2.4.1), however, ANYNDLREC did not capture data from respondents who reported in the prescription stimulants section of the interview that they injected prescription stimulants in the past 12 months or past 30 days unless they reported that they injected cocaine, heroin, methamphetamine, or other drugs in the special drugs section. Also, unlike the data in the special drugs section for the most recent use of a needle to inject drugs, the needle use data for prescription stimulants covered only the past 12 months and past 30 days but did not capture information about the injections of prescription stimulants more than 12 months prior to the interview. Consequently, if respondents indicated in the special drugs section that they never injected cocaine, heroin, methamphetamine, or other drugs and they indicated in the prescription

stimulants section that they never injected stimulants, then the assumption would need to be made from the data for other drugs that these respondents did not inject prescription stimulants more than 12 months ago. Had respondents been given the opportunity in the prescription stimulants section to report injection of stimulants more than 12 months ago, however, some respondents might have reported the injection of stimulants in that period despite reporting in the special drugs section that they never injected cocaine, heroin, methamphetamine, or other drugs.

7.2.5.4 General Needle Use

Questions SD17 through SD20 asked lifetime needle users about general needle use behaviors, including reuse of a needle that the respondent had used before (SD17; edited variable GNNDREUS), use of a needle that the respondent knew or suspected had been previously used by someone else (SD18; edited variable GNNDLSH1), use of bleach to clean the needle the last time that the respondent used a needle to inject drugs (SD19; edited variable GNNDCLLEN), and use of the needle by someone else after the respondent last used a needle (SD20; edited variable GNNDLSH2). In addition, respondents were asked in question SD21 (and in the corresponding "OTHER, Specify" variable, if applicable) to report how they got the last needle that they used for injecting drugs (edited variable GNNDGET).

An important aspect of the processing of variables for general needle use behaviors or the source of the last needle that respondents used consisted of assigning a legitimate skip code of 99 if respondents had never used a needle to inject drugs (Section 2.4.2). In particular, if respondents did not report that they used a needle to inject cocaine, heroin, or methamphetamine, then question SD15 asked whether respondents ever used a needle to inject *any* drug. Even if the lifetime needle use variables for cocaine, heroin, or methamphetamine had missing data (e.g., "don't know" or "refused"), a response of "no" in SD13 could be interpreted to mean that the respondent had never used a needle to inject any drug. Consequently, when SD15 was answered as "no" (OTDGNEDL = 2), then GNNDREUS through GNNDGET were assigned legitimate skip codes.

When OTDGNEDL (corresponding to question SD15) had been set to a value of 4 as described in Section 7.2.5.2, any data in the general needle use variables GNNDREUS through GNNDGET were replaced with a code of 89 (LEGITIMATE SKIP Logically assigned) to indicate that respondents logically should have skipped these items because they appeared to be reporting about legitimate use of drugs with a needle. However, this edit was not implemented if respondents answered any of the general needle use questions as "yes" because affirmative answers to these questions would indicate the misuse of drugs with a needle.

7.3 Substance Dependence and Abuse

The substance dependence and abuse section asked about symptoms of dependence or abuse in the past 12 months that were associated with the use of alcohol, marijuana, cocaine (including crack), heroin, hallucinogens, inhalants, or methamphetamine in the past 12 months or the misuse of prescription psychotherapeutic drugs (i.e., pain relievers, tranquilizers, stimulants, and sedatives) in that period. Lifetime users of cocaine (including crack), heroin, and methamphetamine who did not report use in the past 12 months in the relevant substance use section also were asked questions about dependence or abuse for the relevant drug if they

reported the use of these drugs in the past 12 months in the special drugs section of the interview (Sections 7.2.5).

In addition, the substance dependence and abuse section included items to assess for dependence on cigarettes if respondents reported use of cigarettes in the past 30 days. Details on how respondents were classified as having dependence or abuse for alcohol or illicit drugs (i.e., not including cigarettes) are found in Section B.4.3 in the 2016 NSDUH methodological summary report (Center for Behavioral Health Statistics and Quality, 2017f).

7.3.1 Editing of Substance Dependence and Abuse Variables

The editing procedures in place for substance dependence and abuse variables are first described in this section for respondents who were nonusers in the period of interest (Section 7.3.1.1). Section 7.3.1.2 describes the edits applied to the questionnaire items that were used to develop the composite score for determining cigarette dependence. Section 7.3.1.3 describes the assignment of a legitimate skip code for alcohol through sedative dependence or abuse variables. Finally, Sections 7.3.1.4 and 7.3.1.5 address the creation of dependence and abuse indicators, respectively, for the modPMN drugs that were listed in Section 7.1.

7.3.1.1 Editing of Substance Dependence and Abuse Variables for Nonusers in the Period of Interest

An important aspect of the processing of variables in this section consisted of assigning codes of 91 or 93 (Section 2.4.2) to variables that had been skipped because respondents never used the substance or they used it but not in the period of interest (i.e., more than 30 days ago for cigarettes or more than 12 months ago for other substances).¹³⁵ For alcohol and marijuana, respondents who had used these substances in the past 12 months also were skipped out of the corresponding dependence and abuse questions if they were only infrequent users of these two drugs in the past 12 months. If recency-of-use variables for the misuse of psychotherapeutic drugs were assigned a code of 83 (i.e., More than 12 months ago LOGICALLY ASSIGNED) because the only drugs that they misused in the past 12 months were OTC drugs (Section 6.3.4.2), then any data in the substance dependence and abuse section for that psychotherapeutic drug were overwritten with a code of 83. For cocaine, heroin, and methamphetamine, however, respondents' answers in the substance dependence and abuse section were retained if they were routed into the respective set of substance dependence and abuse questions because they reported past year use in the special drugs section (Section 7.2.5).

For alcohol and marijuana, the final, edited 12-month frequency variables (ALCYRTOT and MJYRTOT, respectively) also were used in assigning codes of 93 or 83 to the substance dependence and abuse variables pertaining to these substances. For example, if the edited variable ALCYRTOT indicated that respondents had used alcohol in the past 12 months but on

¹³⁵ For cocaine, heroin, and methamphetamine, respondents were not asked the corresponding questions in the substance dependence and abuse section if there was no indication of use in the past 12 months either in the relevant earlier substance use section (or sections, in the case of cocaine and crack) or in respondents' answers from the special drugs section. As noted in the main text, however, respondents who did not indicate past year use of cocaine, heroin, or methamphetamine in the corresponding earlier substance use sections but indicated past year use in the special drugs section *were* routed by the CAI instrument into the relevant drug dependence or abuse questions.

fewer than 6 days in that period, the edited substance dependence and abuse variables for alcohol were assigned a code of 93 if they had been skipped. If respondents answered one or more dependence or abuse questions for alcohol but the final value for ALCYRTOT indicated that they had used alcohol on fewer than 6 days in the past 12 months, the previous answers in the dependence and abuse questions were overwritten with a code of 83. Similar edits were done for marijuana if MJYRTOT indicated that respondents used marijuana on fewer than 6 days in the past 12 months.

7.3.1.2 Editing of Cigarette Dependence Variables

The questionnaire items that were used to determine cigarette dependence (also known as "nicotine dependence") in the 2016 NSDUH have been included in the survey since 2001. The method for determining dependence involved calculating a continuous score from the Nicotine Dependence Syndrome Scale (NDSS) (Shiffman, Hickcox, Gnys, Paty, & Kassel, 1995; Shiffman, Waters, & Hickcox, 2003). The score was calculated from the 17 questions shown in [Table 7.1](#) that were asked of respondents who used cigarettes in the past 30 days. For each of these items, respondents who had used cigarettes in the past 30 days were asked to choose an answer from the following scale: 1 = Not at all true; 2 = Somewhat true; 3 = Moderately true; 4 = Very true; or 5 = Extremely true. For details on how the estimate for cigarette dependence was calculated based on the NDSS, see Section B.4.2 in Appendix B of the 2009 summary of national findings (Office of Applied Studies, 2010a; 2010b).

As for other variables in the substance dependence and abuse section, editing of the cigarette dependence variables also involved assignment of a legitimate skip code of 99 according to the skip logic within this section. For example, question DRCGE06A (not shown in [Table 7.1](#)) asked if respondents who reported cigarette use in the past 30 days had any friends who did *not* smoke cigarettes. If DRCGE06A was answered as "no," question DRCGE06B was skipped (i.e., choosing at times not to be around friends who do not smoke because they do not like it when the respondent smokes). The edited variable CIGFNLKE corresponding to DRCGE06B was assigned a code of 99 if DRCGE06A was answered as "no."¹³⁶

No additional editing was done to the cigarette dependence variables if respondents reported cigarette use in the past 30 days or they used cigarettes in the past 30 days but skipped out of questions that did not apply. In particular, no editing was done when respondents entered the same response for all items (e.g., keying a "1" to every item). If respondents entered the same response to all cigarette dependence items, however, that would strongly suggest that they were not paying careful attention to the questions. For example, if a respondent chose "1" (i.e., not at all true) in question DRCGE03 ([Table 7.1](#)), the respondent would be expected to choose a response other than "1" to indicate some degree of truth to the statement in question DRCGE04 about having some sense of control over smoking. Nevertheless, when respondents entered the same response to all cigarette dependence items, these data were retained in order to allow analysts to decide how they wanted to handle these cases.

¹³⁶ Consistent with the principles in Section 2.4.2, CIGFNLKE retained missing values if DRCGE06B was skipped because respondents answered DRCGE06a as "don't know" or "refused."

Table 7.1 Mapping of Cigarette Dependence Questions to Edited Variables

Question	Edited Variable	Question Text
DRCGE01	CIGIRTBL	After not smoking for a while, you need to smoke in order to feel less restless and irritable.
DRCGE02	CIGCRAVE	When you don't smoke for a few hours, you start to crave cigarettes.
DRCGE03	CIGCRAGP	You sometimes have strong cravings for a cigarette where it feels like you're in the grip of a force you can't control.
DRCGE04	CIGINCTL	You feel a sense of control over your smoking – that is, you can "take it or leave it" at any time.
DRCGE05	CIGAVOID	You tend to avoid places that don't allow smoking, even if you would otherwise enjoy them.
DRCGE07	CIGPLANE	Even if you're traveling a long distance, you'd rather not travel by airplane because you wouldn't be allowed to smoke.
DRCGE08	CIGRNOUT	You sometimes worry that you will run out of cigarettes.
DRCGE09	CIGREGDY	You smoke cigarettes fairly regularly throughout the day.
DRCGE10	CIGREGWK	You smoke about the same amount on weekends as on weekdays.
DRCGE11	CIGREGNM	You smoke just about the same number of cigarettes from day to day.
DRCGE12	CIGNMCHG	It's hard to say how many cigarettes you smoke per day because the number often changes.
DRCGE13	CIGSVLHR	It's normal for you to smoke several cigarettes in an hour, then not have another one until hours later.
DRCGE14	CIGINFLU	The number of cigarettes you smoke per day is often influenced by other things – how you're feeling or what you're doing, for example.
DRCGE15	CIGNOINF	Your smoking is not affected much by other things. For example, you smoke about the same amount whether you're relaxing or working, happy or sad, alone or with others.
DRCGE16	CIGINCRS	Since you started smoking, the amount you smoke has increased.
DRCGE17	CIGSATIS	Compared to when you first started smoking, you need to smoke a lot more now in order to be satisfied.
DRCGE18	CIGLOTMR	Compared to when you first started smoking, you can smoke much, much more now before you start to feel anything.

7.3.1.3 Editing of Substance Dependence and Abuse Variables for Alcohol through Sedatives

In addition to the assignment of codes of 91 or 93 for lifetime nonusers or respondents who last used a substance more than 12 months prior to the interview, an important aspect of the processing of the substance dependence and abuse variables for alcohol through sedatives involved assignment of a legitimate skip code of 99 (Section 2.4.2); this code was assigned when respondents qualified for being asked dependence or abuse questions about a given substance but they legitimately skipped out of one more questions about that substance. Aside from assignment of codes of 91, 93, or 99, minimal additional editing was done to the substance dependence and abuse variables.

For example, the symptom of tolerance to the effects of alcohol was measured through two related questions, DRALC06 ("During the past 12 months, did you need to drink more alcohol than you used to in order to get the effect you wanted?") and DRALC07 ("During the past 12 months, did you notice that drinking the same amount of alcohol had less effect on you than it used to?"). An affirmative answer to either question would indicate tolerance. Thus, if respondents had already answered DRALC06 as "yes," there was no need to ask DRALC07. If the edited variable corresponding to question DRALC06 (ALCNDMOR) was coded as 1 (i.e., "yes"), the edited variable corresponding to question DRALC07 (ALCLSEFX) was assigned a legitimate skip code.

During preliminary review of the 2016 NSDUH data, an error was identified in the code that created the 2015 edited variable TRQLLSEFX (whether the same amount of tranquilizers had less effect than before). Results of an impact assessment using corrected data for TRQLLSEFX indicated that the changes to TRQLLSEFX would have a minimal effect on published estimates for tranquilizer use disorders and other substance use disorders (e.g., psychotherapeutic drug use disorders). A corrected version of this variable, named TRQLLSEFT, has been created for the 2015 NSDUH, but imputed or recoded substance use disorder variables were not revised for 2015.¹³⁷ TRQLLSEFT continued to be created for 2016, and imputation or recoding procedures for creating substance use disorder variables used TRQLLSEFT.

7.3.1.4 Creation of Dependence Indicators for the modPMN Drugs

This section describes how the edited dependence indicators were created for the modified predictive mean neighborhood (modPMN) drugs. Appendix D of the 2016 analytic codebook (Center for Behavioral Health Statistics and Quality, 2017b) describes how the edited dependence indicators were created for other drugs. The method for modPMN drugs is similar to the method for zero-fill drugs and involves the following three steps:

1. Membership in the domain is determined. Domain nonmembers receive a skip code for the edited dependence indicator. Generally, the domain for prescription drugs includes past year misusers of the drugs, and the domain for other modPMN drugs includes past year users of the drugs. This step is covered in Section 7.3.1.4.1.
2. The item-level variables are recoded to criterion-level variables, where the criteria are those defined by the *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition (DSM-IV; American Psychiatric Association, 1994). For some of these drugs, 9 item-level variables are recoded to six criteria, and for the others, 10 item-level variables are recoded to six criteria. This step is covered in Section 7.3.1.4.2.
3. Finally, the criterion-level variables are recoded to create dependence indicators, where again this process is defined by DSM-IV. Respondents with three or more positive criteria are defined with dependence. This step is covered in Section 7.3.1.4.3.

¹³⁷ The uncorrected variable TRQLLSEFX also remains in the 2015 data.

The step where the item-level variables are recoded to form the criterion-level variables is identical for both modPMN drugs and zero-fill drugs. However, for the other two steps, there are important differences between the method used for the modPMN drugs and the method used for the zero-fill drugs. These differences are discussed in Sections 7.3.1.4.1 and 7.3.1.4.3.

Table 7.2 is intended to serve as a quick reference guide to the variables used to create the edited dependence indicators. The nine item-level variables used for all drugs are xxxxLOTTM, xxxxGTOVR, xxxxKPLMT, xxxxNDMOR, xxxxLSEFX, xxxxCUTEV, xxxxEMCTD, xxxxPHCTD, and xxxxLSACT, where xxxx is the four-letter abbreviation associated with the drug. The item-level variable used for drugs that include withdrawal symptoms is called xxxxWDSMT. Details appear in Section 7.3.1.4.2.

Table 7.2 Quick Reference Guide to Creation of Edited Dependence Indicators

modPMN Drug	Domain Variable(s)	Item-Level Variables	Edited Dependence Indicator
Other modPMN Drugs			
Hallucinogens	HALLUCYR	Does not include withdrawal symptoms	DEPENDHAL
Inhalants	INHALYR	Does not include withdrawal symptoms	DEPENDINH
Methamphetamine	METHAMYR, METHNDLRC	Includes withdrawal symptoms	DEPENDMTH
Prescription Drugs			
Pain Relievers	PNRNMYR	Includes withdrawal symptoms	DEPENDPNR
Tranquilizers	TRQNMYR	Does not include withdrawal symptoms	DEPENDTRQ
Stimulants	STMNMYR	Includes withdrawal symptoms	DEPENDSTM
Sedatives	SEDNMYR	Includes withdrawal symptoms	DEPENDSED

modPMN = modified predictive mean neighborhood.

7.3.1.4.1 Domain Variables for Dependence Indicators for modPMN Drugs

Generally, membership in the domain for substance dependence is determined by the imputation-revised recency of use for other modPMN drugs and by the imputation-revised recency of misuse for the prescription drugs. A unit respondent is in the domain if, and only if, he or she is a past year (mis)user of the drug. Those who are not in the domain receive a skip code for the edited dependence indicators. Those who never (mis)used the drug receive a code of 91, and those who have (mis)used the drug in their lifetime but not in the past year receive a code of 93.

For all modPMN drugs except methamphetamine, the variable used to determine the domain status is simply the imputation-revised recency of (mis)use variable. The creation of these variables is described in Chapter 6.

For methamphetamine, the respondent is in the domain if, and only if, at least one of the following is true:

- IRMETHAMREC is 1 or 2 (i.e., the respondent is a past year methamphetamine user according to the imputation-revised recency variable).
- METHNDLRC is 1 or 2 (i.e., the respondent used a needle to inject methamphetamine in the past year).

METHNDLRC, described in Section 7.2.5.1, does not undergo any formal imputation. This implies that the domain variable is technically missing if IRMETHAMREC has a value other than 1 or 2 and METHNDLRC is missing. For the sake of simplicity, and because these cases are rare, these cases undergo a sort of mode imputation for the purpose of determining the domain status for dependence; that is, they are assumed to be out of the domain.

This approach to the domain indicator differs slightly from the approach taken for zero-fill drugs. For zero-fill drugs, there is no domain variable associated with the dependence variables. No distinction is drawn between those who are not in the domain and those who are in the domain but were not defined with dependence. However, the same domain information is programmed into the survey instrument; that is, those known to be past year users based on the raw versions of the domain variables described above are asked the dependence questions.

7.3.1.4.2 Creating the Criterion-Level Dependence Variables from the Item-Level Dependence Variables

For modPMN drugs that do not include a withdrawal criterion (hallucinogens, inhalants, and tranquilizers), there are nine item-level variables that are recoded into six criteria. The creation of the edited item-level variables is described briefly in Sections 7.3.1.1 and 7.3.1.3. The six criteria are as follows, where xxxx is the four-letter abbreviation for the drug (HALU, INHL, or TRQL):

1. Spent a great deal of time over a period of a month getting, using, or getting over the effects of the substance (xxxxLOTTM = Yes or xxxxGTOVR = Yes).
2. Unable to keep set limits on substance use or used more often than intended (xxxxKPLMT = Yes).
3. Needed to use substance more than before to get desired effects or noticed that using the same amount had less effect than before (xxxxNDMOR = Yes or xxxxLSEFX = Yes).
4. Able to cut down or stop using the substance every time he or she tried or wanted to (xxxxCUTEV = No).
5. Continued to use substance even though it was causing emotional, nervous, mental health, or physical problems (xxxxEMCTD = Yes or xxxxPHCTD = Yes).
6. Reduced or gave up participation in important activities due to substance use (xxxxLSACT = Yes).

An additional question pertaining to withdrawal symptoms was asked for the four modPMN drugs that include a withdrawal criterion (methamphetamine, pain relievers, stimulants, and sedatives). The withdrawal question asked the respondent if he or she had experienced substance-specific withdrawal symptoms at one time that lasted for longer than a

day after he or she cut back or stopped using. The specific number and type of listed withdrawal symptoms varied by substance. The edited variables are named xxxxWDSMT, where xxxx is the four-letter abbreviation for the drug (METH, PNRL, STIM, or SEDV).

There are no differences between the handling of the zero-fill drugs and the modPMN drugs for the step where the criterion-level dependence variables are created.

7.3.1.4.3 Creating the Edited Dependence Indicators from the Criterion-Level Dependence Variables

A respondent is defined with dependence if, and only if, he or she is in the domain and has at least three positive criteria. Therefore, if p is the number of positive dependence criteria and m is the number of missing dependence criteria, then the following can occur:

- The domain member is defined with dependence if, and only if, $p \geq 3$.
- The domain member is defined as not having dependence if, and only if, $p + m < 3$.
- All other domain members have a missing value for the dependence indicator.

It follows that domain members have a missing value for the edited dependence indicator in the following scenarios:

- The domain member has no positive criteria and is missing three or more criteria.
- The domain member has one positive criterion and is missing two or more criteria.
- The domain member has two positive criteria and is missing one or more criteria.

The handling of the zero-fill drugs is different in that all respondents with $p < 3$ are defined as not having dependence. This is equivalent to replacing all missing criteria with negative values.

The edited dependence indicators have the names DEPENDxxx, where xxx is the three-letter abbreviation associated with the drug (HAL, INH, MTH, PNR, TRQ, STM, or SED).

7.3.1.5 Creation of Abuse Indicators for modPMN Drugs

This section describes how the edited abuse indicators were created for the modPMN drugs. Appendix D of the 2016 analytic codebook (Center for Behavioral Health Statistics and Quality, 2017b) describes how the edited abuse indicators were created for zero-fill drugs. The method for modPMN drugs is similar to the method for predictive mean neighborhood (PMN) drugs. As for the edited dependence indicators (Section 7.3.1.4), both methods involve the following three steps:

1. Membership in the domain is determined. Domain nonmembers receive a skip code for the edited abuse indicator. The domain for the abuse variables discussed in this report is exactly the same as the domain for the dependence variables, but this is not true for the zero-fill drugs covered in Appendix D of the 2016 analytic codebook

(Center for Behavioral Health Statistics and Quality, 2017b). This step is covered in Section 7.3.1.5.1.

2. The item-level variables, described in Section 7.3.1.3, are recoded to criterion-level variables, where the criteria are those defined by the DSM-IV (American Psychiatric Association, 1994). For all modPMN drugs, five item-level variables are recoded to four criteria. This step is covered in Section 7.3.1.5.2.
3. Finally, the criterion-level variables are recoded to create abuse indicators, where again this process is defined by DSM-IV. Respondents with one or more positive criteria are defined with abuse. This step is covered in Section 7.3.1.5.3.

The step where the item-level variables are recoded to form the criterion-level variables is identical for both modPMN and zero-fill drugs. However, for the other two steps, there are important differences between the method used for the modPMN drugs and the zero-fill drugs. These differences are discussed in Sections 7.3.1.5.1 and 7.3.1.5.3.

7.3.1.5.1 Domain Variables for Abuse Indicators for modPMN Drugs

For the abuse indicators discussed here, the domain for the abuse variables is exactly the same as the domain for the dependence variables (Section 7.3.1.4.1). However, this is not true for the abuse indicators for zero-fill drugs that are discussed in Appendix D of the 2016 analytic codebook (Center for Behavioral Health Statistics and Quality, 2017b). The two major differences are the following:

- As for the dependence variables, no distinction is drawn between those who are not in the domain and those who are in the domain but were not defined with abuse, but, again, the survey instrument determines the domain based on responses to questions about past year (mis)use of the drug.
- Respondents defined with dependence are automatically defined as not having abuse. Thus, those respondents defined with dependence can be thought of as being nonmembers of the domain for abuse. This process is defined by DSM-IV.

The abuse indicators covered here are independent of dependence, and they are intended to provide analysts with the capacity to analyze abuse independently of dependence if they choose.

7.3.1.5.2 Creating the Criterion-Level Abuse Variables from the Item-Level Abuse Variables

The five item-level abuse variables used to create the four abuse criteria are xxxxSERPB, xxxxPDANG, xxxxLAWTR, xxxxFMFPB, and xxxxFMCTD, where xxxx is the four-letter abbreviation associated with the drug (HALU, INHL, METH, PNRL, TRQL, STIM, or SEDV). The four criteria are the following:

1. Respondent reported having serious problems due to substance use at home, work, or school (xxxxSERPB = Yes).

2. Respondent reported using substance regularly and then did something where substance use might have put him or her in physical danger (xxxxPDANG = Yes).
3. Respondent reported substance use was causing actions that repeatedly got him or her in trouble with the law (xxxxLAWTR = Yes).
4. Respondent reported having problems caused by substance use with family or friends (xxxxFMFPB = 1) and continued to use substance even though it was thought to be causing problems with family and friends (xxxxFMCTD = Yes).

There are no differences between the handling of the zero-fill drugs and the modPMN drugs for the step where the criterion-level abuse variables are created.

7.3.1.5.3 Creating the Edited Abuse Indicators from the Criterion-Level Abuse Variables

A respondent is defined with abuse if, and only if, he or she is in the domain and has at least one positive criterion. Therefore, if p is the number of positive criteria and m is the number of missing criteria, the following can occur:

- The domain member is defined with abuse if, and only if, $p \geq 1$.
- The domain member is defined as not having abuse if, and only if, $p = 0$ and $m = 0$.
- The domain member has a missing value for the abuse indicator if, and only if, $p = 0$ and $m \geq 1$.

The handling of the zero-fill drugs is different in that all respondents who are either defined with dependence or have $p = 0$ are defined as not having abuse. As for dependence, this is equivalent to replacing all missing abuse criteria with negative values.

Tables 7.3 and 7.4 can be used to compare the values of the abuse indicators for zero-fill drugs with the values of the abuse indicators for the modPMN drugs. Note the following:

- For zero-fill drugs, neither the dependence indicator nor the abuse indicator is ever missing. Missing criteria are simply assumed to be negative.
- For zero-fill drugs, all respondents defined with dependence are automatically defined as not having abuse. The values of the abuse criteria are not relevant.
- For modPMN drugs, both the dependence and abuse indicators may be missing, and the value of the abuse indicator is independent of the value of the dependence indicator.

The edited abuse indicators have the names ABUPOSxxx, where xxx is the three-letter abbreviation associated with the drug (HAL, INH, MTH, PNR, TRQ, STM, or SED). The variables are so named to remind analysts that these edited abuse variables are unrelated to the dependence indicators and therefore do not follow the DSM-IV definition; they simply indicate whether the respondent was "positive" (POS) for abuse (ABU).

Table 7.3 Value of Abuse as Derived from Dependence and the Abuse Criteria, Zero-Fill Drugs

Abuse Criteria	Dependence	
	Yes	No
Met at least one	No	Yes
Met none, and none were missing	No	No
Met none, and at least one was missing	No	No

PMN = predictive mean neighborhood.

Table 7.4 Value of Edited Abuse as Derived from Edited Dependence and the Edited Abuse Criteria, modPMN Drugs

Abuse Criteria	Dependence		
	Yes	No	Missing
Met at least one	Yes	Yes	Yes
Met none, and none were missing	No	No	No
Met none, and at least one was missing	Missing	Missing	Missing

modPMN = modified predictive mean neighborhood.

7.3.2 Imputation of Cigarette Dependence and Substance Dependence and Abuse Variables

The imputation procedures performed on the substance dependence and abuse variables is addressed in three sections. Section 7.3.2.1 describes the imputation process for questionnaire items that comprise the NDSS composite score used to determine cigarette dependence. Section 7.3.2.2 provides details on the cigarette dependence variables, including information about respondent eligibility criteria for answering these questions and the reasons why they were classified as eligible or not eligible to be imputed. Finally, Section 7.3.2.3 describes how the imputation-revised substance dependence and abuse indicators for the modPMN drugs were created from their edited predecessors.

7.3.2.1 Imputation of Cigarette Dependence Variables

Unlike the majority of variables that are imputed for NSDUH, imputation of the cigarette dependence variables was not performed using the PMN or modPMN methodology. For respondents who used cigarettes in the past 30 days and provided complete data for all 17 of the dependence questions that were used to calculate the NDSS scale value, imputation-revised cigarette dependence variables were simply assigned the values from the corresponding edited variables.

For respondents who used cigarettes in the past 30 days and gave a valid response to 16 of the 17 NDSS items, the predicted mean for the one missing item was obtained using the coefficients corresponding to the other 16 nonmissing covariates from the appropriate weighted least squares regression. For example, if CIGIRTBL was the variable whose missing value was to be imputed, CIGIRTBL would be specified as the dependent variable in the model, and the remaining 16 NDSS variables served as the covariates: CIGCRAVE, CIGCRAGP, CIGINCTL, CIGAVOID, CIGPLANE, CIGRNOUT, CIGREGDY, CIGREGWK, CIGREGNM,

CIGNMCHG, CIGSVLHR, CIGINFLU, CIGNOINF, CIGINCRS, CIGSATIS, and CIGLOTMR. The imputation-revised variable was then set to the predicted mean.¹³⁸

Respondents who used cigarettes in the past 30 days but answered 15 or fewer of the cigarette dependence questions were left out of the modeling process entirely. The missing values for these respondents remained missing in the imputation-revised variables that were derived from the edited NDSS variables.

Across all respondents (regardless of how many of the NDSS questions they answered), no response propensity adjustments were performed for the item respondent weights used in the regression-based imputation models. The analysis weights described in Section 3.3.1 were applied instead.

The NDSS mean value was calculated from imputation-revised versions of the 17 nicotine dependence questionnaire variables. The NDSS mean value was set to missing for respondents who used cigarettes in the past 30 days but answered 15 or fewer of the cigarette dependence questions.

7.3.2.2 Summary Information for Cigarette Dependence Variables

In the 2016 NSDUH, 125 cases had valid values for 16 of the 17 questions. These cases were imputed. In 2016, 126 cases had fewer than 16 valid values for the NDSS scale. For these cases, no imputations were performed. Note that it was possible that the respondent was eligible to answer the questions about cigarette dependence because he or she was imputed to be a past month cigarette user. This situation occurred nine times in the 2016 NSDUH.

[Table 7.5](#) summarizes the eligibility of respondents to answer the cigarette dependence questions and reasons why respondents were classified as eligible or not eligible to be imputed. Furthermore, this table provides details about the amount of cigarette dependence data that were missing for eligible respondents. It also provides information on whether the respondent was imputed to be a past month cigarette user. Consequently, the respondent would be eligible to have cigarette dependence data but would have missing data for all the cigarette dependence variables.

Of the 251 eligible respondents in 2016 who did not answer all 17 NDSS questions, half (125 cases, 49.8 percent of the eligible cases with incomplete data) were missing a response for only one of the items ([Table 7.5](#)). Any respondent with more than one of the 17 items missing (126 cases, 50.2 percent of the eligible cases with incomplete data) did not have his or her missing responses replaced with imputed values, and no cigarette dependence score was calculated for those respondents. For the respondents missing only one response, imputation was used to fill in the values for the missing variable, using information from the other 16 nonmissing variables, through weighted least squares regression models. This resulted in 17 regression-based imputation models, where the response variable for each model was the edited

¹³⁸ Because the response variable and covariates were treated as continuous in the models, it is possible for a predicted mean, and therefore an imputation-revised value, to exceed five or be less than one.

variable that corresponded to each item in the NDSS, and the covariates in each model were the remaining NDSS variables.

Table 7.5 Summary of Response Patterns for 2016 NDSS Variables

Number of Valid NDSS Variables	Past Month Smoker	Past Month Smoker Status Imputed	Eligible to Answer NDSS Questions	NDSS Variables Imputed	Frequency	Percentage of Total
N/A	No	No	No	N/A ¹	55,240	81.30
N/A	No	Yes	No	N/A ¹	9	0.01
Subtotal					55,249	81.32
0	Yes	No	Yes	No	18	0.03
0	Yes	Yes	Yes	No	9	0.01
1-15	Yes	No	Yes	No	99	0.15
Subtotal					126	0.19
16	Yes	No	Yes	Yes	125	0.18
17	Yes	No	Yes	N/A ²	12,442	18.31
Total					67,942	100.00

N/A = not applicable; NDSS = Nicotine Dependence Syndrome Scale.

¹ The respondent was not eligible to answer the NDSS questions.

² None of the NDSS variables were missing for this scenario.

7.3.2.3 Imputation of Dependence and Abuse Indicators for modPMN Drugs

This section describes how the imputation-revised substance dependence and abuse indicators for modPMN drugs were created from the edited versions of the same variables. The creation of the edited variables is described in Sections 7.3.1.4 and 7.3.1.5. These variables underwent imputation using the modPMN method, which is described in Sections 3.5, 3.6, and 3.7. The dependence and abuse variables for zero-fill drugs did not undergo formal imputation; instead, missing values were replaced by negative responses as described in Appendix D of the 2016 analytic codebook (Center for Behavioral Health Statistics and Quality, 2017b).

There are seven modPMN drugs that have both dependence and abuse indicators, so a total of 14 imputation-revised variables were created. Section 7.3.2.3.1 describes how these 14 variables were segregated into imputation sets. Section 7.3.2.3.2 describes how the variables were ordered within the imputation sets. The processing order is important because the variables later in the sequence can in some sense be said to have undergone a more precise imputation than variables earlier in the sequence (see Section 3.7.1). Sections 7.3.2.3.3, 7.3.2.3.4, and 7.3.2.3.5 provide details on the three steps of modPMN as described in Sections 3.6.1, 3.6.2, and 3.6.3. Section 7.3.2.3.6 uses flowcharts to show how the three steps were arranged to complete imputations. These flowcharts follow the general modPMN flowcharts presented in [Figures 3.6](#) and [3.7](#). Finally, Section 7.3.2.3.7 describes the 28 variables created by these procedures—the 14 imputation-revised dependence and abuse indicators and their imputation indicators—and also mentions some of the recoded variables created using these variables.

7.3.2.3.1 Dependence and Abuse Indicators for modPMN Drugs: Imputation Sets

Typically, closely related variables are placed in the same imputation set. The dependence and abuse indicators for the same drug were considered to be closely related; methamphetamine and prescription stimulants were considered to be closely related (because methamphetamine is a stimulant); and tranquilizers and sedatives were considered to be closely related. This led to the formation of the five imputation sets shown in Table 7.6. Although these sets are numbered in the table, they can be run in any order (or in parallel), because none of the imputation-revised variables is used as a covariate for any other variable outside of the imputation set.

Table 7.6 Segregation of Dependence and Abuse Indicators for modPMN Drugs into Imputation Sets

modPMN Drug	Dependence	Abuse
Hallucinogens	Set 1	
Inhalants		
Methamphetamine	Set 3	
Prescription Stimulants		
Prescription Pain Relievers	Set 4	
Prescription Tranquilizers	Set 5	
Prescription Sedatives		

modPMN = modified predictive mean neighborhood.

7.3.2.3.2 Ordering of Variables within Imputation Sets

As explained in Section 3.7.1, the sequencing of variables in modPMN within an imputation set is by descending item response rate. The ordering in 2016 was as shown in Table 7.7.

Table 7.7 Sequencing of Dependence and Abuse Indicators within Imputation Sets, 2016

modPMN Drug	Dependence	Abuse
Hallucinogens	Set 1, Variable 2	Set 1, Variable 1
Inhalants	Set 2, Variable 2	Set 2, Variable 1
Methamphetamine	Set 3, Variable 4	Set 3, Variable 3
Prescription Stimulants	Set 3, Variable 2	Set 3, Variable 1
Prescription Pain Relievers	Set 4, Variable 2	Set 4, Variable 1
Prescription Tranquilizers	Set 5, Variable 4	Set 5, Variable 3
Prescription Sedatives	Set 5, Variable 2	Set 5, Variable 1

modPMN = modified predictive mean neighborhood.

7.3.2.3.3 Step 1: Response Propensity Adjustment

This section provides specific information about the response propensity steps for the five dependence and abuse imputation sets. General information about the response propensity adjustment step of modPMN is available in Section 3.6.1. There were a total of 28 response

propensity models fit across these five imputation sets.¹³⁹ As laid out in Section 3.3.1, there are four key inputs to the response propensity adjustment step, whether PMN or modPMN is the imputation method: the analysis weight, the domain indicator, the item response indicator, and the covariate list. The four inputs for the five dependence and abuse imputation sets are described below.

- *Analysis Weight*: For all five imputation sets for dependence and abuse, the input to all response propensity models was the preliminary analysis weight, PANALWT.
- *Domain*: This is described in Section 7.3.1.4.1. The domain includes past year (mis)users based mainly on the imputation-revised recency of (mis)use variables.
- *Item Response Indicator*: This was based simply on whether the edited variable in the question was missing. The process by which this is determined is described in Section 7.3.1.4.3 for the dependence indicators and in Section 7.3.1.5.3 for the abuse indicators.
- *Covariates*: [Table 7.8](#) lists the tables in Appendix D that show both the starting and the final covariates used in each of the 28 response propensity models. Each table includes the covariate lists for both cycles.

Table 7.8 Tables Reporting Covariate Lists for Response Propensity Models, Dependence and Abuse Indicators

modPMN Drug	Dependence	Abuse
Hallucinogens	Table D.54 (4 models)	
Inhalants	Table D.56 (4 models)	
Methamphetamine	Table D.58 (8 models)	
Prescription Stimulants		
Prescription Pain Relievers	Table D.60 (4 models)	
Prescription Tranquilizers	Table D.62 (8 models)	
Prescription Sedatives		

modPMN = modified predictive mean neighborhood.

7.3.2.3.4 Step 2: Prediction Modeling

This section provides specific information about the prediction modeling step for the five dependence and abuse imputation sets. General information about the prediction modeling step of modPMN is provided in Section 3.6.2. There were a total of 140 prediction models¹⁴⁰ fit across these five imputation sets. Each of the 28 response propensity models segregated the domain members into five bins, and separate prediction models were fit within the bins.

For each of the 140 models, the probability of dependence (or abuse) was modeled using dichotomous logistic regression. SUDAAN's LOGISTIC procedure (called RLOGIST in SAS-callable SUDAAN) was used to estimate the regression coefficients and to calculate predicted probabilities for all item nonrespondents. These predicted probabilities were used in the

¹³⁹ There were seven modPMN drugs, two variables per drug (dependence and abuse), and two cycles, leading to $7 \times 2 \times 2 = 28$ response propensity models.

¹⁴⁰ Typically, not all 140 models are fit in a given year. If there are no item nonrespondents in a particular bin, the prediction modeling and collocated stochastic imputation steps are skipped.

subsequent step to impute a positive or negative response. The covariates used in the models are listed in Appendix D. [Table 7.9](#) lists the tables in Appendix D that show both the starting and the final covariates for each of the 140 models.

Table 7.9 Tables Reporting Covariate Lists for Prediction Models, Dependence and Abuse Indicators

modPMN Drug	Dependence	Abuse
Hallucinogens	Table D.55 (20 models)	
Inhalants	Table D.57 (20 models)	
Methamphetamine	Table D.59 (40 models)	
Prescription Stimulants		
Prescription Pain Relievers	Table D.61 (20 models)	
Prescription Tranquilizers	Table D.63 (40 models)	
Prescription Sedatives		

modPMN = modified predictive mean neighborhood.

7.3.2.3.5 Step 3: Collocated Stochastic Imputation

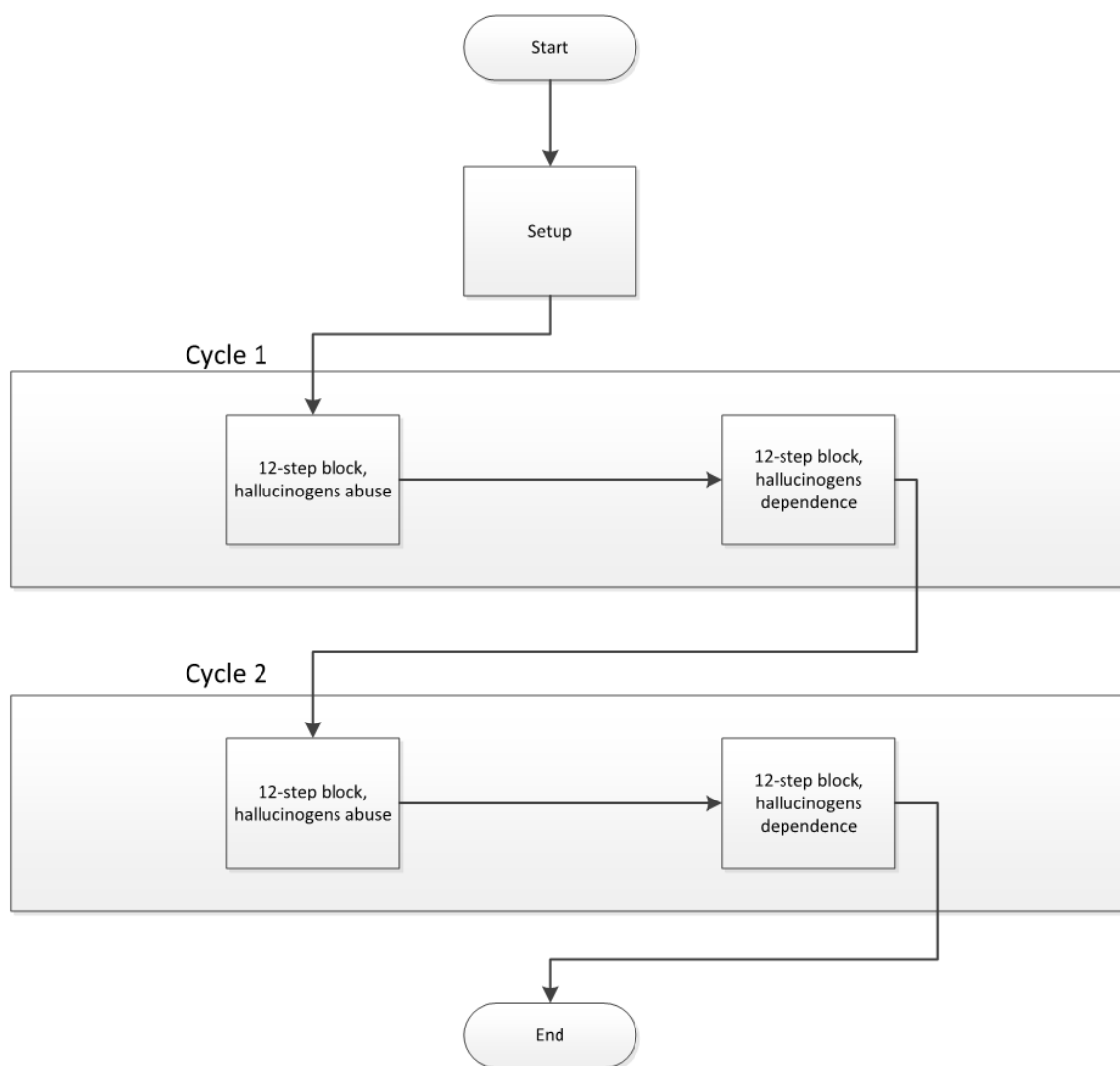
The collocated stochastic imputation step for the dependence and abuse indicators is straightforward and is described in general in Section 3.6.3. Briefly, each of the 140 prediction models is followed by an imputation step where a random number is used to impute a value of "Yes" or "No" for the given dependence or abuse indicator for each item nonrespondent. Imputation is done only at the dependence or abuse indicator level. This approach simplifies the procedures, but one shortcoming is that partial information stored in the item-level edited variables (Sections 7.3.1.1, 7.3.1.2, and 7.3.1.3) and the edited criterion-level variables (Sections 7.3.1.4.2 and 7.3.1.5.2) is ignored. For example, all else being equal, an item nonrespondent with two positive dependence criteria and four missing dependence criteria has the same probability of being imputed with dependence as an item nonrespondent with no positive dependence criteria and three missing dependence criteria. The numbers of cases undergoing collocated stochastic imputation for each of the 14 dependence and abuse indicators are provided in [Tables E.111](#) to [E.124](#).

7.3.2.3.6 Sequencing of Steps for Selected Dependence and Abuse Imputation Sets

This section provides the sequencing of steps for the five dependence and abuse imputation sets. General information about the sequencing of steps for modPMN is available in Sections 3.7.2 and 3.7.3.

Imputation sets 1, 2, and 4 include two variables, so there are four 12-step blocks of programs included (Section 3.7.2). The first of the four is for the first variable within the first cycle; the second is for the second variable within the first cycle; the third is for the first variable within the second cycle; and the fourth and last is for the second variable within the second cycle. Each 12-step block consists of a response propensity step where the domain members are segregated into five bins, followed by a prediction modeling step within each bin, followed by a collocated stochastic imputation step within each bin, and finally followed by a "wrap-up" program that merges the five datasets into one. [Figure 7.1](#) illustrates the sequencing of steps for imputation set 1.

Figure 7.1 Sequencing of Steps for Hallucinogens Dependence and Abuse



Imputation sets 3 and 5 include four variables instead of two. The algorithm is identical to the one used for imputation sets 1, 2, and 4, except that there are four 12-step blocks per cycle instead of two, for a total of eight 12-step blocks.

7.3.2.3.7 Final Dependence and Abuse Variables

Once all five imputation sets have been completed, 28 variables are assembled into a single dataset for delivery to the project's master dataset: the imputation-revised dependence (IRDEPENDxxx) and abuse (IRABUPOSxxx) indicators for the seven modPMN drugs along with their imputation indicators.¹⁴¹ For the 2015 survey, the imputation indicators (IIDPENDXXX and IABUPOSXXX) were assigned to two levels where level 1 mapped to "Questionnaire data or legitimate skip" and level 3 mapped to "Imputation-revised." Beginning in 2016, these imputation indicators were modified so that respondents who were not past year users/misusers of modPMN drugs received values of 9 (legitimate skip) instead of 1

¹⁴¹ See Section 3.1 for details about imputation indicators.

(questionnaire data). Values of 3 remained unchanged to denote imputed responses. The 2015 imputation indicators were also retroactively revised based on the new imputation indicator definitions and redelivered to the 2015 master file. Several frequently used recoded variables are also created from the imputation-revised dependence and abuse indicators, including use disorder variables and abuse indicators that follow the DSM-IV definitions. These recoded variables are described in Appendix D of the 2016 analytic codebook (Center for Behavioral Health Statistics and Quality, 2017b).

This page intentionally left blank

8. Editing of Additional Self-Administered Sections for Substance Use

8.1 Introduction

This chapter discusses editing procedures for the following additional self-administered sections that pertain to substance use:

- blunts,
- prior substance use,
- substance use treatment, and
- consumption of alcohol.

In contrast to variables for substance dependence and abuse that were described in Section 7.3, none of the edited variables in this chapter were imputed to remove missing values.

These interview sections are included in this chapter according to their related topical content rather than in the sequential order in which they appeared to respondents during the interview. The order in which these sections were presented to respondents relative to other self-administered sections of the interview can be found in the computer-assisted interviewing (CAI) specifications for the 2016 NSDUH that are available at <https://www.samhsa.gov/data/>.

8.2 Blunts

Since 2004, all respondents have been asked questions about smoking part or all of a cigar with marijuana in it, commonly referred to as a "blunt." Two new questions also were added to the blunts section in 2013 about use of marijuana that had been recommended by a doctor or other health care professional (subsequently referred to as medical marijuana use). Since 2014, the logic for asking the questions about medical marijuana use has included respondents who confirmed that they smoked blunts in the past 30 days despite reporting in the marijuana section that they last used marijuana more than 12 months ago.

8.2.1 Use of Blunts

Respondents first were asked if they ever smoked part or all of a cigar with marijuana in it, sometimes called a "blunt." If respondents were lifetime users of blunts, they were asked how long it had been since they had smoked a blunt. If respondents indicated blunt use in the past 30 days, they were asked to report the number of days they smoked blunts in the past 30 days. If respondents reported use of both cigars (from the tobacco section) and blunts in the past 30 days, they also were asked if they had smoked a cigar *without* marijuana in it during the past 30 days.

In addition, the section included consistency checks in selected situations where respondents gave reports of blunt use that were inconsistent with what they had reported in the marijuana section. Specifically, respondents could have reported earlier in the interview that they

never used marijuana or hashish and then indicate that they had smoked a blunt in the past 30 days. Similarly, respondents could have reported that their last use of marijuana was more than 30 days ago and then indicate use of blunts in the past 30 days. In these situations, respondents were asked why they had reported earlier that they had never used marijuana or hashish or that they had not used it in the past 30 days, respectively.

An important aspect of the processing of variables in this section pertaining to use of blunts consisted of assigning codes of 91 and 99 (Section 2.4.2) to variables that had been skipped because the questions did not apply. For example, if respondents said they never smoked part or all of a blunt, they were skipped out of the remaining blunts questions. Thus, respondents were assigned a code of 91 (NEVER USED BLUNTS) to the variables BLNTREC (most recent use of a blunt, corresponding to question BL02) and BLNT30DY (frequency of use of blunts in the past 30 days, corresponding to question BL02A) if they reported that they had never smoked blunts. Similarly, if respondents refused to report whether they had ever smoked part or all of a blunt, that refusal code was assigned to BLNTREC and BLNT30DY as well. In addition, if respondents reported being lifetime users of blunts but reported that they last used blunts more than 30 days ago, question BL02a would have been skipped. Therefore, BLNT30DY was assigned a code of 93 (DID NOT USE BLUNTS IN THE PAST 30 DAYS) in this situation.

Minimal editing was done to the variables BLNT30C1, BLNT30C2, RSNOMRJ, and RSNMRJMO, corresponding to questions BL03, BL04, BL05, and BL06, respectively; these variables pertained to the above-mentioned consistency checks between respondents' answers in the blunts and marijuana sections. If respondents reported not using blunts in the past month or reported that they were a lifetime marijuana user, questions BL03 and BL05 would have been skipped. Therefore, BLNT30C1 and RSNOMRJ were assigned a code of 99 in this situation. Similarly, if respondents were not past month users of blunts or reported last using marijuana in the past 30 days in the marijuana section (i.e., a response in the marijuana section that was consistent with reported use of blunts in the past 30 days), then questions BL04 and BL06 would have been skipped. Therefore, BLNT30C2 and RSNMRJMO were assigned a code of 99 in these situations.

If respondents were routed to question BL03 (because they reported past month use of blunts but previously reported never using marijuana) and then they refused to answer question BL03, a refusal was propagated from BLNT30C1 to RSNOMRJ. If respondents answered BL03 as "don't know," RSNOMRJ retained a code of 98. Similarly, if respondents were routed to BL04 (because they reported past month use of blunts but previously reported using marijuana but not in the past 30 days) and they refused to answer, a refusal was propagated from BLNT30C2 to RSNMRJMO. Respondents answering "don't know" to question BL04 retained a code of 98 in RSNMRJMO.

In addition, since 2005, respondents who were lifetime users of blunts have been asked how old they were when they first used blunts (BL08). If respondents reported first using blunts within 1 year of their current age, they were asked to report the specific year and month when they first used, with the allowable years ranging from 2014 to 2016. If respondents reported first using blunts at their current age, their last birthday was in the current year, and they were interviewed after their last birthday, the CAI program assumed that the first use of blunts occurred in the current year (i.e., 2016). These respondents were asked only for the month when

they first used in the current year. The remaining respondents who first used blunts within 1 year of their current age could be routed to one of two possible questions on the specific year they first used. They then were routed to a question to report on the specific month they first used blunts in the year they had reported previously.

Because the routing logic to the different versions of the MFU and YFU questions was mutually exclusive, composite sets of MFU and YFU variables (BLNTYFU and BLNTMFU, respectively) were created from the individual unedited variables. In addition, the final YFU variables were recoded to replace codes of 1 or 2 from the questionnaire that denoted specific years with values for the years (i.e., 2014 through 2016). In situations in which the CAI program skipped the YFU questions because it could be logically inferred that first use of blunts occurred in the interview year (i.e., 2016), a code of "2016" was assigned to BLNTYFU; this was done even if respondents did not know what month they first used in the current year, or if they refused to report what month they first used in the current year. If the MFU and YFU questions had been skipped because respondents first used blunts at ages that were more than 1 year below their current ages, legitimate skip codes were assigned to the BLNTYFU and BLNTMFU variables.

A consistency check also was included for situations in which the apparent age when respondents first used blunts that was calculated from their MFU and YFU and their date of birth (CAI variable MYR1STBL) was inconsistent with the age that respondents directly reported for when they first used blunts (CAI variable AGE1STBL). For example, the consistency check was triggered if a 16-year-old respondent reported first using blunts at age 16 but then reported first using blunts in a year and month that would have meant the respondent was 15 years old when he or she first used blunts. No further editing needed to be done if respondents indicated twice in a row that the AFU that was calculated from the MFU and YFU was correct. The CAI program updated the value for AGE1STBL (corresponding to the edited variable BLNTAGE) to agree with the values for the MFU and YFU.

A calculation of an AFU based on respondents' reported MFU and YFU of blunts was not done if respondents reported that they first used blunts in the same month as their birth dates; in this situation, a unique AFU could not be determined. Similarly, this consistency check was not triggered if respondents had missing data in either of the year or month questions, such as if respondents knew the year when they first used blunts but they did not know the MFU.

If respondents indicated at some point in the consistency check sequence that the value they had originally reported for their AFU was correct (and by extension, that the MFU and YFU was not correct), they had an opportunity to revise the values for their MFU and YFU. These revised reports for MFU and YFU were used in subsequent editing steps. Otherwise, respondents' answers to the original MFU and YFU questions were used in subsequent editing.

Table B.13 in Appendix B discusses edits pertaining to the consistency checks for these initiation variables for blunts. For consistency with how the substance use initiation data have been edited, the default when a respondent did not resolve an inconsistency between the AFU and the MFU and YFU was to favor the AFU in subsequent editing decisions. Unlike the initiation data in the substance use sections at the beginning of the interview, however, the initiation variables for blunts did not undergo subsequent statistical imputation.

Because these initiation variables were not present in 2004, editing of BLNTREC beyond 2004 did not take into account these initiation data for consistency with how BLNTREC was edited in 2004. Keeping the editing procedures comparable with those in 2004 would permit analysts to examine trends since 2004. However, a flag variable was created that indicated when initiation data suggested more recent use than what the respondent indicated in BLNTREC. The flag was coded as 98 (blank) if no inconsistency existed between BLNTREC and the initiation data.

[Table B.14](#) presents additional edits pertaining to variables in the blunts section. Many of these edits were developed when the section was added to the survey in 2004, particularly with regard to the editing of BLNTREC. For example, if respondents reported using blunts but they also reported never using marijuana, then question BL03 was triggered only when respondents reported using blunts in the past 30 days. Respondents who verified that they had never used marijuana would then have their BLNTEVER answer assigned to 4 (No LOGICALLY ASSIGNED) and BNLTRC assigned to 81 (NEVER USED BLUNTS, Logically assigned). [Table B.14](#) also discusses miscellaneous edits for the initiation variables that have been present since 2005, other than those described in [Table B.13](#) that were related to consistency checks.

No editing was done to make the data for blunts consistent with initiation data for cigars or marijuana from their respective earlier sections of the interview, or vice versa. Thus, for example, initiation data in the blunts section could indicate that respondents first smoked a cigar with marijuana in it at an earlier age than they reported for when they first used cigars or marijuana. However, variables in the blunts section of the National Survey on Drug Use and Health (NSDUH) codebooks included a standard footnote to indicate that the data in this section may be inconsistent with data from earlier sections of the interview.

8.2.2 Medical Marijuana Use

In the 2016 NSDUH, if respondents reported use of marijuana in the past 12 months in the marijuana section, or they reported smoking a blunt in the past 12 months, or they confirmed smoking a blunt in the past 30 days (i.e., despite reporting in the marijuana section that they last used marijuana more than 12 months ago), they were asked question MJMM01, regarding whether any of their marijuana use in the past 12 months was recommended by a doctor or other health care professional (i.e., medical marijuana use). If respondents reported any medical marijuana use in the past 12 months (MJMM01 = 1), they were asked question MJMM02 regarding whether *all* of their marijuana use in the past 12 months was recommended by a doctor or other health care professional.

An important aspect of the processing of variables pertaining to medical marijuana use consisted of assigning codes of 91, 93, and 99 (Section 2.4.2) to variables that had been skipped because the medical marijuana questions did not apply to respondents. For example, if respondents never used marijuana and they never smoked part or all of a blunt, they were assigned a code of 91 (NEVER USED MARIJUANA OR BLUNTS) to the variables MEDMJYR (corresponding to MJMM01) and MEDMJALL (corresponding to MJMM02). This edit also was implemented when MJMM01 and MJMM02 had been skipped if respondents reported in the marijuana section that they never used marijuana but data on lifetime use of blunts was unknown. Similarly, if MJMM01 and MJMM02 had been skipped because

respondents last used marijuana more than 12 months ago and there were no reports of use of blunts in the past 12 months, then MEDMJYR and MEDMJALL were assigned a code of 93 (DID NOT USE MARIJUANA OR BLUNTS IN THE PAST 30 DAYS). If respondents were past year users of marijuana or blunts but they answered question MJMM01 as "no," then MEDMJALL was assigned a code of 99 (LEGITIMATE SKIP).

If respondents in 2016 had not been asked MJMM01 and MJMM02, the respective edited variables MEDMJYR and MEDMJALL retained missing values (i.e., blank; see Section 2.2.1) in the following situations:

- The edited variable from the marijuana section for most recent use of marijuana (MJREC) had a missing value; or
- MJREC indicated use of marijuana at some point in the respondent's lifetime but the exact period of most recent use was unknown (MJREC = 9).

Respondents who previously reported marijuana use in the past 12 months could answer question MJMM01 as "yes" (MEDMJYR=1), even if they were living in a state where a law approving marijuana for medical use was not in effect at the time they were interviewed. [Table 8.1](#) lists the 29 states and the District of Columbia where a law approving marijuana for medical use was enacted by the end of 2016, along with the dates when the laws took effect in these states.

In order to better understand the quality of the medical marijuana data, two recoded variables (MEDMJSTA and MEDMJCST) were created that examined (1) reported medical marijuana use, (2) state of residence at the time of interview, (3) dates when medical marijuana laws took effect (where applicable), and (4) the interview dates.

The variable MEDMJSTA was created based on the information in [Table 8.1](#) on the status of state medical marijuana laws in 2016, along with information on the date of interview and the state where respondents were living on the date they were interviewed. Florida approved marijuana for medical use in November 2016, but the law was not scheduled to take effect until January 2017. Therefore, Florida was not included as a state that had a law in effect in 2016.

MEDMJSTA was created for *all* respondents in the 2016 NSDUH, regardless of whether they reported any lifetime or past year marijuana use or past year medical marijuana use. MEDMJSTA was coded as follows:

- MEDMJSTA = 1: The respondent was in a state where a law allowing the use of marijuana for medical reasons was in effect on or before the interview date.
- MEDMJSTA = 2: The respondent was in a state that did not have a law in effect for any period of 2016 that allowed the use of marijuana for medical reasons.
- MEDMJSTA = 3: The respondent was interviewed in a state before that state's medical marijuana law took effect.

Table 8.1 States Where Marijuana Was Approved for Medical Use as of December 31, 2016, and Dates When the Laws Took Effect

State	Date Enacted (Approved, Passed, or Signed into Law)	Effective Date
California	November 5, 1996	November 6, 1996
Alaska	November 3, 1998	March 4, 1999
Oregon	November 3, 1998	December 3, 1998
Washington	November 3, 1998	November 3, 1998
Maine	November 2, 1999	December 22, 1999
Hawaii	June 14, 2000	December 28, 2000
Colorado	November 7, 2000	June 1, 2001
Nevada	November 7, 2000	October 1, 2001
Vermont	May 26, 2004	July 1, 2004
Montana	November 2, 2004	November 2, 2004
Rhode Island	January 3, 2006	January 3, 2006
New Mexico	March 13, 2007	July 1, 2007
Michigan	November 4, 2008	December 4, 2008
New Jersey	January 18, 2010	July 1, 2010
District of Columbia	May 21, 2010	July 27, 2010
Arizona	November 2, 2010	November 2, 2010
Delaware	May 13, 2011	July 1, 2011
Connecticut	May 31, 2012	October 1, 2012
Massachusetts	November 6, 2012	January 1, 2013
New Hampshire	July 23, 2013	July 23, 2013
Illinois	August 1, 2013	January 1, 2014
Maryland	April 14, 2014	June 1, 2014
Minnesota	May 29, 2014	May 30, 2014
New York	July 5, 2014	July 5, 2014
Pennsylvania	April 17, 2016	May 17, 2016
Ohio	June 8, 2016	September 8, 2016
Arkansas	November 8, 2016	November 9, 2016
North Dakota	November 8, 2016	December 8, 2016
Florida	November 8, 2016	January 3, 2017

An additional related variable, MEDMJCST, was created to identify reports of medical marijuana use in the past 12 months from question MJMM01 (MEDMJYR = 1) when respondents were not living in a state where marijuana was approved for medical use at the time they were interviewed (Table 8.1).¹⁴² MEDMJCST compares reports of medical marijuana use in the past 12 months with the states where respondents were living at the time they were interviewed.

¹⁴² MEDMJCST is not available on the 2016 NSDUH public use file.

Values of 1 through 5 in MEDMJCST have the following meanings:

- MEDMJCST = 1: The respondent reported medical marijuana use in the past 12 months and was living in a state that had a medical marijuana law in effect at the time the respondent was interviewed. If a state's law was in effect for only part of the year, MEDMJCST was coded as 1 for respondents who reported medical marijuana use and were interviewed after their state's law took effect.
- MEDMJCST = 2: The respondent reported medical marijuana use in the past 12 months and was living in a state that did not have a medical marijuana law in effect at any time in the current survey year.
- MEDMJCST = 3: The respondent reported medical marijuana use in the past 12 months and was in a state where a medical marijuana law went into effect during the current survey year. However, the respondent was interviewed before their state's law took effect. Stated another way, the report of medical marijuana use in the past 12 months occurred before the law was in effect.
- MEDMJCST = 4: The respondent used marijuana in the past 12 months but reported that use of marijuana had not been recommended by a doctor or other health care professional (MEDMJYR = 2).
- MEDMJCST = 5: Whether or not any of the respondent's past year use of marijuana was recommended by a doctor or other health professional was unknown.

In addition, codes of 91 and 93 were used in MEDMJCST to identify respondents who reported never using marijuana or blunts and not using marijuana or blunts in the past year, respectively.

8.3 Prior Substance Use

The prior substance use section covered the following topics:

- retrospective use of marijuana, cigarettes, alcohol, and cocaine in the year prior to the past 12 months (also referred to below as the past 12 to 24 months); and
- the age, year, and month when respondents last used substances covered in the corresponding substance use section of the interview (i.e., age at last use [ALU], month of last use [MLU], and year of last use [YLU]) if they were lifetime but not past month users of these substances.

One of the important aspects of the processing of variables in this section consisted of assigning codes of 91 and 99 (Section 2.4.2) to variables that had been skipped because the questions did not apply. For example, if respondents never used marijuana or indicated use in the past 30 days, they were skipped out of the questions asking for their age, year, and month when they last used marijuana. Respondents were skipped out of the questions pertaining to the age, year, and month when they last smoked cigarettes daily if (1) they smoked every day in the past 30 days or (2) they never had a period in their lives when they smoked cigarettes every day for at least 30 days. When a recency variable from a given substance use section (e.g., MJREC for marijuana) had a refusal code and the corresponding variables in the prior substance use section

had been skipped, that refusal from the substance use recency was propagated to the edited prior substance use variables (Section 2.4.3). Documentation of edits for specific subsections of the prior substance use section is presented in the remainder of this section.

8.3.1 Retrospective Use

Since 2003, this section has included a question about retrospective use of marijuana in the year prior to the past 12 months (i.e., edited variable MRJYRBFR, corresponding to question LU01). In addition, retrospective questions on use of cigarettes, alcohol, and cocaine in the year before the past 12 months (edited variables CIGYRBFR, ALCYRBFR, and COCYRBFR, respectively) have been included since 2005. These questions were analogous to MRJYRBFR.

If respondents never used the drug of interest (e.g., marijuana), the retrospective variable (e.g., MRJYRBFR) was assigned a code of 91. In addition, the cigarette recency variable CIGREC indicated if respondents last smoked cigarettes more than 3 years ago (CIGREC = 4). Logically, if respondents last smoked cigarettes more than 3 years ago, they would not have smoked a cigarette in the past 12 to 24 months. Therefore, when CIGREC indicated that respondents last smoked cigarettes more than 3 years ago and question LU37 (corresponding to CIGYRBFR) had been skipped, CIGYRBFR was assigned a code of 99 (LEGITIMATE SKIP).

The retrospective variables MRJYRBFR, ALCYRBR, CIGYRBFR, and COCYRBFR were not edited for consistency with other data for these drugs in the prior substance use section. For example, if MRJYRBFR indicated that respondents used marijuana in the past 12 to 24 months but respondents indicated that they last used marijuana at an age in edited variable MRJAGLST that would suggest that they last used marijuana more than 2 years ago, no editing was done to either MRJYRBFR or MRJAGLST. However, a codebook "NOTE" was added to alert analysts that these types of inconsistencies could exist between the related variables for these drugs.

8.3.2 Last Use of Drugs

The prior substance use section included questions about the last use of marijuana, cigarettes, smokeless tobacco, cigars, alcohol, marijuana, cocaine (including crack cocaine if applicable), heroin, hallucinogens (including LSD, PCP, and Ecstasy, if applicable), inhalants, or methamphetamine.¹⁴³ If respondents were lifetime users of specific substances in the corresponding substance use section but had not used these substances in the past 30 days, they were asked in this section for the age, year, and month when they last used these substances (i.e., ALU, MLU, and YLU). If respondents ever had a period of smoking cigarettes daily but had not smoked every day in the past 30 days, they also were asked for the age, the year, and the month when they last smoked cigarettes on a daily basis. Because of changes beginning with the 2015 questionnaire, ALU, YLU, and MLU variables for smokeless tobacco, hallucinogens, Ecstasy, inhalants, and methamphetamine were renamed beginning in 2015.

¹⁴³ Questions about the last misuse of prescription psychotherapeutic drugs (pain relievers, tranquilizers, stimulants, and sedatives) were dropped from the prior substance use section in 2015.

Respondents who last used a given substance more than 30 days ago¹⁴⁴ were asked how old they were when they last used that substance. If respondents reported last using the substance within 1 year of their current age, they were asked to report the specific month and year when they last used, with the allowable years ranging from 2014 to 2016. If respondents reported last using the substance at their current age and their birth month was earlier than the interview month (i.e., they reached their current age in the same year that they were interviewed), the CAI program assumed that the last use of the substance occurred in the current year (i.e., 2016). These respondents were asked only for the month that they last used in the current year. The remaining respondents who last used a substance within 1 year of their current age could be routed to one of two possible questions on the specific year they last used. They then were routed to a question to report on the specific month that they last used the substance in the year they had reported previously.

Because the routing logic to the different versions of the MLU and YLU questions was mutually exclusive, composite sets of MLU and YLU variables were created from the individual unedited variables. In addition, if respondents indicated a specific year that they last used a substance, the final YLU variables for 2016 were recoded to replace codes from the questionnaire with values for the years (i.e., 2014 through 2016). If respondents confirmed that they last used a substance at their current age and were interviewed subsequent to their birthday, a code of "2016" was assigned to the YLU; this was done even if respondents did not know what month they last used in the current year, or if they refused to report what month they last used in the current year. If the MLU and YLU questions had been skipped because respondents last used the substance more than 1 year younger than their current ages, legitimate skip codes were assigned to the final MLU and YLU variables.

Since 2005, consistency checks have been included in the section that were triggered when the values for the MLU and YLU were inconsistent with the ALU. Specifically, for respondents who recently stopped use of a given substance, the CAI program calculated a second ALU based on the MLU and YLU data by comparing these data with the respondent's date of birth. This comparison was not done if the respondent reported last use of the substance in the same month that he or she was born; a unique ALU could not be determined from the MLU and YLU in these situations because it was not known whether the substance use occurred before or after the respondent's birthday. Similarly, a consistency check was not triggered if the respondent had missing data in either of the month or year questions, such as if the respondent knew the year when he or she last used a substance but did not know the MLU.

In the remaining situations in which respondents provided complete data for the MLU and YLU, a consistency check was triggered if the MLU and YLU suggested that respondents stopped use of the substance at an earlier or a later age than what they had previously reported in their ALU question. For example, a consistency check was triggered if a 16-year-old respondent reported last using a substance at age 16 but then reported last using the substance in a month and year that would have meant the respondent was 15 years old when he or she last used the substance. No editing needed to be done if respondents indicated twice in a row that the ALU

¹⁴⁴ Subsequent discussion also applies to respondents whose last period of smoking cigarettes every day occurred more than 30 days ago.

that was calculated from the MLU and YLU was correct. The CAI program updated the value for the ALU (e.g., AGELSTCG for cigarettes) to agree with the values for the MLU and YLU.

If respondents indicated at some point in the consistency check sequence that the value they had reported for their ALU (e.g., question LU03 for cigarettes) was correct, they had an opportunity to revise the values for their MLU and YLU. If a consistency check was triggered between the ALU and data in the MLU and YLU, the MLU and YLU were updated with any year and month data that the respondent entered in the consistency checks (e.g., LUCG07 and LUCG07a for any cigarette use). These data were used in subsequent editing steps. Otherwise, the MLU and YLU data were picked up from the original source variables (e.g., LU03a through LU03d for any cigarette use) for use in subsequent editing.

Table B.15 in Appendix B presents the edits that were implemented when consistency checks were triggered between the ALU, MLU, and YLU. The default when a respondent did not resolve an inconsistency between the ALU and the MLU and YLU was to favor the ALU in subsequent editing decisions.

Respondents also were skipped out of the MLU and YLU questions if they indicated that they last used a substance (or last smoked cigarettes every day) at an age that was more than 1 year younger than their current age. In these situations, the edited MLU and YLU variables were assigned legitimate skip codes.

Prior substance use variables were not edited with respect to *imputed* substance use variables. Suppose, for example, that respondents did not know or refused to report when they first used marijuana, but they gave ages at last use for marijuana in question LU02 that were consistent with their current ages. Although the potential existed for the imputed marijuana AFU (imputed variable: IRMJAGE) to be imputed to a value greater than the ALU, the marijuana ALU (MRJAGLST) was not edited for consistency with IRMJAGE.

Similarly, if a substance's recency variable (e.g., MJREC for marijuana) had been assigned an "indefinite" value of 8 or 9 (indicating use at some point in the past 12 months or lifetime, respectively; see Section 6.2.2.3), the corresponding imputed recency (e.g., IRMJRC for marijuana) could be statistically imputed to indicate past month use (IRMJRC = 1). Although the prior substance use section was designed for respondents who were not past month users, any data in this section were retained for respondents who were statistically imputed to be past month users. In this situation, analysts would have the option of deciding whether to use or disregard data from respondents who were imputed to be past month users.

Table B.16 describes the specific edits that were implemented for the ALU, MLU, and YLU variables in the prior substance use section (i.e., other than the edits described in Table B.15 when consistency checks were triggered between the ALU, MLU, and YLU variables). For example, a consistency check was triggered if respondents entered an ALU that was earlier than the age when they reported first using drugs or cigarettes, or when they first smoked cigarettes daily. Because these ALU questions occurred in a section of the interview after the questions that were used to identify past year initiates, respondents were not allowed to change their answers to the corresponding AFU questions from the earlier substance use sections. Thus, the only way that respondents could resolve the inconsistency between the ALU

and AFU was to change their answer to the ALU question. If respondents indicated that their inconsistent ALU was correct or they entered a new ALU value that was still inconsistent with the AFU, the edited ALU variable (e.g., MRJAGLST for marijuana) was assigned a bad data code. Thus, the relevant AFU from the corresponding earlier substance use section was used as the standard against which the corresponding ALU variable was compared. Similarly, the MFU and YFU questions for a given drug from the earlier substance use section of the interview were used as standards for editing the MLU and YLU variables for that drug.

8.4 Substance Use Treatment

The substance use treatment section asked about the receipt of treatment services or the perceived need for treatment for the use of alcohol or other drugs, not counting cigarettes. Questions about the receipt of treatment services included questions about the receipt of treatment in the respondents' lifetime and in the past 12 months, specific locations where respondents received treatment in the past 12 months, emergency room visits in the past 12 months related to their use of specific drugs, whether they were still in treatment, the length of time since they were last in treatment (if they were not currently in treatment), specific questions about their last (or current) treatment episode (including payment), whether they were enrolled in treatment on October 1, 2015, and whether the only treatment they received in the past 12 months was detoxification.

Questions have been included since 2004 to capture information about respondents' life history of substance treatment. In addition, questions TX52 and TX53 have been included in the substance treatment section since 2006 to ask whether respondents had attended a self-help group in the past 12 months (question TX52), and if so, whether this was for alcohol use, drug use, or both (TX53); these questions were added to capture reports of self-help group attendance from respondents who may not think that attending these groups constitutes "treatment."

The substance treatment section also included questions about respondents' perceived need for treatment in the past 12 months if they never received treatment or did not report that they received treatment in the past 12 months. Questions about respondents' perceived need for treatment included questions about specific drugs for which respondents thought they needed treatment and whether they made specific efforts to receive treatment in the past 12 months. In addition, respondents who received treatment in the past 12 months but did not report that they were currently in treatment were asked whether they felt the need for *additional* treatment in the past 12 months. Those respondents who reported that they felt the need for additional treatment were asked about the specific drugs for which they needed additional treatment and whether they made specific efforts to receive additional treatment. Respondents who felt the need for substance use treatment (or additional treatment) also were asked to report why they did not receive the treatment (or additional treatment) that they felt they needed.

8.4.1 General Skip Logic Edits

As noted previously, the substance treatment section was relevant only for respondents who reported some lifetime use of alcohol or other drugs, not counting cigarettes. Therefore, all of the edited treatment variables were assigned a code of 91 (i.e., NEVER USED ALCOHOL OR DRUGS) if respondents were skipped out of the entire substance treatment section because

they never used alcohol, or they had never used or misused illicit drugs, including the misuse of prescription psychotherapeutic drugs (i.e., pain relievers, tranquilizers, stimulants, or sedatives). If respondents reported the lifetime use or misuse of alcohol or illicit drugs, an important aspect of the processing of the substance treatment variables involved assignment of relevant legitimate skip codes when it could be determined unambiguously from respondents' answers that subsequent questions did not apply.

8.4.1.1 Skip Logic Edits for the Receipt of Substance Use Treatment

Respondents who were lifetime users or misusers of alcohol or illicit drugs were asked if they had ever received treatment for their alcohol or other drug use, not counting cigarettes. If respondents reported that they never received treatment (i.e., TXEVRRCD = 2), the CAI program skipped them out of all remaining questions pertaining to the actual receipt of treatment services. Thus, if respondents clearly indicated that they never received treatment, the skipped treatment service variables were assigned legitimate skip codes (Section 2.4.2). As described in Section 2.4.3, when the treatment service questions were skipped because respondents refused to indicate whether they ever received treatment, the edited variables were assigned a refusal code; if treatment service questions were skipped because respondents did not know whether they ever received treatment, the edited variables retained a value of blank.

Similarly, respondents were not asked subsequent questions about receipt of treatment services in the past 12 months if they did not report having ever received treatment in that period (i.e., TXYRRECVD = 2). Thus, if respondents reported that they did not receive treatment in the past 12 months and there were no other responses in the substance treatment section to suggest that they had (Section 8.4.2), legitimate skip codes were assigned to the variables pertaining to receipt of treatment in specific locations in the past 12 months. The procedures for editing 12-month treatment variables that had been skipped when respondents refused to indicate whether they had received treatment in the past 12 months or did not know whether they had received treatment in this period were the same as those described above.

8.4.1.1.1 Skip Logic Edits for Emergency Room Treatment for Specific Drugs

If respondents reported that they received treatment in the past 12 months, it was possible for them to be asked subsequent questions about treatment in an emergency room in the past 12 months for their use of marijuana, cocaine, heroin, LSD, PCP, or methamphetamine (questions TX05 and TX06). Respondents were not asked these questions if they previously reported that their treatment in the past 12 months was only for their use of alcohol. Thus, "legitimate skip" codes were assigned to the corresponding edited variables pertaining to emergency room use (TXYRERDRG and TXYRERNUM), provided there were no other answers in the substance treatment section that indicated treatment for use of these drugs, which would suggest that respondents should have been asked these questions. Similarly, legitimate skip codes were assigned to the edited variable pertaining to the number of emergency room episodes for treatment of these six drugs (TXYRERNUM) if respondents reported that they never received treatment in an emergency room related to their use of these drugs.

In addition, respondents who reported receiving treatment in the past 12 months for their use of illicit drugs but who never used marijuana, cocaine, heroin, LSD, PCP, or

methamphetamine were not asked the questions about use of hospital emergency room services for the use of these drugs. Rather than assign the usual type of legitimate skip code (i.e., 99 or 89), however, a special code of 6 was assigned in these situations, provided that the respondent had not indicated receipt of treatment for any of these drugs elsewhere in the substance use treatment section. This code had the following meaning: 6 = Never used the relevant drug.

8.4.1.1.2 Skip Logic Edits for Last or Current Treatment for Specific Substances

Respondents who reported receiving treatment in their lifetime or who were currently receiving treatment were not asked certain questions about the receipt of treatment related to their use of specific drugs if they were lifetime nonusers of these drugs.¹⁴⁵ For example, respondents who never used heroin were not asked whether they last received (or were currently receiving) treatment for their use of heroin. Similarly, respondents who reported receiving treatment in the past 12 months but who never used marijuana, cocaine, heroin, LSD, PCP, or methamphetamine were not asked the questions about use of hospital emergency room services for the use of these drugs. As described in the previous section, a special code of 6 (Never used the relevant drug) was assigned in these situations, provided that the respondent had not indicated receipt of treatment for any of these drugs elsewhere in the substance use treatment section.

This coding was done because respondents could be routed into or skipped out of a number of different combinations of questions depending on their reported drug use history. For example, suppose a respondent reported that he or she had received treatment in the past 12 months and was a lifetime user of alcohol, marijuana, cocaine, and hallucinogens and also misused prescription pain relievers in the past 12 months. This respondent would selectively be asked the questions about treatment for these drugs during his or her last treatment or current episode and would not be asked the questions pertaining to treatment for heroin, inhalants, methamphetamine, prescription tranquilizers, prescription stimulants, and prescription sedatives.

8.4.1.2 Skip Logic Edits for the Substance Use Treatment History Variables

Most of the editing of the substance treatment history questions TX45 through TX51A that have been present since 2004 also involved assigning legitimate skip codes where relevant. Consistent with the logic described above, if respondents had never used alcohol or other drugs, these variables were assigned codes of 91 or 991. In addition, if respondents reported in question TX01 that they had never received treatment, these variables were assigned legitimate skip codes (e.g., 99 or 999). Questions TX45 through TX48A also pertained to respondents who had received treatment but not in the past 12 months. Therefore, if respondents reported in question TX02 that they received treatment in the past 12 months, the edited variables corresponding to questions TX45 through TX48A were assigned legitimate skip codes. In addition, when respondents reported receiving treatment in the past 12 months, subsets of the variables corresponding to questions TX49 through TX51A were assigned legitimate skip codes based on answers in question TX03 regarding receipt of treatment in the past 12 months for alcohol, drugs, or both. Variables corresponding to TX49 through TX51A also were assigned legitimate

¹⁴⁵ For brevity, lifetime nonuse in this section also refers to respondents who did not report the lifetime misuse of prescription psychotherapeutic drugs.

skip codes based on indications in the initial substance use sections that respondents never used alcohol or never used any illicit drugs. Similar assignment of legitimate skip codes occurred for the variables corresponding to TX49 through TX51A if respondents received treatment in their lifetime but not in the past 12 months, or depending on respondents' answers in the earlier substance use sections or in special drugs for methamphetamine. For example, if respondents answered question TX02 as "no" regarding receipt of treatment in the past 12 months, the variables corresponding to TX49 through TX51A were assigned legitimate skip codes; by definition, these respondents had to have answered the lifetime treatment question TX01 as "yes."

8.4.1.3 Skip Logic Edits for the Perceived Need for Substance Use Treatment

The variables pertaining to the perceived need for substance use treatment were compared with data on receipt of treatment services in the past 12 months to identify when the perceived need for treatment variables were legitimately skipped. For example, if respondents had received treatment services in the past 12 months, the questions about perceived need for treatment in that period did not apply. Thus, legitimate skip codes were assigned to the variables pertaining to the perceived need for any alcohol or other drug treatment when respondents had received treatment in the past 12 months. Similarly, if respondents received treatment in the past 12 months and they reported that they were still in treatment (TXCURRENT = 1), the questions about perceived need for additional services did not apply, and legitimate skip codes were assigned to the corresponding edited variables.

Respondents who had not indicated that they received treatment in the past 12 months and who were lifetime users of alcohol or some other drug also were skipped out of questions regarding their perceived need for additional treatment. Again, the edited variables corresponding to perceived need for additional services were assigned legitimate skip codes. Those respondents who had not indicated that they received treatment in the past 12 months were asked the general question about whether they perceived themselves as needing treatment for their use of alcohol or other drugs (edited variable NDTXYRADG). If they did not see themselves as needing treatment, they were skipped out of questions pertaining to perceived need for treatment for specific drugs in the past 12 months. Again, legitimate skip codes were assigned to the edited variables that had been skipped.

Similarly, respondents were globally skipped out of questions TX11 through TX22 (regarding their perceived need for any treatment for alcohol or specific other drugs) if they reported in question TX02 that they received treatment in the past 12 months. Therefore, the edited variables corresponding to questions TX11 through TX22 (NDTXYRALC through NDTXEFFRT) were assigned legitimate skip codes.

Legitimate skip codes of 99 or 89 were assigned to the perceived need for treatment variables in situations in which respondents were lifetime nonusers of a particular substance. For example, if respondents indicated that they needed treatment for their use of alcohol or drugs, they were asked about their perceived need for treatment only for those specific drugs that they had ever used; legitimate skip codes were assigned to the skipped drug-specific variables that respondents had never used. Thus, for example, if a respondent had never used heroin but reported needing treatment in the past 12 months for alcohol or drugs (TX08 = 1), a legitimate

skip code was assigned to the edited variable pertaining to the perceived need for treatment for heroin (NDTXYRHER).

Procedures consistent with those described in Section 2.4.3 also were implemented when questions about the perceived need for treatment were potentially applicable, but respondents refused to report whether they had ever used a particular drug. For example, if a respondent had not received treatment in the past 12 months, reported needing treatment in the past 12 months for alcohol or other drugs, but refused to report whether he or she had ever used heroin, the item about perceived need for treatment for heroin was skipped. Because the respondent refused to report about lifetime use or nonuse of heroin, the edited variable NDTXYRHER was assigned a refusal code.

8.4.2 Lifetime, Past Year, or Current Treatment

8.4.2.1 Most Recent Receipt of Treatment

If respondents reported the receipt of substance use treatment in their lifetime (i.e., TXEVRRCVD = 1), then subsequent edits to questions TX07 (edited variable TXCURRENT) and TX24 (edited variable TXRCVDREC) established the most recent period in which they received treatment. [Table B.17](#) in Appendix B discusses edits that were relevant to the determination of when respondents last received treatment or whether respondents were currently receiving treatment.

In particular, the editing procedures generally favored responses that indicated a more recent receipt of treatment. For example, respondents could report that they received treatment in the past 12 months (TX02 = 1) but then subsequently report that the last time they received treatment was more than 12 months ago (TX24 = 3). For these respondents, the recency of treatment was inferred to be at some point within the past 12 months (TXRCVDREC = 8). Respondents also could provide an answer other than "yes" when asked in question TX02 whether they had received treatment in the past 12 months and then indicate that they were currently receiving treatment (TX07 = 1) or that they last received treatment in the past 30 days or more than 30 days ago but within the past 12 months (TX24 = 1 or 2). In these situations, the respondents were logically inferred in TXYRRECVD (corresponding to question TX02) to have received treatment in the past 12 months.

In addition, respondents could answer question TX07 as "don't know" or "refused" but indicate in question TX38 for the outcome of the last treatment episode that he or she was still in treatment. TXCURRENT was assigned a code of 3 (Yes LOGICALLY ASSIGNED) to indicate that the respondent was logically inferred to still be in treatment.

The edits also inferred that the last treatment episode logically occurred more than 12 months prior to the interview date when there were no other data to suggest that respondents received treatment in the past 12 months. Specifically, respondents could answer TX02 as "no" to indicate that they did not receive treatment in the past 12 months but answer question TX24 as "don't know" or "refused." When this occurred, TXRCVDREC was assigned a code of 13, where 13 = More than 12 months ago LOGICALLY ASSIGNED).

Similarly, respondents could answer "don't know" or "refused" when asked whether they had received treatment in the past 12 months and then report that they last received treatment more than 12 months ago. In this situation, a negative response was logically inferred for the variable pertaining to receipt of treatment services in the past 12 months (TXYRRECVD = 4), where 4 = No LOGICALLY ASSIGNED. In turn, the skipped variables pertaining to the receipt of substance use treatment in the past 12 months were assigned legitimate skip codes.

Also, if respondents answered TX07 as "don't know" or "refused" but their last treatment episode from TXRCVDREC was more than 30 days prior to the interview, then it was logically inferred that the respondents were no longer in treatment. In this situation, TXCURRENT was assigned a code of 4 (No LOGICALLY ASSIGNED) to indicate that the respondent was logically inferred no longer to be in treatment.

8.4.2.2 Past Year Treatment for Alcohol, Illicit Drugs, or Both

If respondents reported in question TX02 that they received substance use treatment in the past 12 months, they were asked in question TX03 (edited variable TXYRALDGB) to report whether they received treatment in the past 12 months for their use of alcohol, their use of drugs, or both. Respondents were asked whether they received treatment in the past 12 months in specific locations for their use of alcohol, illicit drugs, or both only if they reported in TX03 that they received treatment in the past 12 months for their use of both alcohol and drugs (TX03 = 3). Consequently, the variables for the receipt of treatment in the past 12 months in specific locations for the use of alcohol, illicit drugs, or both were assigned legitimate skip codes when respondents had consistent data (see below) for the receipt of treatment in the past 12 months only for their use of alcohol or only for their use of illicit drugs.

However, all respondents who reported treatment in the past 12 months were asked TX03 regardless of whether they had previously reported using alcohol but not illicit drugs (in which case, they would have been expected to have received treatment only for their use of alcohol) or vice versa. No editing was done to TXYRALDGB when the answers in TX03 disagreed with respondents' reports of their lifetime use of alcohol or illicit drugs.

Table B.17 in Appendix B discusses additional edits pertaining to TXYRALDGB. Although TXYRALDGB was not edited with respect to data from earlier substance use sections, for example, TXYRALDGB *was* edited with respect to substance use treatment data pertaining to the substances for which respondents received treatment during their last or current treatment episode. Logically, if respondents reported in TX02 that they received treatment in the past 12 months, then their last treatment episode (i.e., if they were not currently receiving treatment) also was in the past 12 months. In particular, the following edits were implemented for TXYRALDGB when the answer in TX03 was inconsistent with the data on the substances for which respondents received their last or current treatment.

- If respondents reported in TX03 that they received treatment only for their use of alcohol in the past 12 months but they reported receiving treatment for one or more illicit drugs during their last or current treatment episode, then TXYRALDGB was assigned a special code of 11.

- If respondents reported in TX03 that they received treatment only for their use of illicit drugs in the past 12 months but they reported receiving treatment for alcohol during their last or current treatment episode, then TXYRALDGB was assigned a special code of 12.

When values of 1 or 2 in TX03 were replaced with codes of 11 or 12 in TXYRALDGB because of inconsistent data, then the skipped variables for the receipt of treatment in the past 12 months in specific locations for the use of alcohol, illicit drugs, or both retained codes of 98 (i.e., blank) rather than being assigned legitimate skip codes.

8.4.2.3 Locations of Past Year Treatment

If respondents reported in question TX02 that they received substance use treatment in the past 12 months, they were asked in questions TX04A to TX04I to report the specific locations where they received treatment in that period. If respondents reported in question TX04I that they received treatment in some other place in the past 12 months, they were asked to specify the description of the place where they received treatment.

[Table B.17](#) in Appendix B discusses edits pertaining to the receipt of treatment in specific locations in the past 12 months. Consistent with the procedures described in Section 2.4.5, for example, if respondents did not report receiving treatment in the past 12 months in a particular location in questions TX04A to TX04H but they specified that location in question TX04I as some other place where they received treatment in the past 12 months, then the corresponding edited variable for that location was assigned a code of 3 (Yes LOGICALLY ASSIGNED).

Respondents who reported receiving treatment in their lifetime also were asked in question TX25 to report the main place where they last received treatment or the main place where they were currently receiving treatment. If respondents chose "Some other place" in TX25, they were asked to specify the description of the place. As noted in Section 2.3.1.4, if respondents reported that the main place where they received treatment during their last or current treatment episode was some other place and they specified a location that corresponded to one of the locations in TX25, then TXLTVMNPL was assigned the relevant code ranging from 21 (a hospital overnight as an inpatient) to 28 (a self-help group).

The edited variables corresponding to questions TX04A to TX04I also were edited with respect to data in TXLTVMNPL. Logically, if respondents last received treatment in the past 12 months or they were currently receiving treatment, then the main location from TXLTVMNPL also applied to the locations in TX04A to TX04I where respondents received treatment at any point in the past 12 months. Therefore, if respondents who received treatment in the past 12 months reported that the main place where they received treatment during their last or current episode was not reported in TX04A to TX04I, then the corresponding edited variable for TX04A to TX04I was assigned a special "yes" code of 5. This logically assigned code of 5 distinguished these edits from the edits that were described previously when respondents specified a particular location where they received treatment in the past 12 months as "some other place." If the respondent was inferred from TXLTVMNPL to have received treatment in some other place, then the "OTHER, Specify" response from TXLTVMNPL also was moved

over to the corresponding "OTHER, Specify" data for the location of treatment in the past 12 months (edited variable TXYROTHSP).

In addition, respondents could report that they received treatment in the past 12 months but could answer "no" to all questions in TX04A to TX04I for the specific locations where they received treatment in that period. When this occurred, the edited variable corresponding to question TX04I (TXYROTHER) was assigned a special code of 7 to indicate that the location where the respondent received treatment in the past 12 months was unknown.

8.4.3 Last or Current Treatment Episode

Questions about the last or current treatment episode were asked principally of respondents who reported that they received treatment in the past 12 months (question TX02 answered as "yes"); the logic also routed respondents to the last or current treatment questions if they did not know or refused to report in question TX02 whether they had received treatment in the past 12 months. If respondents received treatment in the past 12 months (or answered question TX02 as "don't know" or "refused") and reported in question TX07 that they were currently in treatment,¹⁴⁶ subsequent questions asked about the main location where they were receiving treatment, specific drugs for which they were receiving treatment, the primary drug for which they were receiving treatment (if treatment for more than one drug was reported), the length of time that they had been in treatment thus far, and anticipated payment sources for their current treatment. If respondents were asked question TX07 and did not report currently being in treatment, these subsequent questions pertained to their last treatment episode, such as the duration of their last treatment and the payment sources for their last treatment. Respondents who did not report that they were currently in treatment also were asked about the outcome of their last treatment.

8.4.3.1 Receipt of Treatment for Specific Substances

Respondents who reported receiving treatment in their lifetime or who were currently receiving treatment were asked questions about their receipt of treatment for their use of the specific substances for which they reported lifetime use or misuse. These respondents also were asked whether they last received or were currently receiving treatment for their use of any other drug.

Table B.17 in Appendix B presents edits that were specific to the variables for the receipt of treatment for the use of specific substances. For example, if respondents did not report receiving treatment for a particular substance during their last or current treatment episode but they specified this substance as "some other drug" for which they received treatment, then the corresponding edited variable for that substance was assigned a code of 3 (Yes LOGICALLY ASSIGNED). Similarly, if respondents did not report receiving treatment for a particular substance but they reported in question TX37 (edited variable TXLTMAIN) that this substance was the main substance for which they last received or were currently receiving treatment, then they were inferred to have received (or to be receiving) treatment for that particular substance in the corresponding substance-specific variable. Because of the addition of the new

¹⁴⁶ Question TX07 asks, "Are you currently receiving treatment or counseling for your [TXFILL1]?" where [TXFILL1] could be replaced with "alcohol use," "drug use," or "alcohol or drug use."

methamphetamine section in 2015, however, no editing was done to the variable for treatment for the use of prescription stimulants (TXLTYSTIM) if respondents specified that they received treatment for their use of methamphetamine; prior to 2015, respondents were inferred to have received treatment for their use of prescription stimulants if they received treatment for methamphetamine.

In addition, respondents could fail to report treatment for any of the specific substances that they used or misused in their lifetime. The following edits were implemented when this occurred:

- If the respondent reported receiving treatment only for alcohol in the past 12 months in question TX03 (edited variable TXYRALDGB), a special logically inferred "yes" code of 5 was assigned to the variable for alcohol treatment during the last treatment episode (TXLTYALCO).
- If the respondent reported receiving treatment only for drugs other than alcohol in the past 12 months in TX03, a special code of 5 was assigned to the "some other drug" variable (TXLTYOTHR) to indicate that the drug for which the respondent received treatment was unknown.
- Otherwise, a special code of 7 was assigned to TXLTYOTHR, the "some other drug" variable, to indicate that the last or current treatment for alcohol or specific other drugs was unknown.

8.4.3.2 Outcome of the Last Treatment and Length of Time in Treatment

Composite variables combining data from more than one individual item were created for the variables pertaining to (1) the outcome of the last treatment episode, for respondents who were not currently in treatment (edited variable TXLTYOCOM); and (2) the length of time that respondents had been in treatment or currently had been in treatment thus far (TXLTYDAYS).

For TXLTYOCOM, respondents could select a response category from a list, including that the last treatment had some other outcome. Only those respondents who reported some other outcome were routed into a second item where they were asked to specify the other outcome of their treatment. Consequently, the final variable for the outcome for that last episode included data from both the existing response categories that respondents were allowed to choose and the valid "other" responses that they specified. If respondents chose the other category but specified something that was coded with a missing value (i.e., "bad data," "don't know," "refused," or blank), a final code of "other" was retained for TXLTYOCOM.

The variable pertaining to the length of time that respondents had been in treatment (TXLTYDAYS) was derived from a series of questions that asked respondents to indicate whether they wanted to give their answer in terms of days, months, or years, and from questions that asked for the number of days, months, or years that they were in treatment. TXLTYDAYS was expressed as the number of days that respondents were in treatment. If respondents answered in terms of a number of months, their reported number of months was multiplied by 30. If respondents answered in terms of a number of years that they had been in treatment, their reported number of years was multiplied by 365.

[Table B.17](#) in Appendix B presents edits that were specific to the variables for the outcome of the last treatment and the length of time for the last or current treatment episode. For example, if respondents answered in terms of a number of months in treatment, the treatment duration data were compared for consistency with the respondent's age. Specifically, the number of months in treatment was divided by 12 to yield an estimated number of years in treatment. If the reported number of years in treatment exceeded the respondent's current age, then TXLTYDAYS was assigned a bad data code. If the difference between the respondent's current age and the number of years in treatment was 10 or fewer years, then this data pattern was flagged. Such respondents would have been reporting that they had *not* been in treatment for 10 or fewer years. However, TXLTYDAYS was not set to bad data for this latter situation.

8.4.3.3 Payment for the Last or Current Treatment

Respondents who reported receiving treatment in their lifetime or who were currently receiving treatment were asked in questions TX42A through TX42K to report whether different sources paid (or were paying) for this last or current treatment. If questions TX42A through TX42J were all answered as no, respondents were asked in TX42K whether their last or current treatment was free. Thus, if respondents reported one or more sources of payment in TX42A through TX42J, then the edited variable TXPAYFREE corresponding to TX42K was assigned a legitimate skip code.

[Table B.17](#) in Appendix B presents additional edits that were specific to the variables for the payment for the last or current treatment episode. For example, if respondents did not report that a particular source paid (or is paying) for their last or current treatment episode but they specified that payment source as "some other source," then the corresponding edited variable for that specific payment source was assigned a code of 3 (Yes LOGICALLY ASSIGNED).

8.4.4 Editing of Miscellaneous Treatment Variables that Existed Prior to 2004

Two additional variables pertaining to the receipt of substance use treatment were included in this section of the interview prior to the 2004 survey. Question TX43 (edited variable TXENRLOCT) asked whether respondents who had received treatment in their lifetime were enrolled in a substance use treatment program on October 1 of the previous calendar year (i.e., 2014 for the 2015 survey). Question TX44 (edited variable TXYRONDTX) asked respondents who received treatment in the past 12 months whether the only treatment that they received for their use of alcohol, drugs, or both was detoxification.

No editing was done to TXENRLOCT, other than to assign a mnemonic name. Minimal editing was done to TXYRONDTX. As described in [Table B.17](#) in Appendix B, however, respondents could report in TX44 that the only treatment they received in the past 12 months was detoxification, or they could answer TX44 as "don't know" or "refused." If such respondents also reported that they attended self-help groups in the past 12 months, then TXYRONDTX was assigned a code of 4 (No LOGICALLY ASSIGNED). Because self-help groups typically are not places where people go to receive detoxification, it was inferred that these respondents had received some treatment other than detoxification.

8.4.5 Lifetime History of Treatment and Additional Self-Help Questions

As noted previously, questions have been included since 2004 to capture information about respondents' life history of substance use treatment. Additional self-help group questions also have been included since 2006 to capture reports of self-help group attendance from respondents who may not think that attending these groups constitutes "treatment."

8.4.5.1 Lifetime History of Treatment

Respondents who had ever received treatment for their use of alcohol or other drugs but did not receive treatment in the past 12 months were asked questions regarding their receipt of treatment for their use of alcohol, other drugs, or both, depending on the substances they had previously reported using in earlier substance use sections. Where relevant, these respondents subsequently were asked to report the ages when they first received treatment for alcohol, drugs, or both. Respondents who reported that they had received treatment in the past 12 months were asked similar questions. Questions about treatment life history for respondents who reported that they had received treatment in the past 12 months also took into account their answers to question TX03, regarding whether they received treatment in the past 12 months for their use of alcohol, drugs, or both.

[Table B.18](#) presents edits that were specific to the lifetime substance use treatment history variables that were added to the survey in 2004. For example, respondents could report that they first received treatment for their use of alcohol at ages that were earlier than when they first reported using alcohol. No editing was done to these data. However, flags were created to indicate whether the ages for first treatment of alcohol or other drugs were consistent with reported ages at first use from the earlier substance use sections, and if not, the flags indicated the degree of inconsistency between these data. The meaning of the values in these flag variables is discussed in more detail in [Table B.18](#).

In addition, data from substance treatment variables that existed prior to 2004 were used to edit these added variables. However, data from these added lifetime substance use treatment history variables were not used to edit the substance treatment variables that existed prior to 2004. Consequently, variables that also existed prior to 2004 (as well as in 2004 through 2016) were created in a manner that was comparable with how these variables were created in prior years.¹⁴⁷

8.4.5.2 Additional Self-Help Variables

An important aspect of editing the variables TXSHGWENT (corresponding to question TX52) and TXSHGALDB (corresponding to TX53) involved assigning relevant codes based on the skip logic for these items. For example, if respondents were routed to question TX52 and answered it as "no," then TXSHGALDB was assigned a legitimate skip code. Similarly, if TX52

¹⁴⁷ An exception is that the variables in the substance use treatment section were renamed in 2015 because changes to the 2015 questionnaire for hallucinogens, inhalants, methamphetamine, and prescription psychotherapeutic drugs could have affected the comparability of the substance treatment data between 2015 and prior years.

was refused, that refusal was propagated to TXSHGALDB. In addition, if the lifetime treatment variable TXEVRRCVD (corresponding to question TX01) had been set to bad data, then any nonblank values also were replaced with codes for bad data in TXSHGWENT and TXSHGALDB (see [Table B.17](#) in Appendix B).

If question TX52 had been skipped because respondents had reported earlier in the substance treatment section that they received treatment in a self-help group in the past 12 months (from question TX04H), then TXSHGWENT was assigned a code of 5, where 5 = Yes LOGICALLY ASSIGNED (from skip pattern). In addition, TX53 was skipped when question TX04H was answered as "yes." Therefore, if values of 1, 2, or 3 existed in TXYRSLFAD (corresponding to question TX04H1 and indicating treatment in a self-help group in the past 12 months for alcohol only, drugs only, or both alcohol and drugs, respectively), TXSHGALDB (corresponding to TX53) was assigned corresponding codes of 11, 12, or 13. "Bumping" these values by 10 in the assignment of codes from TXYRSLFAD to TXSHGALDB indicated to analysts that a logical inference was made in TXSHGALDB; in contrast, codes of 1, 2, or 3 in TXSHGALDB came from respondents answering TX53 directly. Other values greater than 3 that existed in TXYRSLFAD were moved over to TXSHGALDB without further editing. In particular, if TXYRSLFAD (from question TX04H1) had been legitimately skipped because respondents reported in question TX03 that they were treated in the past 12 months for alcohol only or drugs only, then a legitimate skip code could be assigned to TXSHGALDB as well.

Variables pertaining to the lifetime receipt of treatment (TXEVRRCVD), treatment in the past 12 months (TXYRRECVD), treatment in a self-help group in the past 12 months (TXYRSLFHP), and treatment for alcohol, drugs, or both in a self-help group in the past 12 months (TXYRSLFAD) were not edited to make them consistent with data in questions TX52 and TX53. This decision was in agreement with the decision that was made in 2004 not to edit substance treatment variables that existed in the section prior to 2004 for consistency with data from questions TX45 through TX51A that were added to the substance treatment section in 2004.

However, a flag variable (TXSHGFLAG) was created to alert analysts to situations in which data were inconsistent between TXSHGWENT and TXYRSLFHP. The default value of TXSHGFLAG was 98 (blank). In addition, TXSHGFLAG was assigned values of 91 if respondents had never used alcohol or other drugs, and it was assigned values of 81 if the logical inference had been made that respondents had never used alcohol or other drugs. Assignment of values other than 81, 91, or 98 to TXSHGFLAG is described below.

- TXSHGFLAG was set to a value of 1 if respondents answered TX52 as "yes" (i.e., attended a self-help group for their own use of alcohol or other drugs in the past year) but the respondents answered question TX04H as "no" (i.e., did not receive treatment in a self-help group in the past 12 months). In this situation, the answers in TX04H and TX52 may be seen as contradictory, except that respondents may not have thought of self-help group attendance as "treatment" when they were answering question TX04H.
- TXSHGFLAG was set to a value of 2 if respondents answered TX52 as "yes" but they reported either that they did not receive any treatment in the past 12 months or in their lifetime. A separate level was created for this pattern because these respondents

may not have thought of self-help groups as "treatment" when they were answering question TX01 (lifetime receipt of treatment) or TX02 (receipt of treatment in the past 12 months). Consequently, respondents who attended self-help groups in the past 12 months may have answered TX01 or TX02 as "no."

- TXSHGFLAG was set to a value of 3 if the edited variable TXYRSLFHP (corresponding to question TX04H) indicated that respondents were logically inferred to have received treatment in a self-help group in the past year, but TXSHGWENT also did not indicate that these respondents attended a self-help group in the past 12 months. In this situation, TXSHGWENT was inconsistent with TXYRSLFHP, but only because a logical inference had been made that TXYRSLFHP should have been answered as "yes."
- TXSHGFLAG was set to a value of 4 if TX52 was answered as "yes" but the edited variable TXYRSLFHP had a value for "don't know," "refused," or "blank." In particular, TXYRSLFHP could have been blank if question TX04H had been skipped and it had subsequently been inferred that respondents got treatment in the past 12 months. In these types of situations, data in TXSHGWENT and TXYRSLFHP were not necessarily contradictory.
- TXSHGFLAG was set to a value of 5 if question TX04H was answered as "no" and TX52 was answered as "don't know" or "refused." Although a logical inference was not made in TXSHGWENT that respondents did not attend self-help groups in the past 12 months, analysts would have the option of inferring that respondents absolutely did not attend self-help groups in the past 12 months for their own use of alcohol or other drugs based on TXYRSLFHP. Alternatively, analysts could decide to treat self-help group attendance data as missing based on TXSHGWENT.

8.4.6 Perceived Need for Substance Treatment

Questions and corresponding edited variables were added in 2015 for the perceived need for treatment or additional treatment for the use of methamphetamine, separate from questions about the perceived need for treatment for the use of prescription stimulants. The questions in 2016 pertaining to respondents' perceived need for substance use treatment did not change relative to 2015.

[Table B.19](#) presents edits that were specific to the variables pertaining to the perceived need for treatment services. As noted in Section 8.4.1.3, for example, respondents were skipped out of questions TX11 through TX22 if they reported that they received treatment in the past 12 months. If respondents had not originally reported receiving treatment in the past 12 months but were logically inferred to have done so (see [Table B.17](#)), these respondents would have been routed to questions TX11 through TX22. Rather than replace respondents' answers with codes for "bad data," however, special codes were assigned to indicate that respondents were routed into questions about their perceived need for treatment for use of specific drugs when they were logically inferred to have received treatment in the past 12 months. This procedure would allow analysts to decide whether to use or disregard these data in their analyses.

8.4.7 Reasons for Not Receiving Treatment

Since 2002, respondents have been asked to report the reasons why they did not receive substance treatment services despite feeling the need for treatment (question TX22A). Similarly, respondents who reported that they received treatment but needed additional treatment have been asked why they did not receive additional treatment or counseling (question TX23A).

Based on a review of what respondents had specified in quarter 1 of 2002 as leading "other" reasons for not getting substance treatment or additional treatment, additional questions (TX22B and TX23B) have been included in the interview since 2003 to capture commonly endorsed other reasons for not receiving treatment. Therefore, if respondents answered question TX08 as "yes" (i.e., perceived the need for substance treatment), they eventually were routed to question TX22A, where they could report up to 10 reasons for not receiving substance treatment, including "some other reason or reasons." If respondents chose the "some other reason or reasons" option in question TX22A, they were routed next to question TX22B, where they could choose additional reasons for not getting substance treatment; again, respondents were given the option in question TX22B to report "some other reason or reasons." If respondents chose this "other" response category in TX22B, they were asked to specify the *most important* other reason why they did not get treatment. Thus, respondents were asked to specify only the most important other reason for not getting substance treatment. Similar logic was in place for questions TX23A and TX23B, when respondents reported in TX09 that they perceived a need for additional substance treatment.

Questions TX22A, TX22B, TX23A, and TX23B were "enter all that apply" questions in which respondents could choose more than one reason from each list. Each response option (e.g., "You had no health care coverage, and you couldn't afford the cost") was captured as a separate variable. The edited variables corresponding to the individual response options in TX22A, TX22B, TX23A, and TX23B were coded as 1 (Response entered) or 6 (Response not entered), if at least one item was chosen from the TX22A, TX22B, TX23A, and TX23B lists (Section 2.4.4).

In addition, the 10th response option in questions TX22A and TX23A ("some other reason or reasons") was principally considered to be a "toggle" to questions TX22B and TX23B, respectively. Therefore, separate "some other reason or reasons" variables were not created to correspond to the last response category in TX22A and TX22B. Similarly, separate variables were not created to correspond to the last category in TX23A and TX23B.

The edits described below were implemented in 2003 to take into account the addition of questions TX22B and TX23B. The edits described use TX22A and TX22B as examples but also applied to TX23A and TX23B.

- If respondents chose the 10th response option in TX22A, any response that was entered from the TX22B series was coded as 1, and anything that respondents did not choose from the TX22B list was coded as 6.
- If respondents chose a response from TX22A but did not choose the 10th response category in TX22A, the variables corresponding to the response categories in TX22B (i.e., NDTXNONED through NDTXOTRSN) all were given a code of 6 (Response

not entered), rather than being assigned "legitimate skip" codes. That is, TX22A and TX22B were considered together to be one big series of reasons.

- If respondents chose category 10 in question TX22A, chose at least one reason from TX22B, but did not choose category 6 in TX22B (some other reason or reasons), the edited variable NDTXOTRSN (some other reason or reasons for not receiving substance treatment) was assigned a code of 6. That is, it was inferred in this situation that the list of specific reasons in TX22B was adequate for capturing why respondents did not get substance treatment. For example, if a respondent chose category 10 in TX22A and then chose only category 2 in TX22B ("You thought you could handle the problem without treatment"), it would be reasonable to infer that this response in TX22B was the only other reason why the respondent did not get treatment.
- If respondents chose response category 10 in question TX22A, it was possible for them to answer TX22B as "don't know" or "refused" (i.e., did not know or refused to report what the other reasons were). When this occurred, the "some other reason" variable NDTXOTRSN was set to 1 (Response entered) in order to retain information that the respondent chose "some other reason or reasons" somewhere in the series. Remaining variables corresponding to the TX22B series retained codes of 94 ("don't know") or 97 ("refused").
- If respondents answered question TX22A as "don't know" or "refused," question TX22B was skipped. Therefore, the relevant code of 94 or 97 was propagated onto the variables corresponding to the TX22B list.
- If NDTXOTRSN had a value of 6 (see above), the "OTHER, Specify" variable NDTXRIMP was assigned a legitimate skip code. If NDTXOTRSN had a refusal code, that refusal was propagated onto NDTXRIMP.
- If NDTXOTRSN had a code of 1 when the respondent answered "don't know" or "refused" to the TX22B series, the "OTHER, Specify" variable NDTXRIMP retained a code of 98 (blank).

Consistent with general editing procedures, if respondents reported a reason that corresponded to a reason in the lists for TX22A/TX22B or TX23A/TX23B, that reason was logically inferred to have been chosen in the relevant edited variable (Section 2.4.5). Suppose, for example, that a respondent had not received treatment but felt the need for it, and the respondent specified that one of the reasons for not receiving treatment was that he or she was not ready to stop using alcohol or drugs. If the respondent had not chosen this response in TX22A, the edited variable NDTXNSTOP (corresponding to response category 5 in TX22A) was assigned a code of 3 (Response entered LOGICALLY ASSIGNED). Similarly, if the respondents specified that they did not get treatment because they thought they could handle the problem without treatment but had not chosen that reason in TX22B, the edited variable NDTXHANDL (no substance treatment because the respondent thought he or she could handle the problem without treatment) was assigned a code of 3.

Conversely, if respondents did not report "some other reason" why they did not receive treatment in the past 12 months (edited variable NDTXOTRSN = 6, corresponding either to response category 10 in question TX22A not being chosen or response category 6 in question TX22B not being chosen), legitimate skip codes were assigned to the edited "OTHER, Specify"

variable NDTXMIMPT (corresponding to question TX22SP). Similar edits were done for the "OTHER, Specify" variable pertaining to reasons for not receiving additional treatment if respondents reported that they felt the need for additional treatment but did not indicate "some other reason" for not receiving additional treatment.

8.5 Consumption of Alcohol

The consumption of alcohol section covered a variety of topics related to respondents' use of alcohol. Administration of questions in this section was limited to respondents who had ever used alcohol, and a subset of questions was limited to respondents who reported alcohol use in the past 30 days. Respondents were asked questions about the following topics:

- number of drinks consumed on the last occasion of alcohol use in the past 30 days;
- the last occasion of alcohol use in the past 30 days by adolescents or young adults younger than the legal drinking age of 21, including who respondents were with, where they drank, and how they obtained the alcohol;
- use of other drugs in combination with alcohol (i.e., at the same time or within a couple of hours of using alcohol) in the past 30 days; and
- history of binge alcohol use (i.e., consumption of five or more drinks on a single occasion for males or four or more drinks on a single occasion for females).

The editing procedures for the consumption of alcohol section are documented below.

8.5.1 General Skip Logic Edits

One of the important aspects of the processing of variables in the consumption of alcohol section consisted of assigning codes of 91, 93, and 99 to variables that had been skipped because the questions did not apply. In particular, certain questions may not apply to respondents for multiple reasons, such as if they never used alcohol or if they were aged 21 or older.

Therefore, [Table B.20](#) in Appendix B discusses the prioritization of skip logic edits in this section. For example, if respondents had never used alcohol, variables in the consumption of alcohol section were assigned a code of 91, 991, or 9991 (NEVER USED ALCOHOL), regardless of whether questions were not applicable for other reasons (e.g., if respondents were aged 21 or older). However, this edit kept a constant number of cases coded as indicating nonuse of alcohol for most variables in the section.

As noted in [Table 6.2](#) in Chapter 6, respondents who reported that they used alcohol in the past 30 days but confirmed use on 0 days in that period were inferred to have used alcohol at some point in the past 12 months (ALCREC = 8). However, the CAI program in 2016 continued to treat these respondents as past month alcohol users and routed respondents to questions CA01 through CA09 (where applicable), as though these respondents used alcohol in the past 30 days. This pattern applied to a maximum of about 300 respondents in 2016.¹⁴⁸ If data existed in

¹⁴⁸ Other CAI skip logic for these consumption of alcohol questions continued to apply if respondents confirmed the use of alcohol on 0 days in the past 30 days. For example, questions CA02A through CA07 for underage alcohol users continued not to be applicable to respondents with this pattern who were aged 21 or older.

questions CA01 through CA09, then these data were retained in the corresponding edited variables. These consumption of alcohol data would be consistent with a final imputed alcohol recency that indicated use in the past 30 days (IRALCRC = 1). Analysts could decide what to do with these consumption of alcohol data if respondents were imputed to have last used alcohol more than 30 days ago but within the past 12 months (IRALCRC = 2). Beginning in 2017, respondents will not be asked these consumption of alcohol questions if they reported in the alcohol section of the interview that they used alcohol on 0 days in the past 30 days.

8.5.2 Number of Drinks for the Last Alcohol Use in the Past 30 Days

Other than implementation of edits based on the skip logic (see [Table B.20](#)), only limited editing was done for the variables pertaining to the number of drinks that respondents consumed the last time they drank in the past 30 days (question CA01; edited variable CADRLAST). However, data from CADRLAST were used in editing the binge alcohol history variables (Section 8.5.5).

8.5.3 Source and Location of Last Underage Alcohol Use in the Past 30 Days

Other than implementation of edits based on the skip logic (see [Table B.20](#)), only limited editing was done for the variables pertaining to alcohol use in the past 30 days among adolescents or young adults aged 12 to 20 (questions CA02A through CA07SP). In particular, no editing was done to CABUYFRE (corresponding to question CA03), CAGVMONY (corresponding to question CA04), CABUYWHO (corresponding to question CA05A), or CABPLACE (corresponding to question CA05B) if respondents reported in question CA04 that they gave money to someone else who bought their last alcoholic beverage for the respondents, but then they reported in CAFRESP (corresponding to CA07SP) that the "other" way that they got the last alcoholic beverage they drank was by purchasing it, such as with fake identification. The rationale for not editing these variables was that there may not be enough detail from what respondents specified to know how these variables should be edited. For example, question CA04 about buying alcohol was asked only if respondents reported in question CA03 that they paid for the last alcohol that they drank. Even if the respondent specified buying alcohol in CAFRESP, however, it would not be known whether the respondent bought the alcohol with his or her own money or whether the alcohol was purchased with someone else's money.

In question CA07 (edited variable CAFREWHO), respondents were asked to give one response for how they got the last alcoholic beverage that they drank, including the option to report that they got it "some other way." Respondents who reported that they got the alcohol some other way were asked to specify what this other way was (edited variable CAFRESP). If a response was given in CAFRESP that corresponded to one of the existing response options in question CA07, then CAFREWHO was assigned the relevant response plus a value of 10. For example, if CAFRESP indicated that respondents got their last alcoholic beverage from a nonrelative who was 21 or older who gave the alcohol to the respondent (i.e., equivalent to CA07 = 3), then CAFREWHO was assigned a code of 13.

Questions CA02B (where underage respondents drank alcohol the last time) and CA09 (the specific drugs that respondents of any age used at the same time or within a couple of hours of when they last drank alcohol) were "enter all that apply" questions. As in other sections, the

responses were captured as discrete variables. The individual edited variables for these "enter all that apply" variables were coded as 1 or 6 to indicate that the response was entered or not entered, respectively (Section 2.4.4).

For the edited variables corresponding to question CA02B, if respondents specified some other place where they drank alcohol the last time that corresponded to one of the existing response options in question CA02B, and they had not chosen that response, then the corresponding edited variable was assigned a code of 3, where 3 = Response entered LOGICALLY ASSIGNED. For example, if respondents reported that they drank alcohol in someone else's home the last time they drank but they had not chosen this response in question CA02B, then the edited variable CADROTHM was assigned a code of 3. In addition, codes of 94 and 97 (for "don't know" and "refused," respectively) were assigned to the entire list of variables corresponding to question CA02B if respondents did not know or refused to report where they were the last time they drank alcohol.

8.5.4 Use of Illicit Drugs with Alcohol in the Past 30 Days

Because of changes beginning with the 2015 questionnaire that affected comparability of estimates for illicit drug use measures, edited variables corresponding to questions CA08 and CA09 were renamed in 2015. For the edited variables corresponding to question CA09, codes of 94 or 97 were retained only for those drugs that respondents used in the past 30 days (or were potentially past month users, such as if a drug's recency variable had been coded as 8, meaning that the respondent used the drug at some point in the past 12 months, which could include within the past 30 days). Otherwise, codes of 94 or 97 in the individual drug variables corresponding to question CA09 were replaced with a code of 89 (LEGITIMATE SKIP Logically assigned) if respondents used some drug in the past 30 days but not the drug that corresponded to the variable of interest. Suppose, for example, that respondents used marijuana and inhalants in the past 30 days and did not use any other illicit drugs in that period. If question CA08 was answered as "yes" and CA09 was refused, then the edited variables CADRKMARJ and CADRKINHL retained a code of 97 and the remaining drug variables were assigned a code of 89.

Special consideration also was given to the editing of the variables CADRKCOCN, CADRKHERN, and CADRKMETH for the use of cocaine, heroin, or methamphetamine, respectively, during or within a couple of hours of the episode of alcohol use in the past 30 days. In particular, respondents who answered question CA08 as "yes" would have the response option "Cocaine or 'crack'" (captured in CADRKCOCN) available to them if they previously reported past month use of cocaine in any of the following places: (1) in the section for any cocaine use, (2) in the section for crack cocaine use, or (3) in the special drugs section (i.e., use of cocaine with a needle in the past 30 days). Similarly, respondents who answered question CA08 as "yes" could choose the response option "Heroin" (captured in CADRKHERN) if they reported past month use of heroin in the heroin section or if they reported past month sniffing, smoking, or injection of heroin in the special drugs section. For methamphetamine, the response option "Methamphetamine" (captured in CADRKMETH) was available to respondents who indicated past month methamphetamine use in the methamphetamine section or in the special drugs section (i.e., use of methamphetamine with a needle in the past 30 days).

Consequently, respondents would have the response options for CADRKCOCN, CADRKHERN, or CADRKMETH available to them if they reported past month use of the relevant drug(s) in the special drugs section, even if they had not reported past month use in the corresponding earlier section(s) of the interview for those specific drugs. In most situations, no further editing was done to these data, even though the corresponding earlier substance use data did not indicate use in the past 30 days.

8.5.5 Binge Alcohol History

The threshold for binge alcohol use in the past 30 days for females changed from consumption of five or more drinks on an occasion in 2014 to consumption of four or more drinks on an occasion beginning in 2015; the threshold for males since 2015 has remained at consumption of five or more drinks on an occasion. In turn, the binge alcohol history questions CA10 through CA11D since 2015 have asked females about their history of consumption of four or more drinks on a single occasion; for males, questions CA10 to CA11D continued to ask about the history of consumption of five or more drinks on an occasion. Questions CA12 through CA14D from the 2014 questionnaire for the history of females' consumption of four or more drinks on an occasion were redundant with the revised binge alcohol history questions for females and were dropped from the 2015 questionnaire.

If respondents were lifetime alcohol users and other data from the alcohol section or from CADRLAST did not indicate that respondents had been binge drinkers in the past 30 days,¹⁴⁹ then they were asked in question CA10 (edited variable CABINGEVR) whether they had ever engaged in binge drinking in their lifetime. Those respondents who had ever been binge drinkers (or had reported binge drinking in the past 30 days in the alcohol section or in CADRLAST) were asked in question CA11 (edited variable CABINGAGE) to report how old they were the first time when they first engaged in binge drinking. If CABINGAGE was within 1 year of respondents' current age, they were routed to questions CA11A through CA11D (edited variables CABINGYFU and CABINGMFU) to report the year and month when they first engaged in binge drinking.

As noted previously, the routing logic for these questions took into account respondents' answers in the earlier alcohol section. Specifically, the CAI program created a flag variable called BINGEFLAG. The default value for this flag was 0. The flag was coded as 1 if question AL08 in the alcohol section indicated that respondents had been binge drinkers in the past 30 days. Also, if question AL07 indicated that respondents usually had four or more drinks (for females) or five or more drinks (for males) on those days when they drank in the past 30 days, then this flag was coded as 1.

The edited variable in the consumption of alcohol section corresponding to this flag was named CABINGFLG. If the variables ALCUS30D (corresponding to question AL07) and ALCBNG30D (corresponding to question AL08) indicated binge drinking in the past 30 days but CABINGFLG did not already have a value of 1, then CABINGFLG was set to a logically assigned value of 11. If ALCUS30D and ALCBNG30D had been set to bad data as a result of

¹⁴⁹ For brevity, remaining text in this section refers to reports of binge alcohol use or binge drinking, defined as consumption of four or more drinks on a single occasion on at least 1 day in the past 30 days for females or as consumption of five or more drinks on a single occasion on at least 1 day in this period for males.

edits in the alcohol section (see Section 6.2.4.7 and [Table B.5](#) in Appendix B), then CABINGFLG also was set to bad data. In turn, any questions that had been answered in CA11 through CA11D were set to bad data.

[Table B.21](#) discusses additional edits associated with the binge alcohol history variables in the consumption of alcohol section. The first set of edits pertains to the consistency checks between the age when respondents reported that they first engaged in binge drinking and the age at initiation of binge drinking based on the year and month when respondents reported that they first engaged in this behavior.

[Table B.21](#) then presents edits for these binge alcohol variables relative to respondents' ages and to corresponding alcohol use data from earlier in the interview. For example, if respondents reported in question CA11 that they first engaged in binge drinking at an age that was earlier than what they reported for when they first had a drink of any alcohol (from the alcohol use variable ALCTRY), then CABINGAGE was set to bad data. Any nonmissing values in CABINGYFU and CABINGMFU also were replaced with codes for bad data. Similarly, suppose respondents first used any alcohol at their current age. If they had a lifetime history of binge drinking (including any indication of this behavior in the past 30 days) but CA11 had a missing value (e.g., "don't know" or "refused"), then it could be logically inferred in CABINGAGE that the first episode of binge alcohol use occurred at the respondents' current age.

In addition, "editing indicator" (EI) variables were created to indicate when nonmissing values were logically assigned to the binge alcohol variables in the consumption of alcohol section. The default value for these EI variables was 1 (i.e., questionnaire data). The EI variables were assigned a value of 2 (Logically assigned data) if a logical inference was made that involved assignment of a nonmissing value to the binge alcohol history variables.

In keeping with the general principle of not editing across sections of the interview (Section 2.4.1), no editing was done to the data in the earlier alcohol section when a respondent's reported first binge alcohol use suggested more recent use of alcohol than was indicated in the recency variable ALCREC. For example, ALCREC may not indicate use in the past month or past year, but respondents could report that they first had five or more drinks or four or more drinks on a single occasion at their current age. Although these data from the consumption of alcohol would suggest that these respondents used alcohol in the past 12 months, no editing was done to ALCREC based on these data.

9. Editing of Self-Administered Sections for Risk and Protective Factors for Substance Use

9.1 Introduction

This chapter discusses editing procedures for the following self-administered sections that pertain to risk and protective factors for substance use:

- risk/availability,
- special topics,
- social environment (administered to adults aged 18 or older),
- youth experiences (administered to adolescents aged 12 to 17), and
- parenting experiences (administered to parents when an adolescent aged 12 to 17 also was selected for a separate interview).

These interview sections are included in this chapter according to their related topical content rather than in the sequential order in which they appeared to respondents during the interview. The order in which these sections were presented to respondents relative to other self-administered sections of the interview can be found in the computer-assisted interviewing (CAI) specifications for the 2016 National Survey on Drug Use and Health (NSDUH) that are available at <https://www.samhsa.gov/data/>.

None of the edited variables in this chapter were imputed to remove missing values. Although the special topics section of the interview includes questions about driving under the influence of alcohol or illegal drugs, it was included in this chapter on risk and protective factors because most of the questions in this section pertain to arrests or other involvement with the criminal justice system (e.g., being on probation or parole in the past 12 months).

9.2 Risk/Availability

The risk/availability section asked about the perceived risk of harm associated with use of alcohol or specific illegal drugs, perceived ease of obtaining different illegal drugs, whether respondents were approached by someone in the past 30 days who was trying to sell an illegal drug, and general risk-taking types of behaviors. The latter included questions on the frequency with which respondents got a "kick out of doing things that are a little dangerous," how often they tried to test themselves "by doing something a little risky," and their frequency of seatbelt use.

Minimal processing of data was done to variables in this section. Specifically, the unedited variables were assigned final, mnemonic variable names (e.g., RSKPKCIG corresponding to question RK01A, which asked about the perceived risk of harm associated with

smoking one or more packs of cigarettes per day). No further editing or processing was done to the variables in this section.

9.3 Special Topics

The special topics section asked about arrests in the respondents' lifetime and in the past 12 months, including arrests for specific offenses in the past 12 months (not counting minor traffic violations). This section also included questions about respondents being on probation or parole in the past 12 months, operating vehicles under the influence of alcohol or illegal drugs¹⁵⁰ in the past 12 months, and respondents' knowledge about their states' marijuana laws.

9.3.1 Editing of Variables for Criminal Justice Involvement

Questions in the special topics section of the interview that pertained to respondents' involvement with the criminal justice system included questions about arrests in the respondents' lifetime and in the past 12 months and whether respondents were on probation or parole in the past 12 months. Being arrested and booked for breaking the law was defined for respondents as being taken into custody and processed by the police or someone connected with the courts, even if respondents were then released. Consequently, respondents who reported that they were arrested and booked were not necessarily jailed for an offense, nor were they necessarily convicted of a crime.

If respondents reported that they had never been arrested in their lifetime and they did not report being on probation or parole in the past 12 months (see below), the edited variables pertaining to arrests in the past 12 months were assigned legitimate skip codes (Section 2.4.2). Other standard edits described in Section 2.4.3 pertaining to situations where respondents answered "don't know" or "refused" to the lifetime arrest question were applied to the past year arrest variables that had been skipped.

Similarly, if respondents reported being arrested in their lifetime but reported being arrested 0 times in the past 12 months, the questions pertaining to arrests for specific offenses in the past 12 months were assigned legitimate skip codes. Respondents who did not know how many times they were arrested in the past 12 months or who refused to answer this question were asked whether they were arrested for specific offenses in the past 12 months. This was consistent with the logic in 1999 and since 2001 but differed from the logic in 2000, when respondents who answered "don't know" or "refused" to the question about the number of specific arrests in the past 12 months were skipped out of questions about arrests for specific offenses in the past 12 months.

Table B.22 in Appendix B discusses additional edits for the criminal justice variables in the special topics section. For example, respondents could report that they had never been arrested in their lifetime but could report that they were on probation, parole, or supervised release in the past 12 months. Because someone could not be on probation or parole without first

¹⁵⁰ Questions SP06A and SP06C specifically asked about driving under the influence of "illegal" drugs. Because respondents in some states that have legalized the recreational use of marijuana among adults may not consider marijuana to be an "illegal" drug, the questions about driving under the influence of alcohol or other substances have undergone further revision in the 2016 NSDUH.

having been arrested for a crime, these respondents were logically inferred to have been arrested in their lifetime. When this situation occurred, the skipped variables pertaining to arrests in the past 12 months retained a value of blank.

Since 2002, respondents who reported in question SP02 that they were arrested and booked for breaking the law at least once in the past 12 months and gave negative answers to every question about specific arrests (including arrests for "some other offense") have been routed to a follow-up question, SP03R, that asked respondents to verify their previous answer from SP02. If respondents did not indicate that their previous answer from question SP02 was correct, they were routed to question SP03S, where they were allowed to update their answer regarding the number of times they were arrested or booked in the past 12 months. If respondents indicated in SP03S that they had been arrested or booked 0 times in the past 12 months, the edited variable NOBOOKYR was assigned a value of 0. Further, when SP03S indicated that respondents had been arrested and booked 0 times in the past 12 months, it was logically inferred that all items pertaining to specific arrests in the past 12 months should have been skipped. Therefore, all of the variables associated with specific arrests in the past 12 months were assigned a code of 89 (Section 2.4.2).

If respondents indicated in SP03R that their previous answer from SP02 was correct, or if they reported being arrested and booked for at least one offense in SP03S, they were asked to specify at least one offense for which they were arrested and booked in the past 12 months (questions SP03RSP or SP03SSP). If respondents specified a legitimate offense in SP03RSP or SP03SSP after giving negative answers to every question about specific arrests, the "OTHER, Specify" variable BKOTHOFF was updated to incorporate the response from SP03RSP or SP03SSP. A code of 3 (Yes LOGICALLY ASSIGNED) also was assigned to the "some other offense" variable (BKOTH) to indicate that this edit had taken place (Section 2.4.5).

Any information that respondents specified in SP03RSP or SP03SSP also was used to edit variables pertaining to offenses that respondents previously had been asked about. For example, if respondents had answered question SP03A as "no" (i.e., had not been arrested and booked for motor vehicle theft in the past 12 months) but then specified in SP03RSP or SP03SSP that motor vehicle theft was one of the offenses for which they had been arrested, the edited variable BKMVTHFT was assigned a code of 3.

If respondents reiterated in SP03R or SP03S that they had been arrested and booked for at least one offense in the past 12 months but still did not report a legitimate offense in questions SP03RSP or SP03SSP (including situations in which they answered SP03RSP or SP03SSP as "don't know" or "refused"), then a code of 5 was assigned to BKOTH. This code of 5 had the following meaning: 5 = Offense unknown. Stated another way, the response from SP02 or SP03S was retained in NOBOOKYR to indicate that the respondents were arrested in the past 12 months, but it was not possible to determine the specific crime for which they were arrested.

9.3.2 Editing of Variables for Driving under the Influence

Questions about driving under the influence from 2015¹⁵¹ were replaced in 2016 with new questions (SP06A to SP06G) that asked about driving under the influence of alcohol, marijuana, cocaine (including crack, where applicable), heroin, hallucinogens, inhalants, and methamphetamine, respectively. Respondents who reported in question SP06A that they drove under the influence of alcohol in the past 12 months and also reported in SP06B to SP06G that they drove under the influence of any of the other substances mentioned previously in that same period were asked in question SP06H if they drove under the influence of alcohol only. Because of the changes to these questions in 2016, the variables in 2016 for driving under the influence of alcohol or other substances are not comparable with the variables from 2015.

Respondents were skipped out of questions pertaining to driving under the influence of alcohol or the specific other substances in SP06B to SP06G if they reported in previous sections of the interview that they did not use the relevant substance(s) in the past 12 months. For example, a respondent who did not report use of marijuana in the past 12 months was not asked SP06B about driving under the influence of marijuana in that period.

For cocaine, heroin, and methamphetamine, however, respondents who had previously reported that they last used these substances more than 12 months prior to the interview also were asked the questions about driving under the influence of these substances if they reported past year use in the special drugs section of the interview (Section 7.2). In addition, the editing of the variables for driving under the influence of cocaine, heroin, or methamphetamine required respondents who reported lifetime nonuse in the corresponding substance use sections to have no reports of lifetime use of these substances with a needle in the special drugs section (Section 7.2.5.1). Because of the additional logic for how respondents could be asked the questions about driving under the influence of cocaine, heroin, or methamphetamine, edits to the variables for driving under the influence were implemented for two groups: (1) alcohol, marijuana, hallucinogens, and inhalants; and (2) cocaine, heroin, and methamphetamine.

9.3.2.1 Driving under the Influence of Alcohol, Marijuana, Hallucinogens, or Inhalants

The edits described below were implemented for the variables for driving under the influence of alcohol, marijuana, hallucinogens, or inhalants when respondents were determined to have skipped out of the corresponding questions because of nonuse in their lifetime or in the past 12 months. Edits for the variable for driving under the influence of alcohol only (i.e., but not illicit drugs) are discussed separately in Section 9.3.2.3.

- If respondents were lifetime nonusers of the relevant substance, then the corresponding skipped variable pertaining to driving under the influence of that substance was assigned a code of 91 to indicate that the respondents were lifetime nonusers of that substance (Section 2.4.2). For example, if respondents never used

¹⁵¹ Questions SP06A through SP06C in 2015 asked respectively about driving under the influence of alcohol and illegal drugs used together, driving under the influence of alcohol (or alcohol only), and driving under the influence of illegal drugs (or illegal drugs only). In some states, however, respondents may not consider marijuana to be an "illegal" drug.

alcohol, then the edited variable DRVINALCO (corresponding to question SP06A) was assigned a code of 91.

- If respondents last used the relevant substance more than 12 months prior to the interview, then the corresponding skipped variable was assigned a code of 93 to indicate that respondents did not use that substance in the past 12 months (Section 2.4.2).

If respondents were not asked the question about driving under the influence of alcohol, marijuana, hallucinogens, or inhalants, and the corresponding recency variable had been assigned a value of 9 (Used at some point in the lifetime LOGICALLY ASSIGNED), then the edited variable for driving under the influence retained a code of 98 (blank). The variables for driving under the influence of alcohol, marijuana, hallucinogens, or inhalants also were assigned "bad data" codes (Section 2.3) for a small number of respondents in 2016 (i.e., fewer than five) if the corresponding recency variable had been assigned a value of 9 (Used at some point in the lifetime LOGICALLY ASSIGNED) but data existed in the variable for driving under the influence. Data for driving under the influence in the past 12 months would indicate that these respondents had used the corresponding substance in that period. For most of these affected respondents in 2016, this edit was implemented because of problems with interview dates (Section 2.3.3). This edit also was implemented for the variable for driving under the influence of hallucinogens (DRVINHALL, corresponding to question SP06E) because of a response of "don't know" or "refused" in a consistency check between "parent" and "child" hallucinogen questions (Section 6.2.2.5). In this situation, the affected variables for parent and child hallucinogen recency had been edited to a value of 9 but the CAI program continued to treat the respondents as past year users. In addition, if variables for a given drug had been assigned codes for bad data because of patterned responses (Section 2.3.2), then the corresponding variable for driving under the influence was assigned a code for bad data.

9.3.2.2 Driving under the Influence of Cocaine, Heroin, or Methamphetamine

Similar logic to that described in Section 9.3.2.1 was implemented for determining when respondents had legitimately skipped out of the corresponding questions for driving under the influence of cocaine, heroin, or methamphetamine because of nonuse in their lifetime or in the past 12 months. As noted previously, however, edits to the variables for driving under the influence of these substances also took into account data from the special drugs section of the interview. Thus, for example, if respondents reported lifetime nonuse of cocaine, heroin, or methamphetamine in that substance's section of the interview, but the corresponding needle recency variable from the special drugs section had been edited to indicate use with a needle more than 12 months ago, then the skipped variable for driving under the influence (e.g., DRVINCOCN, corresponding to question SP06C) was coded as 93 (DID NOT USE IN THE PAST 12 MONTHS) instead of 91 (NEVER USED) as part of the editing.

As was also indicated previously, respondents who reported in sections of the interview for cocaine, heroin, or methamphetamine that they last used these substances more than 12 months ago but who reported use at some point in the past 12 months in the special drugs questions were asked the corresponding questions for driving under the influence. For example, if respondents reported in the cocaine section of the interview that they last used cocaine "more than 12 months ago," but they reported using cocaine with a needle in the past 30 days or more

than 30 days ago but within the past 12 months, then data from question SP06C were retained in the edited variable DRVINCOCN.

As for the edits described in Section 9.3.2.1, if respondents were not asked the question about driving under the influence of cocaine, heroin, or methamphetamine, and the corresponding recency variable had been assigned a value of 9 (Used at some point in the lifetime LOGICALLY ASSIGNED), then the edited variable for driving under the influence retained a code of 98 (blank). The edited variable for driving under the influence also retained a code of 98 if respondents reported that they last used cocaine, heroin, or methamphetamine more than 12 months ago in the corresponding recency variable but they had missing data for recency questions for these substances in the special drugs section.¹⁵²

9.3.2.3 Driving under the Influence of Alcohol Only

As noted previously, respondents were asked question SP06H about driving under the influence of only alcohol if they reported in question SP06A that they drove under the influence of alcohol in the past 12 months and they also reported that they drove under the influence of marijuana, cocaine, heroin, hallucinogens, inhalants, or methamphetamine in that same period. The edits described below were implemented for the variable DRVINALON for driving under the influence of only alcohol when respondents were determined to have skipped out of question SP06H.

- If respondents were lifetime nonusers of alcohol, then DRVINALON was assigned a code of 91 (NEVER USED ALCOHOL). In this situation, both DRVINALON and DRVINALCO (drove under the influence of alcohol in the past 12 months) were coded as 91.
- If respondents last used alcohol more than 12 months prior to the interview, then DRVINALON was assigned a code of 93 (DID NOT USE ALCOHOL IN THE PAST 12 MONTHS). In this situation, both DRVINALON and DRVINALCO were coded as 93.
- If respondents reported that they did not drive under the influence of alcohol in the past 12 months (DRVINALCO = 2), then DRVINALON was assigned a code of 99 (LEGITIMATE SKIP) (Section 2.4.2). This edit was implemented regardless of whether respondents reported that they drove under the influence of other substances (i.e., marijuana, cocaine, heroin, hallucinogens, inhalants, or methamphetamine) in the past 12 months.
- If respondents reported that they drove under the influence of alcohol in the past 12 months (DRVINALCO = 1), but they did not drive under the influence of other

¹⁵² For the variable for driving under the influence of heroin in the past 12 months (DRVINHERN, corresponding to question SP06D), a small number of respondents in 2016 (fewer than five) were given a code of 93 (DID NOT USE HEROIN IN THE PAST 12 MONTHS) instead of the variable being coded as 98 (blank) because the heroin recency (HERREC) indicated use more than 12 months ago. However, some of the data for smoking, sniffing, or using heroin with a needle were missing. This issue would not have resulted in respondents being reclassified as having driven under the influence of heroin in the past 12 months. Edits for DRVINHERN will be updated for 2017 to treat such respondents as having unknown information for driving under the influence of heroin in the past 12 months.

substances in that period (including situations where respondents did not use these substances in the past 12 months or in their lifetime), then DRVINALON was assigned a code of 99. In this situation, respondents logically drove only under the influence of alcohol in the past 12 months.

- If DRVINALCO had missing data but respondents did not drive under the influence of other substances in the past 12 months, then DRVINALON was assigned a code of 99. In this situation, respondents were not eligible to be asked question SP06H because they did not drive under the influence of other substances in the past 12 months, despite having unknown information for whether they drove under the influence of alcohol in that period.

DRVINALON retained a code of 98 (blank) in the situations described below when question SP06H had been skipped.

- Use of alcohol in the past 12 months was unknown and either of the following applied to driving under the influence of other substances in that period: (1) respondents drove under the influence of other substances, or (2) driving under the influence of one or more other substances also was unknown. In these situations, both DRVINALON and DRVINALCO were coded as 98.
- Respondents used alcohol in the past 12 months, DRVINALCO had missing data (e.g., question SP06A was answered as "don't know" or "refused"), and the same conditions that were described in the preceding bullet applied to driving under the influence of other substances in that period.
- Respondents drove under the influence of alcohol in the past 12 months (DRVINALCO = 1), they did not report driving under the influence of any other substances in that period, but information was unknown for whether respondents drove under the influence of some of these other substances.

9.3.3 Editing of Variables for State Marijuana Laws

The variable pertaining to a state's maximum legal penalty for first offense possession of an ounce or less of marijuana was renamed from MXMJPENL to MXMJPNLT. The variable name was changed to reflect the addition of a response category for "No penalty" and to alert analysts that data in MXMJPNLT from 2006 onward may not be comparable with data in MXMJPENL prior to 2006.

Minimal editing was done to the variables MXMJPNLT, MXMJSURE, and MEDMJUSE, corresponding to questions SP07, SP08, and SP09, respectively; these variables pertained to knowledge about state marijuana laws and penalties. If respondents indicated in MXMJPNLT that they did not know the maximum legal penalty in their states for possession of an ounce or less of marijuana for personal use, or if they refused to answer this question, the edited variable MXMJSURE (regarding respondents' degree of certainty about their answer to question SP07) was assigned a legitimate skip code. In addition, if interviewers had entered incorrect information in the FIPE4 question regarding the state where the respondent's sampled dwelling unit was located, the variables MXMJPNLT, MXMJSURE, and MEDMJUSE were

assigned bad data codes.¹⁵³ This latter edit was done because the state that respondents were asked about in these questions was governed by the state that interviewers entered in FIPE4. Hence, if interviewers entered incorrect state information in FIPE4, the answers that respondents provided in questions SP07 through SP09 were deemed to be questionable. For example, if a respondent lived in California (FIPE4 = 5) but the interviewer entered that the respondent's sampled dwelling unit was in Colorado (FIPE4 = 6), the respondent would be asked for information on marijuana laws in Colorado.

9.4 Social Environment

As noted in Section 9.1, the social environment section was administered only to adults. This section included questions about involvement in criminal or potentially criminal activities, attitudes about adults trying marijuana once or twice, and religious involvement. A shorter version of this section has been administered since 2005 with the deletion of items related to neighborhood cohesiveness. Consequently, the name of this section has been changed from "social and neighborhood environment" prior to 2005 to "social environment" since 2005.

As was the case in prior years, minimal processing of data was done to variables in this section. The primary data processing involved assignment of legitimate skip codes for respondents who were aged 12 to 17 (Section 2.4.2). Adults were asked all questions within the social environment section. As noted in Section 2.3.2.2, responses were replaced with codes for bad data if adult respondents keyed "1" (or multiple-digit responses of "1" if they occurred) to all of the questions that were asked in the section. Although a code of "1" would be plausible for all variables in this section, responses were set to bad data if the pattern of keying "1" wherever possible continued from prior sections.

9.5 Youth Experiences

As noted in Section 9.1, the youth experiences section was administered only to respondents aged 12 to 17. This section included questions about school enrollment and related issues (e.g., opinions about the importance of assigned schoolwork) in the past 12 months, including homeschooling; other social and family characteristics (e.g., substance use behaviors of other students or friends, personal attitudes about substance use, parental attitudes about substance use); people with whom the youth could confide about a serious problem; exposure to alcohol- and other drug-related prevention messages in school or outside school; and personal behaviors (e.g., involvement in criminal or potentially criminal activities, involvement in extracurricular activities) that might be associated positively or negatively with the use of alcohol or other drugs. The youth experiences section also included questions about youths' religious involvement in the past 12 months and opinions about religious issues.

In 2016, the following instruction was added to YE09 (school attendance in the past 12 months) to better define school: "The next questions are about school. By 'school,' we mean elementary school, junior high or middle school, high school, or a college or university. Please

¹⁵³ Creation of the edited variable STATELOC from FIPE4 is discussed in detail in Chapter 4.

include home schooling as well." This change was determined not to have affected the comparability of school-related variables between 2016 and prior years.

Minimal processing of data was done to variables in this section. The primary data processing involved assignment of legitimate skip codes based on the CAI routing logic. That included (1) assignment of legitimate skip codes to variables in the entire section for respondents who were aged 18 or older and (2) assignment of legitimate skip codes to youths' data based on routing logic within the youth experiences section.

As noted in Section 2.3.2.2, responses in the youth experiences section were replaced with bad data codes when respondents keyed answers of "1" (or multiple-digit responses of "1," where applicable) to every question that they were asked within a section. In addition, adolescents who keyed answers of "2" wherever possible in the youth experiences questions had their answers replaced with bad data codes. For respondents to key all possible items as "2" in the youth experiences section would mean that they keyed this response to 42 items in a row, including that they had engaged in all "problem" activities (such as getting into fights, carrying a handgun, or selling illegal drugs) exactly "1 or 2 times" in the past year. These respondents who keyed "2" wherever possible also would have participated in exactly two school-based activities, two community-based activities, two church- or faith-based activities, and two other kinds of activities in the past year.

Some special issues were encountered in editing the variables corresponding to question YE22, which pertained to people whom youths could turn to if they had a serious problem. Specifically, youths were asked to enter all the different types of people to whom they could turn to (e.g., a parent, a friend). This question also included a response category for youths who felt that there was no one they could talk to about a serious problem.

The questions indicating the youths' relationships to people whom they could turn to if the youths had a serious problem were "enter all that apply" questions. The individual edited variables for relationships were coded as 1 or 6 to indicate that the response was entered or not entered, respectively (Section 2.4.4). If the entire list of responses was blank (e.g., if a youth broke off the interview before getting to these questions), the edited variables retained a code of "blank."

Youths could indicate that there was no one they could talk to about a serious problem but then indicate that they could talk to one or more of the people or types of people in the list from question YE22. In this situation, the variable pertaining to the first item in the list ("There is nobody I can talk to about a serious problem") was assigned a code of 11 (if that response was chosen along with another response from the list). Similarly, a code of 11 was assigned to the edited relationship variables (e.g., my mom, my dad) when they were chosen along with the response that there was nobody that the youth could talk to.

9.6 Parenting Experiences

The parenting experiences section was intended to be administered only in dwelling units (DUs) where (1) two people had been selected for an interview, (2) a 12- to 17-year-old had been selected for an interview (regardless of whether the youth completed the interview), and (3) the

respondent being interviewed was the parent or legal guardian of the 12- to 17-year-old who also was selected for an interview. Editing of the parenting experiences data first involved editing the field interviewer (FI) checkpoint variables (FIPE1, FIPE2, and FIPE3) completed by the interviewers toward the beginning of the interview.¹⁵⁴ The variables in the parenting experiences section then were edited based on the final values assigned to the edited FIPE variables.

The audio computer-assisted self-interviewing (ACASI) "lockout" feature that was described in Section 2.4.6 did not affect how the FI checkpoint data were edited but did affect the editing of the parenting experiences data. Issues associated with this ACASI "lockout" that are relevant to the editing of the parenting experiences variables are described in Section 9.6.2.

9.6.1 Editing of the Field Interviewer Checkpoint Variables

Interviewers were instructed to enter into these checkpoints the relevant information described above for determining whether respondents were eligible to be administered the parenting experiences questions. These checkpoint variables were edited for consistency with the pair-selection and pair-respondent sample variables (PAIRSEL and PAIRRESP, respectively). These checkpoints were interviewer-administered and not self-administered. Editing of these checkpoints was related to the edits for the parenting experiences questions (which were self-administered), however, because the final values in the edited checkpoints were critical for determining whether respondents were in fact eligible to be asked the parenting experiences questions.

Editing of the FIPE1 Checkpoint (and Related Edits). First, the FIPE1 variable was edited for consistency with the pair-selection variable PAIRSEL. Specifically, this checkpoint pertained to whether two people were selected for an interview at that DU. There were no situations in 2016 when two people were interviewed at a given DU without two people having first been selected. Therefore, editing FIPE1 involved reviewing only information on the number of people selected for an interview at that DU based on PAIRSEL.

If the pair-selection data indicated that two people were selected from that DU, then FIPE1 should have been answered as "yes." Therefore, if the pair-selection data indicated that two people were selected and FIPE1 was not answered as "yes," a code of 3 (i.e., Yes LOGICALLY ASSIGNED) was assigned to the edited FIPE1 variable (SKPX2PER). Similarly, if the pair-selection data indicated that only one person was selected from that DU, then FIPE1 should have been answered as "no." Therefore, if the pair-selection data indicated that only one person was selected and FIPE1 was not answered as "no," the editing procedures logically inferred that "no" should have been the answer. If the edited version of FIPE1 indicated that two people were not selected for an interview, then the edited versions of FIPE2 (SKPX1217) and FIPE3 (SKPXPRNT) were assigned legitimate skip codes. If data existed in FIPE2 or FIPE3 when the edited SKPX2PER was inferred to be answered as "no," SKPX1217 and SKPXPRNT were assigned a code of 89 (i.e., LEGITIMATE SKIP Logically assigned) to signify that these two checkpoints should have been skipped (Section 2.4.2).

¹⁵⁴ "Checkpoint" refers to an item completed by the interviewer about the location of the sampled dwelling unit (SDU) or characteristics of the sample within the SDU; these checkpoints are not seen by the respondent and are used to determine what the respondent is asked in subsequent sections of the interview.

Editing of the FIPE2 Checkpoint (and Related Edits). Next, FIPE2 was edited for consistency with PAIRSEL, PAIRRESP, and the age of the respondent. Specifically, this checkpoint pertained to whether a 12- to 17-year-old was selected for an interview at that DU, *regardless of whether the selected youth actually responded*. Edits of the FIPE2 checkpoint data involved review of both the pair-selection data (PAIRSEL) and the pair-respondent data (PAIRRESP) in case either indicated that a 12- to 17-year-old was selected or interviewed.

The age of the respondent was taken into account because interviewers were skipped past this checkpoint if respondents were aged 12 to 17. Therefore, the edited version of FIPE2 (SKPX1217) was assigned legitimate skip codes (i.e., 99 if FIPE2 was blank and 89 if FIPE2 was not blank) when the respondent was a youth.

The remaining edits for FIPE2 were implemented when the respondent was an adult. If both PAIRSEL and PAIRRESP indicated that a 12- to 17-year-old was neither selected nor interviewed, it reasonably could be inferred that FIPE2 should have been answered as "no." If FIPE2 was not already answered as "no," the edits assigned a code to SKPX1217 to indicate that a response of "no" was logically inferred. This included situations in which the pair-selection data indicated that a 12- to 17-year-old was not selected, and a completed interview was obtained from only one respondent, who was not aged 12 to 17, regardless of whether PAIRSEL and PAIRRESP were totally consistent. For example, if the pair-selection data indicated that an 18- to 25-year-old and a 26- to 34-year-old were selected, but a single interview was obtained from a 35- to 49-year-old, the pair-selection and pair-respondent data were not totally consistent, but neither would suggest that a 12- to 17-year-old should have been selected. When the edited SKPX1217 indicated that a 12- to 17-year-old was not selected, including situations described above in which the edits inferred that no 12- to 17-year-old was selected, then legitimate skip codes were assigned to the edited variable SKPXPRNT corresponding to FIPE3 (code of 99 if FIPE3 was blank; or 89 if it was not blank).

If either PAIRSEL or PAIRRESP indicated that a 12- to 17-year-old was selected or interviewed, it could be inferred that FIPE2 should have been answered as "yes." Therefore, if FIPE2 was not already answered as "yes," a special code was assigned to SKPX1217 to indicate that a response of "yes" was logically inferred. This included the following situations: (1) PAIRSEL indicated that a 12- to 17-year-old was selected, and PAIRRESP indicated that an interview was obtained from a 12- to 17-year-old, regardless of whether PAIRSEL and PAIRRESP matched exactly (e.g., a 12- to 17-year-old and a 26- to 34-year-old were selected, but interviews were obtained from a 12- to 17-year-old and a 35- to 49-year-old); and (2) PAIRSEL indicated that a 12- to 17-year-old was selected, but a single interview from an adult was obtained at the DU, regardless of whether the adult category from PAIRSEL matched the category in PAIRRESP (e.g., a 12- to 17-year-old and 26- to 34-year-old were selected, but a single interview was obtained from a 35- to 49-year-old). In the latter situation, the respondent result (from PAIRRESP) was not totally consistent with what would be expected based on the pair selection, but PAIRRESP would not provide any information to directly contradict the indication from PAIRSEL that a 12- to 17-year-old was selected.

If PAIRSEL and PAIRRESP disagreed when two people were interviewed, with one indicating the selection or interview of a 12- to 17-year-old but the other variable did not, then special codes were assigned to SKPX1217. When this type of inconsistency occurred, a code of

11 was assigned to SKPX1217 when FIPE2 was originally answered as "yes," and a code of 12 was assigned when FIPE2 was originally answered as "no."

Suppose, for example, that PAIRSEL indicated that a 12- to 17-year-old and a 35- to 49-year-old were selected for the interview, but PAIRRESP indicated that an 18- to 25-year-old and a 35- to 49-year-old were actually interviewed, with the interviewer keying FIPE2 = 1 in the adult's interview (i.e., "yes," a 12- to 17-year-old was selected for an interview at this DU). In this situation, the "yes" in FIPE2 was consistent with who was *selected* (according to the information provided by the screening respondent), but it was not consistent with the ages provided by the respondents themselves. Therefore, the edited variable SKPX1217 would be set to a value of 11 in this example.

This latter edit preserved the information that the interviewer originally entered but also denoted that an inconsistency existed between PAIRSEL and PAIRRESP. This edit also was designed to preserve any possible parenting experiences data when both FIPE2 and FIPE3 (see below) were answered as "yes" but there was an inconsistency between PAIRSEL and PAIRRESP. When an inconsistency occurred between PAIRSEL and PAIRRESP, an analyst would have discretion about whether to use parenting experiences data in an analysis.

Editing of the FIPE3 Checkpoint. This checkpoint pertained to whether the respondent was the parent or legal guardian of the 12- to 17-year-old who also was selected to be interviewed at that DU. Respondents who were aged 12 to 17 were skipped out of FIPE3 and did not have an opportunity to be routed into the parenting experiences section. Therefore, when FIPE3 had been skipped because the respondent was 12 to 17, the edited FIPE3 variable SKPXPRNT was assigned a legitimate skip code.

No further editing of FIPE3 was done when PAIRSEL indicated that a 12- to 17-year-old was selected and PAIRRESP had some result *other than* that of two adults having been interviewed at that DU. The rationale for this approach was that FIPE3 was based on who the actual respondent was, provided that a 12- to 17-year-old was selected. For example, if PAIRSEL indicated that a 12- to 17-year-old and a 26- to 34-year-old were selected, but a 35- to 49-year-old and a 12- to 17-year-old were interviewed, and FIPE3 was answered as "yes" (i.e., this adult respondent is the parent of the youth who was selected), that 35- to 49-year-old respondent may indeed have been a parent or legal guardian of the youth who was selected. This principle also would have held if the selected youth did not respond. Therefore, any data that were present in the parenting experiences section would be preserved.

In contrast, the following situations could occur when FIPE3 was inconsistent with either PAIRSEL or PAIRRESP: (1) PAIRSEL indicates that a youth/adult pair was selected, but two adult interviews were obtained at that DU, or (2) PAIRRESP indicated that a youth/adult pair was interviewed, but PAIRSEL indicated that an adult/adult pair was selected. When either of these inconsistencies occurred, a code of 11 was assigned to SKPXPRNT when FIPE3 was originally answered as "yes," and a code of 12 was assigned when FIPE3 was originally answered as "no."

Suppose, for example, that PAIRSEL indicated that an 18- to 25-year-old and a 35- to 49-year-old were selected for the interview, but PAIRRESP indicated that a 12- to 17-year-old and a

35- to 49-year-old were actually interviewed, and the interviewer keyed FIPE2 = 1 and FIPE3 = 1 in the adult's interview. Stated another way, the interviewer indicated that "yes," a 12- to 17-year-old was selected for an interview at this DU, and "yes," this 35- to 49-year-old respondent was the parent of the 12- to 17-year-old youth who was selected. In this situation, FIPE3 was consistent with PAIRRESP but not PAIRSEL. Furthermore, based on who was interviewed at that DU, the 35- to 49-year-old may indeed be the parent of the 12- to 17-year-old who also was interviewed at that DU. In this situation, the edited SKPXPRT would be set to a value of 11 to denote that this type of inconsistency has occurred. Again, this edit would preserve any possible parenting experiences data—especially in situations in which an adult/child *respondent* pair was obtained.

9.6.2 Editing of the Variables in the Parenting Experiences Section

The variables in the actual parenting experiences section were edited according to the final values assigned to SKPX2PER, SKPX1217, and SKPXPRT based on the edits described in Section 9.6.1. In particular, if these three variables indicated that the respondent was not eligible to be administered the parenting experiences questions, then the edits assigned the appropriate legitimate skip codes to the parenting experiences variables. This included replacing blank values with legitimate skip codes when a code of 12 had been assigned SKPXPRT and the parenting experiences section has been skipped. The rationale for this latter edit was that even if FIPE3 was answered as "no" when PAIRSEL and PAIRRESP were inconsistent, the adult respondent still may not have been the parent or legal guardian of the youth who also was selected for an interview at that DU.

Conversely, if a respondent had been skipped out of the parenting experiences section and the edited FIPE variables SKPX2PER, SKPX1217, or SKPXPRT indicated that the respondent was *potentially* eligible to be administered the parenting experiences questions (i.e., the respondent skipped the section based on the original answers in the FIPE questions, but other data suggested that the respondent may have been eligible to be asked these questions), then the edited parenting experiences variables retained a value of "blank." For example, if FIPE2 had been keyed as "no" and it was inferred for SKPX1217 that a 12- to 17-year-old was selected (i.e., SKPX1217 = 3), then FIPE3 and the parenting experiences questions also would have been skipped. In this situation, the respondent's eligibility or ineligibility to be administered the parenting experiences questions could not be determined because the field interviewer (FI) was not routed to the final checkpoint. Therefore, it could not be determined whether the respondent should have been asked the parenting experiences questions or should have been skipped.

As noted previously, the interview has included a "lockout" feature since 2003 that does not allow interviewers or respondents to go back into the ACASI sections and change their answers once that section of the interview had been completed. However, interviewers could go back to the beginning of the interview after respondents had been administered the ACASI sections and change FIPE1 through FIPE3 in a manner that made the final value in FIPE3 inconsistent with the presence of data in the parenting experiences section. Specifically, FIPE3 ("Is this respondent the parent or legal guardian of the 12- to 17-year-old child who was selected for an interview?") could be answered as something other than "yes," with at least some data existing in the parenting experiences section. According to the CAI logic, however, the section

was to be administered only when the interviewer indicated that the respondent was the parent or guardian of the selected 12- to 17-year-old (i.e., FIPE3 = 1).

The following is an example of a scenario where parenting experiences data could exist when FIPE3 was not answered as "yes." If the FI initially answered FIPE1 through FIPE3 as "yes" (i.e., two people were selected at that DU, a 12- to 17-year-old was selected at that DU, and the respondent is the parent/guardian of that 12- to 17-year-old), the respondent would be routed through the ACASI parenting experiences questions. In the household roster section of the "back-end" demographics section, however, a "hard error" would be triggered if the second person selected to be interviewed was not identified in the roster. Before the interview could proceed, the interviewer would need to change the information in the household roster to make it consistent with the information in FIPE1 through FIPE3, or else the interviewer would need to go back and change the information in FIPE1 through FIPE3 to make it consistent with the roster. In particular, an interviewer could resolve this inconsistency by going back and changing one of the answers in FIPE1 through FIPE3 from "yes" to "no." Because the interviewer and respondent are locked out of ACASI, however, the parenting experiences data would be saved as it had been entered originally. This type of scenario was reproducible in testing of the CAI instrument and yielded a result where FIPE3 was not answered as "yes" but data existed in one or more questions in the parenting experiences section.

If interviewers changed the value in FIPE3 to "no," this would not present a problem in editing the parenting experiences data because the corresponding edited variable SKPXPRNT indicated that the respondent was not the parent or legal guardian of the youth who was selected for an interview. Consequently, parenting experiences data were edited to infer that these respondents should have legitimately skipped the parenting experiences section. Any data that existed in the parenting experiences section were overwritten with a code of 89 (or 9989, etc.).

However, changes made by the interviewer to FIPE1 through FIPE3 that resulted in SKPXPRNT having a final value of 98 (i.e., blank) were more problematic. Because the edits for SKPX2PER, SKPX1217, and SKPXPRNT were consistent with the pair-selection and pair-respondent data (PAIRSEL and PAIRRESP, respectively), the decision was made to retain the value of 98 in SKPXPRNT. Therefore, any nonblank values that existed in the parenting experiences section were replaced with codes for bad data.

One set of variables involved skip logic within the parenting experiences section. Specifically, respondents were skipped out of question PE04 (length of most serious discussion about the dangers of tobacco/alcohol/other drug use) when question PE03 had a value of 1 (i.e., talked with child 0 times in the past year about the dangers of tobacco/alcohol/other drug use), or if PE03 was answered as "don't know" or "refused." Standard procedures for assigning legitimate skip codes (Section 2.4.2) or propagating refusal codes (Section 2.4.3) were implemented in the edited version of question PE04 (PXSERDIS) depending on the response in PE03 (edited variable P XKIDYR).

Parents were asked to report the birth date of the youth who was selected for an interview at that DU (question PE01). However, the birth year that respondents could enter for the youth in question PE01 was restricted to ages that would be more consistent with selection of a 12- to 17-year-old (but also allowed for birth dates that would include 18-year-olds, in case a 17-year-old

respondent just recently had a birthday). Thus, respondents were prevented from entering birth dates that would be extremely inconsistent with selection of a 12- to 17-year-old (such as entry of the current interview year for the birth year).

A refinement has been implemented for the parenting experiences edit logic since 2003 to take into account the situation in which all remaining parenting experiences questions had been skipped because the respondent did not provide a date of birth for the selected youth in question PE01 or did not provide an age for the youth in question PE01B. Prior to 2003, the skipped parenting experiences variables were assigned legitimate skip codes. Since 2003, the parenting experiences variables have retained codes for "blank" when this pattern occurred. The effect of this refinement in 2003 was to make the frequencies of legitimate skip values in PXCHCIG (corresponding to question PE02) and subsequent parenting experiences variables agree with the total count of codes of 2 or 12 (i.e., "no"), or 89 or 99 (i.e., legitimate skip) in SKPXPRT.

The CAI program also calculated an age for the youth who was selected for an interview based on the youth's date of birth (as reported by the parent) and the interview date at the start of the parenting experiences section. Respondents were asked to confirm this age (question PE01A). If parents did not confirm the age that the CAI program calculated for the youth, they were asked to provide a corrected age for the youth who was selected for an interview (question PE01B). Similarly, if respondents did not know or refused to report the date of birth of the selected youth, they were asked to report an age in question PE01B without having to indicate the youth's date of birth.

This information was captured in the created variable PXCHLDAG. Specifically, PXCHLDAG contained the age based on the reported date of birth for the youth and the interview date (if respondents confirmed that this age was correct), or else PXCHLDAG contained the age supplied by the respondent from question PE01B. If respondents supplied a corrected age for the youth in question PE01B that was between 12 and 18 and it mismatched the age of the youth that was calculated from the birth date and interview date information, the edited variables containing the birth date information for the youth (PXBMONTH, PXBDAY, and PXBYR) were assigned bad data values. If respondents answered question PE01B as "don't know" or "refused" when they were asked to provide a corrected age for the selected youth, that response of "don't know" or "refused" was assigned to PXCHLDAG. In addition, if respondents answered question PE01A as "don't know" or "refused" when they were asked to confirm the age of the youth who was selected for the interview, the age for the youth that the CAI program had previously calculated was retained in PXCHLDAG (see above). When values in parenting experiences variables had been set to bad data because SKPXPRT was blank, PXCHLDAG also was assigned a code for bad data.

A recoded variable (PXCMPAGE, for "compare age") also was created that compared the selected youth's age (from PXCHLDAG) with the respondent's age for the second interview that was conducted at that DU. If two interviews were obtained at that DU and a 12- to 17-year-old was selected for an interview, then PXCMPAGE was calculated as the absolute value of the difference between PXCHLDAG and the actual age of the second respondent, within defined categories (i.e., 0 year difference in ages; 1 year difference in ages; 2 year difference in ages; 3 to 4 year difference in ages; and 5 or more year difference). If the adult respondent answered "don't know" or "refused" to the question about the youth's date of birth, or if the youth's date of birth

information was set to bad data because of invalid dates, these codes were reflected in PXCMPAGE.

For the large majority of cases where an interview was obtained from a 12- to 17-year-old, PXCMPAGE indicated no difference between the age based on the date of birth reported by the parent and the youth's age that was recorded in the second interview at that DU. Nevertheless, information about more extreme differences in ages as recorded by PXCMPAGE (e.g., a difference of 2 or more years between the two ages) could be used by analysts in deciding whether to use the parenting experiences data in an analysis. When the second interview was from an 18-year-old, PXCMPAGE was assigned a value of 18. When the second interview was from an adult older than age 18 (i.e., and the parent was supposed to be reporting about a 12- to 17-year-old), the edit program assigned a code of 50 to PXCMPAGE. Again, these codes were designed to give analysts discretion in using or disregarding parenting experiences data when the second interview at a DU came from an adult.

If a 12- to 17-year-old was supposed to be selected at a given DU but only the adult was interviewed, PXCMPAGE was assigned a code of 93. This code was assigned because there were no data to corroborate the youth's date of birth reported by the parent.

If the edited FIPE variables from above indicated that the respondent was not eligible to be administered the parenting experiences questions, then PXCMPAGE was assigned a code of 99 (i.e., legitimate skip). That included situations in which the edited FIPE3 was assigned a code of 12 because of an inconsistency between PAIRSEL and PAIRRESP, and the parenting experiences section had been skipped (see above). Otherwise, if the parenting experiences section was all blank or if PXCMPAGE was undefined for some other reason, then PXCMPAGE was assigned a code of 98. This code of 98 in PXCMPAGE meant "other missing." This code of 98 also was applied in PXCMPAGE when SKPXPRNT was blank and parenting experiences data had been replaced with codes for bad data.

10. Editing of Self-Administered Sections for Physical and Mental Health

10.1 Introduction

This chapter discusses editing procedures for the following self-administered sections that pertain to physical and mental health:

- health,
- adult mental health service utilization (administered to adults aged 18 or older),
- adult mental health (administered to adults),
- adult or adolescent depression (separate sections administered either to adults or to adolescents aged 12 to 17, respectively), and
- youth mental health service utilization (administered to adolescents).

None of the edited variables in this chapter were imputed to remove missing values.

These interview sections are included in this chapter according to their related topical content rather than in the sequential order in which they appeared to respondents during the interview. The order in which these sections were presented to respondents relative to other self-administered sections of the interview can be found in the computer-assisted interviewing (CAI) specifications for the 2016 National Survey on Drug Use and Health (NSDUH) that are available at <https://www.samhsa.gov/data/>.

10.2 Health

The health section in the 2016 NSDUH covered a wide variety of health-related topics:

- pregnancy status for females aged 12 to 44,
- height and weight,
- utilization of medical services in the past 12 months,
- screening for substance use or talks with a health care professional about substance use in the context of health care visits in the past 12 months, and
- specific health conditions.

10.2.1 Pregnancy Status

Female respondents aged 12 to 44 were asked whether they were currently pregnant and, if so, the number of months that they had been pregnant, as well as lifetime and past year histories of specific health conditions.

An important aspect of processing the pregnancy variables for females aged 12 to 44 involved assignment of legitimate skip codes, where relevant (Section 2.4.2). For example, males of any age and women over the age of 44 were assigned legitimate skip codes to the pregnancy variables. Similarly, if females aged 12 to 44 reported that they were not currently pregnant (PREGNANT = 2), legitimate skip codes were assigned to the variable pertaining to the number of months that they were pregnant (PREGMOS).

In the pregnancy variables, if women reported currently being pregnant, the allowable range for the number of months that they were pregnant ranged from 1 to 9 months. Thus, women who reported that they were currently pregnant were not allowed to report that they had been pregnant for "0" months.

10.2.2 Height and Weight

The 2016 NSDUH included questions for respondents to report their height and weight. Respondents were given four options for reporting their height: (1) in feet and inches, (2) in meters and centimeters, (3) in inches only, or (4) in centimeters only. The edited variable HTANSWER was a two-level variable that indicated whether respondents preferred to report their height in (1) English units (feet and inches or inches only) or (2) metric units (meters and centimeters or centimeters only).

Respondents could choose to report their weight in pounds or in kilograms. If female respondents aged 12 to 44 reported that they were currently pregnant (Section 10.2.1), they were asked to report their weight before they became pregnant. As for HTANSWER, the edited variable WTANSWER was a two-level variable that indicated whether respondents preferred to report their height in (1) English units (pounds) or (2) metric units (kilograms).

A single variable for height in inches (HTINCHES) was created from respondents' reports of their height. If respondents chose to report their height in feet and inches, their reports were converted to inches (e.g., 5 feet, 10 inches = 70 inches). If respondents chose to report their height in meters and centimeters, their answers first were converted to centimeters (e.g., 1 meter, 50 centimeters = 150 centimeters). Reported heights in centimeters (including heights in meters and centimeters that were converted to centimeters) were converted to heights in inches using the conversion factor of 1 centimeter = 0.393700787 inches.

Since 2014, the allowable ranges for reporting heights in inches (questions HLTH06A or HLTH06B) reflect whether respondents preferred to report their height in feet and inches or only in inches. Similarly, the allowable ranges for reporting heights in centimeters (questions HLTH08A or HLTH08B) reflect whether respondents preferred to report their height in meters and centimeters or only in centimeters. For example, if respondents chose to report their height in feet and inches, respondents could choose a number of feet ranging from 2 to 8 and a number of inches ranging from 0 to 11. Respondents who chose to report their height only in inches could choose a height ranging from 24 to 96 inches.

The allowable range for reporting weight in pounds (questions HLTH10 and HLTH13) was 40 to 999 pounds. Similarly, the allowable range for reporting weight in kilograms

(questions HLTH12 and HLTH14) was 18.00 to 999.00 kilograms. As a consequence, codes for missing data for respondents' final weights were four digits in length.

HTINCHES could have missing values based on the types of responses that are described below.

- If respondents had missing data for their preferred way of reporting their height (edited variable HTANSWER), remaining questions for height were skipped. The corresponding code for missing data was assigned to HTINCHES (Section 2.2.1). For example, if HTANSWER = 94 (don't know), then HTINCHES was assigned a code of 994 because heights of 90 or more inches were allowed in the data.
- If respondents preferred to report their height in feet and inches or in meters and centimeters, it was possible for them to report the major unit for their height (e.g., feet) but to have a missing value for the minor unit (e.g., inches). The corresponding code for missing data was assigned to HTINCHES. For example, if a respondent reported being 5 feet tall but did not know the number of additional inches beyond 5 feet for his or her height, then HTINCHES was assigned a code of 994.

A single variable for weight in pounds (WTPOUNDS) was created from respondents' reports of their weight (or the weight of pregnant females before they became pregnant). Reported weights in kilograms were converted to weights in pounds using the conversion factor of 1 kilogram = 2.20462262 pounds.

As for HTINCHES, relevant codes for missing data were assigned to WTPOUNDS according to the type of nonresponse. For example, if WTANSWER had a code of 94 (don't know), then WTPOUNDS was assigned a code of 9994. As noted previously, codes for missing data in WTPOUNDS were four digits in length to account for the allowable ranges for weight in the 2016 survey.

10.2.3 Utilization of Medical Services

Prior to 2013, the NSDUH questionnaire included questions about respondents' utilization of hospital emergency room services and overnight inpatient hospitalizations. Questions about respondents' outpatient medical visits in the past 12 months have been included in the questionnaire since the 2013 NSDUH, along with questions about use of other medical services that were described previously in this paragraph.

An important aspect of processing the health care utilization variables involved assignment of legitimate skip codes where relevant (Section 2.4.2). Specifically, respondents who did not report that they were hospitalized overnight in the past 12 months (edited variable INHOSPYR) were not asked for the number of times they were hospitalized in that period (edited variable NMNGTHSP). If respondents reported that they were not hospitalized overnight in the past 12 months (INHOSPYR = 2), the variable NMNGTHSP was assigned a legitimate skip code. If respondents refused to report whether they were hospitalized overnight in the past 12 months (INHOSPYR = 97), that refusal was propagated onto NMNGTHSP (Section 2.4.3).

In addition, question HLTH19 (edited variable NMVSOPT), regarding the number of times that respondents visited health care professionals in outpatient settings, was a continuous variable. If respondents did not know or refused to report the number of times they visited health care professionals in outpatient settings in the past 12 months, they were asked in a follow-up question (HLTH19DK) to give their "best guess" of the number of outpatient visits that they had in the past 12 months. To encourage respondents to report a number of visits, HLTH19DK was categorical and included ranges for respondents to report their number of visits (e.g., 2–3 times, 4–5 times). HLTH19DK also had response categories of 0 ("I have not visited a health care professional in the past 12 months") and 1 (i.e., 1 time). If NMVSOPT indicated a specific number of outpatient visits in the past 12 months, then the edited variable NMVSOEST corresponding to HLTH19DK was assigned a legitimate skip code.

For the questions about the number of nights that respondents were inpatients in a hospital or the number of outpatient visits in the past 12 months, allowable values that respondents could report ranged from 1 to 366. No editing was done to the variables NMNGTHSP or NMVSOPT when respondents reported that they had spent 365 nights in a hospital in the past 12 months, or that they visited a health care professional in an outpatient setting 365 or 366 times, respectively. In addition, no editing was done if the sum of the number of nights as an inpatient and the number of outpatient visits in the past 12 months was greater than 365. However, if respondents reported spending 366 nights as an inpatient in a hospital in the past 12 months, NMNGTHSP was set to 365. Because NMVSOPT refers to the number of outpatient *visits* rather than the number of days, no editing was done to NMVSOPT if respondents reported 366 outpatient visits in the past 12 months.

10.2.4 Screening for Substance Use or Talks about Substance Use with a Health Care Professional

In addition to questions being included since the 2013 NSDUH for outpatient health care visits in the past 12 months, questions have been included in this section since 2013 regarding whether respondents who had health care visits in the past 12 months were asked by a doctor or other health care professional about their tobacco, alcohol, or other substance use in the context of these medical visits (i.e., screening for substance use), regardless of whether they reported use of tobacco, alcohol, or illicit drugs. Respondents also were asked questions about talks that they had with a doctor or other health care professional about their use of tobacco, alcohol, marijuana, or other illicit drugs if they reported health care visits in the past 12 months and they reported use of tobacco products, use of alcohol, or use of specific illicit drugs (i.e., marijuana, cocaine, crack cocaine, heroin, hallucinogens, inhalants, or methamphetamine) (for HLTH23).

Table 10.1 lists these questions, the names for the corresponding edited variables, and descriptions for these questions. The wording for question HLTH20C changed from respondents being asked by a doctor or other health care professional about respondents' use of "illegal drugs" in 2015 to their use of "marijuana or illegal drugs" in 2016. Investigation of the effect of this change to HLTH20C supported but did not conclusively prove that this change affected the comparability of the data from HLTH20C between 2015 and 2016. Because real changes in the population could not be ruled out, the edited variable corresponding to HLTH20C was not renamed for 2016.

Table 10.1 Mapping of Edited Variables for Health Care Screening or Talks about Substance Use with a Health Care Professional

Question	Edited Variable	Description
HLTH20A	HPUSETOB	Doctor asked in person or on a form in the past 12 months if the respondent smokes cigarettes or uses any other tobacco products.
HLTH20B	HPUSEALC	Doctor asked in person or on a form in the past 12 months if the respondent drinks alcohol.
HLTH20C	HPUSEDRG	Doctor asked in person or on a form in the past 12 months if the respondent uses marijuana or other illegal drugs.
HLTH21	HPQTTTOB	Doctor advised the respondent in the past 12 months to quit smoking cigarettes or quit using other tobacco products.
HLTH22, response option 1	HPALCAMT	Doctor asked the respondent in the past 12 months how much the respondent drank alcohol.
HLTH22, response option 2	HPALCFRQ	Doctor asked the respondent in the past 12 months how often the respondent drank alcohol.
HLTH22, response option 3	HPALCPRB	Doctor asked the respondent in the past 12 months whether the respondent had any problems because of his or her drinking.
HLTH22, response option 4	HPALCCUT	Doctor advised the respondent in the past 12 months to cut down on his or her drinking.
HLTH22, response option 5	HPALCTX	Doctor offered to give the respondent more information about alcohol use and treatment for alcohol use problems in the past 12 months.
HLTH22, response option 95	HPALCNOT	Doctor did not discuss the respondent's alcohol use with the respondent in the past 12 months.
HLTH23	HPDRGTALK	Doctor discussed use of specific illicit drugs with the respondent in the past 12 months.

NOTE: HLTH22 is an "enter all that apply" question. Therefore, separate variables were created for each response option. Illicit drugs for HLTH23 referred to marijuana, cocaine, crack, heroin, hallucinogens, inhalants, or methamphetamine.

Respondents were defined as having had health care visits in the past 12 months if any of the following occurred:

- the respondent reported in question HLTH16 (edited variable NMERTMT) that he or she had been treated in a hospital emergency room one or more times in the past 12 months;
- the respondent answered question HLTH17 (edited variable INHOSPYR) as "yes," indicating that the respondent stayed overnight or longer as an inpatient in a hospital in the past 12 months; or
- the respondent reported in question HLTH19 (edited variable NMVSOPT) or in question HLTH19DK (edited variable NMVSOEST) that he or she had one or more visits to a health care professional in an outpatient setting in the past 12 months.

Thus, an important aspect of processing the variables pertaining to screening for substance use and talks about substance use with health care professionals in health care settings involved assignment of legitimate skip codes where relevant (Section 2.4.2). Further details are

provided elsewhere in this section for assignment of legitimate skip codes, depending on the questions.

10.2.4.1 Screening for Substance Use

Eligibility for being asked questions HLTH20A, HLTH20B, and HLTH20C was based only on reports of health care visits in the past 12 months, regardless of whether respondents used tobacco, alcohol, or other substances. HPUSETOB, HPUSEALC, and HPUSEDRG were assigned a code of 99 (Legitimate skip) if respondents reported (1) no hospital emergency room visits in the past 12 months (i.e., NMERTMT = 0), (2) no overnight hospitalizations in the past 12 months (i.e., INHOSPYR = 2), and (3) no outpatient visits in the past 12 months (i.e., NMVSOPT = 0 or NMVSOEST = 0). HPUSETOB, HPUSEALC, and HPUSEDRG retained missing values if these questions had been skipped because respondents did not report having any health care visits in the past 12 months but (1) NMERTMT or INHOSPYR had missing values (i.e., "don't know" or "refused") or (2) both NMVSOPT and NMVSOEST had missing values.

10.2.4.2 Talks about Tobacco Use

Respondents who had health care visits in the past 12 months and reported *past year* use of cigarettes, smokeless tobacco, or cigars, or *past month* use of pipe tobacco (i.e., there is no question in NSDUH to identify respondents who used pipe tobacco more than 30 days ago but within the past 12 months) were eligible to be asked question HLTH21 (HPQTTTOB), regarding whether a doctor or other health care professional advised these respondents to quit smoking cigarettes or to stop using other tobacco products.^{155,156} Data in HPQTTTOB for 2016 were edited according to this skip logic as described below.

- If question HLTH21 had been skipped and respondents *never* used cigarettes, smokeless tobacco, cigars, or pipe tobacco, then HPQTTTOB was assigned a code of 91, where 91 = NEVER USED TOBACCO PRODUCTS. This code was assigned regardless of whether respondents reported any health care visits in the past 12 months.
- If question HLTH21 had been skipped because respondents did not report health care visits in the past 12 months and the most recent use of any tobacco products was more than 12 months ago (which could include reports that respondents never used some

¹⁵⁵ In 2013, respondents who had health care visits in the past 12 months and reported *lifetime* use of tobacco products were eligible to be asked HTLH21. For documentation of the procedures for editing HPQTTTOB based on the logic in the 2013 NSDUH, see the editing and imputation report in the 2013 NSDUH methodological resource book (Center for Behavioral Health Statistics and Quality, 2015).

¹⁵⁶ Although separate questions about the use of snuff and chewing tobacco were replaced in 2015 with questions about the use of any smokeless tobacco products, the variable corresponding to question HLTH21 retained the name HPQTTTOB because the questions did not change for cigarettes, cigars, or pipe tobacco. In particular, most users of tobacco products in the past 30 days in 2015 used cigarettes; see the report on key substance use and mental health indicators (Center for Behavioral Health Statistics and Quality, 2016b).

types of tobacco products),¹⁵⁷ then HPQTT0B was assigned a code of 93, where 93 = DID NOT USE TOBACCO IN THE PAST 12 MONTHS.

- If question HLTH21 had been skipped because respondents had no health care visits in the past 12 months but they used tobacco products in the past 12 months,¹⁵⁸ then HPQTT0B was assigned a code of 99, where 99 = LEGITIMATE SKIP. Thus, in the editing procedures, precedence was given to assigning codes of 91 or 93 to HPQTT0B ahead of assigning a code of 99.

HPQTT0B retained missing values if HLTH21 had been skipped because of the following:

- respondents reported health care visits in the past 12 months but it was unknown whether respondents were users of tobacco products in the past year; or
- respondents reported past year use of one or more tobacco products but it was unknown whether respondents had any health care visits in the past 12 months.

10.2.4.3 Talks about Alcohol Use

For question HLTH22, respondents who had health care visits in the past 12 months and reported past year use of alcohol were eligible to be asked the question. Because HLTH22 was an "enter all that apply" question, separate variables were created that corresponded to the response options in HLTH22. When HLTH22 was skipped, variables corresponding to HLTH22 were edited as described below.

- If question HLTH22 had been skipped and respondents *never* used alcohol, then the variables corresponding to HLTH22 were assigned a code of 91, where 91 = NEVER USED ALCOHOL. This code was assigned regardless of whether respondents reported any health care visits in the past 12 months.
- If question HLTH22 had been skipped because respondents last used alcohol more than 12 months prior to the interview date, then the edited variables corresponding to HLTH22 were assigned a code of 93, where 93 = DID NOT USE ALCOHOL IN THE PAST 12 MONTHS. Again, this code was assigned regardless of whether respondents reported any health care visits in the past 12 months.
- If question HLTH22 had been skipped because respondents had no health care visits in the past 12 months but they used alcohol in the past 12 months, then the edited variables corresponding to HLTH22 were assigned a code of 99, where 99 = LEGITIMATE SKIP. Thus, in the editing procedures for HLTH22, precedence was given to assigning codes of 91 or 93 ahead of assigning a code of 99.

¹⁵⁷ For brevity, this includes reports that respondents did not use pipe tobacco in the past 30 days and there were no reports of past year use of other tobacco products.

¹⁵⁸ For brevity, this includes reports that respondents used pipe tobacco in the past 30 days.

The edited variables for HLTH22 retained missing values if HLTH22 had been skipped because of the following:

- respondents reported health care visits in the past 12 months but it was unknown whether respondents were past year alcohol users; or
- respondents reported past year alcohol use but it was unknown whether respondents had any health care visits in the past 12 months.

In addition, the last category in question HLTH22 allowed respondents to report that a doctor or other health professional did not discuss the respondents' alcohol use with them during any health care visits that they had in the past 12 months (HLTH22 = 95). The CAI logic did not allow respondents to choose this last response if they chose any of the preceding five response options for discussions about alcohol use, or vice versa (Table 10.1). If this last response option was chosen in HLTH22, then the edited variable HPALCNOT corresponding to this response option retained a code of 95. The edited variables HPALCAMT through HPALCTX were assigned a code of 99 (LEGITIMATE SKIP). If respondents chose one or more of the first five response options in HLTH22, then HPALCNOT was assigned a code of 99.

The edited variables for question HLTH22 were examples of variables with the new codes for "enter all that apply" questions that were described in Section 2.4.4 in Chapter 2. If respondents chose any of the responses for categories 1 through 5 from HLTH22, then these responses that were chosen were assigned a code of 1 (Yes). Other responses from the list that were not chosen were assigned a code of 2 (No [not entered]).

10.2.4.4 Talks about Illicit Drug Use

Respondents were eligible to be asked question HLTH23 (HPDRGTALK) if they had health care visits in the past 12 months and they reported past year use of any of the following: marijuana, cocaine (including crack), heroin, hallucinogens, inhalants, or methamphetamine. For cocaine, heroin, and methamphetamine, lifetime users who did not report past year use in the corresponding substance use sections were eligible to be asked HLTH23 if they reported past year use of these substances in the special drugs section (e.g., use of cocaine with a needle in the past 12 months) (Section 7.2). When HLTH23 was skipped, HPDRGTALK was edited as described below.

- If question HLTH23 had been skipped and respondents *never* used any of the illicit drugs that were described in the preceding paragraph, then HPDRGTALK was assigned a code of 91, where 91 = NEVER USED MJ/COC/HER/HALL/INH/METH. This code was assigned regardless of whether respondents reported any health care visits in the past 12 months.
- If question HLTH23 had been skipped because the most recent use of any of these illicit drugs was more than 12 months ago (which could include reports that respondents never used some of these types of illicit drugs), then HPDRGTALK was assigned a code of 93, where 93 = DID NOT USE MJ/COC/HER/HAL/INH/METH IN THE PAST 12 MONTHS.

- If question HLTH23 had been skipped because respondents had no health care visits in the past 12 months but they used any of these illicit drugs in the past 12 months, then HPDRGTALK was assigned a code of 99, where 99 = LEGITIMATE SKIP. Thus, in the editing procedures for HPDRGTALK, precedence was given to assigning codes of 91 or 93 ahead of assigning a code of 99.

HPDRGTALK retained missing values if HLTH23 had been skipped because of the following:

- respondents reported health care visits in the past 12 months but it was unknown whether respondents were users of any of these illicit drugs in the past year; or
- respondents reported past year use of one or more illicit drugs but it was unknown whether respondents had any health care visits in the past 12 months.

10.2.5 Specific Health Conditions

Beginning in 2015, respondents were asked about the following:

- sexually transmitted diseases in the past 12 months (question HLTH24), and
- the occurrence of the following diseases or conditions at any point in the respondent's lifetime (question HLTH25):
 - any kind of heart condition or heart disease;
 - diabetes or sugar diabetes;
 - chronic bronchitis, emphysema, or chronic obstructive pulmonary disease, also known as COPD;
 - cirrhosis of the liver;
 - hepatitis B or C;
 - kidney disease, not including bladder infection or incontinence;
 - asthma;
 - HIV or AIDS;
 - cancer or a malignancy of any kind; or
 - hypertension, also called high blood pressure.

Respondents who reported in HLTH25 that they ever had cancer were asked to choose from a list to report what kind of cancer(s) they had (question HLTH26), including other forms of cancer not on the list and to report how old they were when that type of cancer was first diagnosed (e.g., question HLTH28M for when their lung cancer was first diagnosed). Some of the cancers listed in HLTH26 were specific to males (prostate, testis) or to females (cervix, ovary, uterus), such that respondents were not shown the sex-specific cancers in HLTH26 that were assumed not to apply to them (e.g., females were not shown the response options for prostate or testicular cancer). Respondents who reported having some other kind of cancer were asked to specify the other kind of cancer that they had. Respondents who reported having any

cancer in their lifetime also were asked if they had cancer in the past 12 months (question HLTH29).

Except for hypertension, respondents who reported that they had a specific condition in HLTH25 were asked how old they were when that particular condition was first diagnosed (e.g., question HLTH30 for when the respondent's heart disease was first diagnosed), or in the case of HIV or AIDS, how old they were when they first found out that they had HIV or AIDS. For hypertension, respondents first were asked whether they were currently taking medication for their high blood pressure (question HLTH40); only those respondents who were currently taking medication for their high blood pressure were asked how old they were when their high blood pressure was first diagnosed (question HLTH41). If respondents reported that a particular condition was first diagnosed at an age that was greater than their current age, then the corresponding edited variable was assigned a code for bad data.

In addition, for heart disease, respondents who ever had heart disease were asked whether they had any kind of heart disease or a heart condition in the past 12 months (question HLTH31). Respondents who reported that they ever had asthma were asked in question HLTH38 whether they still had asthma.

The questions for the lifetime occurrence of specific health conditions in HLTH25 or specific cancers in HLTH26 were "enter all that apply" questions, in which respondents could choose more than one condition or more than one type of cancer from the list. Therefore, separate variables were created for each of the response choices for health conditions or specific cancers in HLTH25 and HLTH26 (e.g., `CANCEREVR` for any lifetime cancer in HLTH25 and `CALUNG` for lung cancer in HLTH26). The corresponding edited variables were examples of variables with the new codes for "enter all that apply" questions that were described in Section 2.4.4 in Chapter 2. If respondents chose any of the responses from HLTH25 or HLTH26, then these responses that were chosen were assigned a code of 1 (Yes). Other responses from the list that were not chosen were assigned a code of 2 (No [not entered]). Respondents also had the option to key an answer of "95" in HLTH25 (edited variable `NONABOVEV`) to indicate that they never had any of the health conditions listed in the question.

Data for specific health conditions were assigned codes for bad data for respondents who reported that they had every condition in HLTH25. However, there were no occurrences of this pattern in 2016.

An important aspect of processing the variables for specific health conditions involved assignment of legitimate skip codes, where relevant (Section 2.4.2). Specific edits for assigning legitimate skip codes are described below.

- If `NONABOVEV` = 95 because respondents reported that they never had any of the conditions in HLTH25, then all remaining health condition variables were assigned legitimate skip codes.
- If a specific health condition was assigned a code of 2 but respondents reported other health conditions in HLTH25, then all related variables pertaining to that health condition were assigned legitimate skip codes. For example, if `CANCEREVR` was

assigned a code of 2, then all variables pertaining to cancer were assigned legitimate skip codes.

- If `CANCEREVR` = 1 and respondents did not report having specific kinds of cancer in `HLTH26`, then the variables pertaining to these forms of cancer that respondents did not report having in their lifetime were assigned legitimate skip codes.
- If `CANCEREVR` = 1, variables for kinds of cancer that were specific to males (i.e., prostate, testicular) or to females (cervical, ovarian, uterine) typically were assigned legitimate skip codes according to the respondent's sex. For example, legitimate skip codes were assigned to variables pertaining to prostate cancer and testicular cancer if these forms of cancer were not reported and respondents were female. However, data were not overwritten with logically assigned legitimate skip codes or codes for bad data if respondents reported forms of cancer that did not agree with their current reported sex.
- If respondents reported that they were first diagnosed with cancer or with a heart condition at their current age, they were not asked whether they had the respective condition in the past year. The corresponding edited variables (`CANCERYR` and `HRTCONDYR`, respectively) were assigned a code of 5, where 5 = Yes LOGICALLY ASSIGNED (from skip pattern).

A limited amount of "back editing" of the cancer variables also was conducted based on "OTHER, Specify" data for other forms of cancer (Section 2.4.5). A code of 3 (Yes LOGICALLY ASSIGNED) was assigned to a given cancer variable if a respondent explicitly specified that he or she ever had that form of cancer. For example, if a female respondent did not report having uterine cancer but explicitly specified uterine cancer as some other form of cancer that she had in her lifetime, then the edited variable `CAUTERUS` was assigned a code of 3. Because of the potential complexity in mapping other forms of cancer that respondents specified to specific organs or body systems, however, this editing was limited to situations in which "OTHER, Specify" codes were assigned that corresponded to the specific cancers listed in `HLTH26`. Nevertheless, analysts would have the option to use these "OTHER, Specify" codes in their analyses of NSDUH data for specific types of cancer.

10.3 Adult Mental Health Service Utilization

The section on adult mental health service utilization asked adult respondents about (1) their receipt of specific sources of inpatient or outpatient mental health services in the past 12 months, (2) the length of time that respondents spent in specific inpatient mental health settings or the number of outpatient visits that respondents made to specific types of outpatient mental health providers, (3) payment sources for mental health services, (4) use of prescribed medication for a mental health condition, (5) unmet demand for services (i.e., the respondent felt the need for mental health services but did not receive them), (6) use of alternative sources of treatment (e.g., acupuncture), and (7) how respondents were prompted to get treatment. If the lifetime substance treatment question `TX01` indicated that respondents had received treatment for their use of alcohol or other drugs (or if they answered the question as "don't know" or "refused"), respondents were instructed not to include substance use treatment in their answers about mental health treatment.

Sources of inpatient mental health treatment or counseling that were asked about in the section included (1) a private or public psychiatric hospital, (2) a psychiatric unit within a general hospital, (3) a medical unit within a general hospital, (4) another type of hospital, (5) a residential treatment center, or (6) "some other type of facility." Sources of outpatient mental health treatment or counseling that were asked about in the section included (1) an outpatient mental health clinic or center, (2) the office of a private therapist not associated with a clinic, (3) a doctor's office that was not part of a clinic, (4) an outpatient medical clinic, (5) a partial day hospital or day treatment program, or (6) "some other place." Sources of alternative treatment that were asked about in the section included treatment from (1) an acupuncturist or acupressurist; (2) a chiropractor; (3) an herbalist; (4) an in-person support group or self-help group; (5) an Internet support group or chat room; (6) a spiritual or religious advisor, such as (but not limited to) a pastor, priest, or rabbi; (7) a telephone hotline; (8) a massage therapist; or (9) "other" (i.e., some other source).

An important aspect of processing the variables in this section involved assignment of legitimate skip codes, where relevant (Section 2.4.2). That included (1) assignment of legitimate skip codes to variables in the entire section for respondents who were aged 12 to 17 and (2) assignment of legitimate skip codes to adult respondents' data based on routing logic within the adult mental health service utilization section. For example, if respondents reported that they did not stay overnight or longer in a hospital or other facility to receive mental health counseling in the past 12 months (AUINPYR = 2), all subsequent variables pertaining to inpatient mental health services were assigned legitimate skip codes.

In addition, if respondents did not report receiving treatment in a particular facility or setting in the past 12 months, the questions pertaining to the number of times they received treatment in that setting were skipped. For example, if respondents reported receiving outpatient mental health services in the past 12 months (AUOPTYR = 1) but did not indicate that they received outpatient services in a day treatment program, the edited variable pertaining to receipt of day treatment services (AUOPDTMT) was assigned a legitimate skip code. If respondents in 2016 reported receiving outpatient services in one or more locations from the lists they were provided but they did not report receiving services in "some other place," the edited "OTHER, Specify" variable AUOPYRSP for outpatient services in some other place was assigned a legitimate skip codes.

Similarly, if respondents reported only one source of payment for inpatient or outpatient mental health services, there was no need to ask them who paid for (or would pay for) most of the inpatient or outpatient services that they received. For example, if respondents reported that they received outpatient mental health services in the past 12 months but reported only that private insurance paid for their outpatient mental health services, the edited variable pertaining to the principal payment source (AUOPMOS) was assigned a legitimate skip code.

In questions pertaining to the specific places where respondents received inpatient or outpatient mental health services in the past 12 months, they were allowed to enter more than one place from the list where they received services. Similarly, respondents could select more than one response from lists of payment sources for their inpatient or outpatient services. Information for each of these mental health service locations or payment sources was subsequently captured as a discrete variable. For example, information about receipt of inpatient

mental health services in a psychiatric hospital, the psychiatric unit of a general hospital, the medical unit of a general hospital, another type of hospital, a residential treatment center, or some other type of facility was captured in the variables AUINPSYH, AUINPGEN, AUINMEDU, AUINAHSP, AUINRESL, and AUINSFAC, respectively. These individual edited variables corresponding to these "enter all that apply" questions in the adult mental health service utilization section were coded as 1 or 6 to indicate that the response was entered or not entered, respectively (Section 2.4.4).

Codes of 94 and 97 (for "don't know" and "refused," respectively) were assigned to an entire list of variables if respondents did not know or refused to report what specific places they received mental health services or what specific sources paid (or would pay) for their mental health treatment (Section 2.2.2). If an entire list was blank but respondents had previously reported receiving inpatient services (e.g., if respondents broke off the interview), then the lists of variables pertaining to locations for inpatient services or payment for inpatient services retained a code of 98 (i.e., "blank"); similar logic was applied if respondents reported receiving outpatient mental health services but the location or payment variables were entirely blank.

Adult respondents also were asked if there was any time in the past 12 months when they felt the need for mental health treatment but did not get services. Respondents who answered this question as "yes" then were asked to indicate the reason (or reasons) why they did not get treatment. Based on a review of what respondents had specified in quarter 1 of 2002 as leading "other" reasons for not getting mental health treatment, an additional question (ADMT27A) has been included in the section since 2003 to capture commonly endorsed other reasons for not getting treatment. Therefore, if respondents answered question ADMT26 as "yes" (i.e., perceived the need for mental health treatment), they were routed first to question ADMT27, where they could report up to nine reasons for not receiving mental health treatment, including "some other reason or reasons." If respondents chose the "some other reason or reasons" option in question ADMT27, they were routed next to question ADMT27A, where they could choose additional reasons for not getting mental health treatment; again, respondents were given the option in question ADMT27A to report "some other reason or reasons." If respondents chose this "other" response category in ADMT27A, they were asked to specify the *most important* other reason why they did not get treatment. Respondents since 2003 also have been asked to specify only the most important other reason for not getting mental health treatment.

Questions ADMT27 and ADMT27A were "enter all that apply" questions, in which respondents could choose more than one reason from each list. Each response option (e.g., "You couldn't afford the cost") was captured as a separate variable. The edited variables corresponding to the individual response options in ADMT27 and ADMT27A were coded as 1 (Response entered) or 6 (Response not entered), if at least one item was chosen from the lists in ADMT27 or ADMT27A (Section 2.4.4).

In addition, the ninth response option in question ADMT27 ("some other reason or reasons") was considered principally to be a "toggle" to question ADMT27A. Therefore, separate "some other reason or reasons" variables were not created to correspond to the last response category in ADMT27 and the last category in ADMT27A.

The edits described below have been implemented since 2003 to take into account the data from question ADMT27A.

- If respondents chose the ninth response option in ADMT27, any response that was entered from the ADMT27A series was coded as 1, and anything that respondents did not choose from the ADMT27A list was coded as 6.
- If respondents chose a response from ADMT27 but did not choose the ninth response category in ADMT27, the variables corresponding to the response categories in ADMT27A (i.e., AUUNOND through AUUNSOR) all were given a code of 6 (Response not entered), rather than being assigned "legitimate skip" codes. That is, ADMT27 and ADMT27A were considered together to be one big series of reasons.
- If respondents chose category 9 in question ADMT27, chose at least one reason from ADMT27A, but did not choose category 7 in ADMT27A (some other reason or reasons), the edited variable AUUNSOR (some other reason or reasons for not receiving mental health treatment) was assigned a code of 6 (Response not entered). That is, it was inferred in this situation that the list of specific reasons in ADMT27A was adequate for capturing why respondents did not get mental health treatment. For example, if a respondent chose category 9 in ADMT27 and then chose only category 2 in ADMT27A ("You thought you could handle the problem without treatment"), it would be reasonable to infer that this response in ADMT27A was the only other reason why the respondent did not get treatment.
- If respondents chose response category 9 in question ADMT27, it was possible for them to answer ADMT27A as "don't know" or "refused" (i.e., did not know or refused to report what the other reasons were). When this occurred, the "some other reason" variable AUUNSOR was set to 1 (Response entered) in order to retain information that the respondent chose "some other reason or reasons" somewhere in the series. Remaining variables corresponding to the ADMT27A series retained codes of 94 ("don't know") or 97 ("refused").
- If respondents answered question ADMT27 as "don't know" or "refused," question ADMT27A was skipped. Therefore, the relevant code of 94 or 97 was propagated onto the variables corresponding to the ADMT27A list.
- If AUUNSOR had a value of 6 (see above), the "OTHER, Specify" variable AUUNRIMP was assigned a legitimate skip code. If AUUNSOR had a refusal code, that refusal was propagated onto AUUNRIMP.
- If AUUNSOR had a code of 1 when the respondent answered "don't know" or "refused" to the ADMT27A series, the "OTHER, Specify" variable AUUNRIMP retained a code of 98 (blank).

Consistent with the editing procedures in prior years, if AUUNMTYR indicated that there was not a time in the past 12 months when respondents felt the need for mental health treatment but did not receive services (AUUNMTYR = 2), the edited variables corresponding to questions ADMT27 and ADMT27A were assigned legitimate skip codes. Similarly, if AUUNMTYR was refused, that refusal was propagated onto the skipped variables from questions ADMT27 and ADMT27A.

Also, since 2004, respondents have been shown a list of alternative sources of mental health treatment in question ADMT29A (edited variable AUALTYR) and were asked whether they had received treatment, counseling, or support from other sources such as these in the past 12 months. Respondents who answered ADMT29A as "yes" were asked question ADMT29B, pertaining to specific sources of alternative mental health treatment in the past 12 months. ADMT29B was an "enter all that apply" type of question. Therefore, the individual variables AUALACUP (acupuncturist or acupressurist), AUALCHIR (chiropractor), AUALHERB (herbalist), AUALSGRP (in-person support group or self-help group), AUALINET (Internet support group or chat room), AUALRELG (spiritual or religious advisor, such as a pastor, priest, rabbi), AUALHLIN (telephone hotline), AUALMASG (massage therapist), and AUALOTH (other source) pertaining to the individual sources of alternative treatment in ADMT29B were assigned codes of 1 or 6, as described previously in this section and in Section 2.4.4, when AUALTYR = 1 (i.e., "yes"). When AUALTYR = 2 (i.e., "no"), AUALCHIR through AUALOTH and AUALOTSP (the "OTHER, Specify" variable for other sources of alternative treatment) were assigned legitimate skip codes. When AUALTYR was refused, that refusal was propagated to AUALACUP through AUALOTSP. When AUALOTH was coded as 6 (Response not entered), a legitimate skip code was assigned to AUALOTSP.

Coding of AUALOTSP, regarding other alternative practitioners, was based on information from the National Institutes of Health's (NIH's) National Center for Complementary and Integrative Health (NCCIH) (<https://nccih.nih.gov/health/integrative-health>).¹⁵⁹ NCCIH groups complementary health approaches into the following categories:

- *Natural Products*: Herbalists fall under this type of therapy, along with use of dietary supplements and probiotics.
- *Mind and Body Practices*: Forms of mind and body practices include the following:
 - acupuncture;
 - massage therapy;
 - meditation;
 - movement therapies, such as Pilates and Rolfing Structural Integration;
 - relaxation techniques, such as breathing exercises and guided imagery;
 - spinal manipulation practiced by health care professionals such as chiropractors and osteopaths;
 - tai chi or qi gong practices of movement from traditional Chinese medicine;
 - yoga; and
 - other mind and body practices, such as healing touch and hypnotherapy.

¹⁵⁹ As defined by NCCIH, if a non-mainstream practice is used together with conventional medicine, it is considered "complementary." If a non-mainstream practice is used in place of conventional medicine, it is considered "alternative," although true alternative medicine is rare (i.e., without the use of any conventional medical treatment). "Integrative" health care involves bringing conventional and complementary approaches together in a coordinated way.

- *Other Complementary Health Approaches*: These include traditional healers, Ayurvedic medicine from India, traditional Chinese medicine, homeopathy, and naturopathy.

In addition, the "OTHER, Specify" variable AUOPYRSP (other source of outpatient mental health treatment or counseling) had previously included a code 11 for support groups, self-help groups, or group counseling. However, AUALSGRP pertained only to support groups or self-help groups that were not commonly part of treatment or counseling from a mental health professional. In contrast, group counseling likely would be administered from a mental health professional. For this reason, respondents who reported group counseling were assigned to a category in AUOPYRSP that has been present since 2003 (43 = Group counseling, self-help not specified). Only those respondents who reported receiving treatment from support groups or self-help groups retained a code of 11 for AUOPYRSP. Therefore, documentation for code 11 in AUOPYRSP has read as "Support group/self-help group" since 2004. Respondents who reported in AUOPYRSP that they had received treatment or counseling from a support group or self-help group were logically inferred in AUALSGRP to have received treatment from this source, if AUALSGRP had not already been coded as 1; this issue is described further in [Table B.23](#) in Appendix B.

In subsequent analyses of the adult mental health service utilization data, respondents were not classified as having received outpatient mental health treatment if the only "outpatient" location that they reported was a support group or self-help group. In contrast, group counseling was considered a valid other form of outpatient treatment. Therefore, to facilitate analysis of trends in adults' receipt of outpatient mental health treatment in the past 12 months, the variable AUOPYRSP also was revised in 2003 to reclassify respondents into category 43 if they reported group counseling and to retain a code of 11 in AUOPYRSP only for those respondents who reported receiving services from a support group or a self-help group.

[Table B.23](#) discusses additional edits that were relevant to the adult mental health service utilization variables. For example, respondents could report receipt of outpatient mental health services in "some other place" and then specify a location (e.g., a private therapist's office) that they had not already chosen as a place where they received services. In these situations, respondents were logically inferred to have received services at that location (Section 2.4.5). For example, if respondents had not already indicated that they received outpatient mental health treatment in the office of a private therapist, the edited variable AUOPTHER was assigned a code of 3 (Response entered LOGICALLY ASSIGNED).

10.4 Adult Mental Health

From 2004 to 2007, the mental health section for adults consisted of two primary components. First, a 12-month Kessler-6 (K6) distress scale was administered, and then questions about lifetime and 12-month major depressive episode (MDE) were asked. Since 2008, however, the K6 questions have collected data on distress in the past 30 days and in the past 12 months (see Section B.4.5 in the 2008 national findings report for details) (Office of Applied Studies, 2009). Respondents were routed into the 12-month version of the K6 if they reported having a period in the past 12 months when they felt more depressed, anxious, or emotionally stressed than they felt in the past 30 days.

In 2008, adult respondents also were administered one of two impairment scales—an abbreviated World Health Organization Disability Assessment Schedule (WHODAS) (see Section B.4.6 in the 2008 national findings report for details) (Office of Applied Studies, 2009) or the Sheehan Disability Scale (SDS)—and suicidal ideation questions. These impairment and suicide questions were placed after the K6 questions, but before the MDE questions. A random split-sample design was implemented in 2008 for adults where respondents in sample A were administered the WHODAS scale and respondents in sample B were administered the SDS. All adult respondents were administered the suicidal ideation questions after the impairment items, but before the MDE items. Starting with 2009, the SDS items were no longer included in the mental health section, and the WHODAS has been used to assess impairment.

Apart from propagating refusals from the lead questions for the WHODAS, minimal processing of data was done to the mental health variables. The primary data processing involved assignment of legitimate skip codes for respondents who were aged 12 to 17. Legitimate skip codes also were assigned to data from adults based on other skip logic within the section. If adults broke off the interview before they reached the WHODAS questions, then all of the WHODAS variables retained a code of 98 (i.e., "blank").

10.5 Adult and Adolescent Depression

The adult and adolescent depression sections have been present in the interview since 2004. Questions in these sections were based on those used in the National Comorbidity Survey Replication for adults (NCS-R) and the National Comorbidity Survey Replication Adolescent Supplement (NCS-A) (<https://www.hcp.med.harvard.edu/ncs/>). These depression sections were included in NSDUH to produce lifetime and 12-month prevalence estimates of major depressive episode (MDE), severity of 12-month MDE, age at first MDE, lifetime number of episodes, current and 12-month treatment, and the respondent's perception of treatment effectiveness.

There were some differences in wording between related items in these sections, such as use of simpler wordings for the adolescent depression questions. For example, question AD17 in the adult depression section asked respondents how "severe" their "emotional distress" was during their worst periods lasting 2 weeks or longer when they had problems with their mood. The corresponding question YD17 from the adolescent depression section asked adolescents how "strong" their "bad feelings" were during these periods. Despite these differences in wording, similar naming conventions were used for the variables in these two sections. For example, the edited variable corresponding to question AD17 from the adult depression section was ADWRDST (where WR = worst period, and DST = distress), and the edited variable corresponding to question YD17 from the adolescent depression section was YOWRDST, even though YD17 did not ask about distress. Thus, the only difference in the names for analogous edited variables in these sections was in the use of the two-letter prefix that defined which section a given variable came from: "AD" for variables from the adult depression section and "YO" for adolescent depression variables.

There also were differences in how the CAI program created indicators of MDE for adolescents and adults based on differences in the NCS-A and NCS-R. Specifically, the criteria for defining respondents as having the symptom of loss of interest or pleasure in most things was less restrictive for adolescents than for adults. In particular, the *Diagnostic and Statistical*

Manual of Mental Disorders, fourth edition (DSM-IV; American Psychiatric Association, 1994), criteria for MDE (American Psychiatric Association, 1994) place more emphasis for adolescents on the cognitive aspects of depression, such as boredom or apathy, rather than on somatic or physical complaints, such as sleep loss, that may be manifest in adults with MDE. For example, somatic or physical complaints, such as sleep loss, may be due to factors during adolescence other than depression. Consequently, somatic or physical complaints that may be associated with MDE among adults function less well as indicators of MDE among adolescents than do cognitive indicators. For this reason, the CAI logic gave adolescents additional opportunities to be classified as having the symptom of loss of interest based on their answers to questions that were not taken into account in classifying adults as having this symptom.

Despite these differences between the adult and adolescent depression sections, the basic logic for asking questions was similar between the two sections. Therefore, the remainder of this section discusses edits for both of these sections together. Except where differences are discussed in terms of how variables were edited for these sections, the same basic edits discussed below applied to variables in both sections.

An important aspect of the processing of variables in these sections consisted of assigning legitimate codes based on the routing logic within these sections. In particular, adults were assigned legitimate skip codes to the edited variables in the adolescent depression section, and adolescents were assigned legitimate skip codes to the edited variables in the adult depression section.

As an additional example, respondents were asked a series of questions to identify changes in appetite or weight. They first were asked whether they had a much smaller appetite than usual during the most recent time when their problems were the worst (questions AD26A and YD26A, corresponding to edited variables ADWRELES and YOWRELES, respectively). If respondents answered the relevant question as "yes," they were skipped out of subsequent questions about increases in appetite and weight gain. If respondents reported having less appetite, they were asked whether they lost weight without trying to, and if so, whether their weight loss was due to them being sick or on a diet; respondents who indicated that their weight loss was due to sickness or a diet were not asked to report how many pounds they lost in weight.

Conversely, if respondents did not report in AD26A/YD26A that they had less appetite than usual, they were asked whether they had a much larger appetite than usual (questions AD26B and YD26B, corresponding to edited variables ADWREMOR and YOWREMOR, respectively). Respondents who reported that they had a much larger appetite were skipped out of remaining questions related to weight loss. These respondents subsequently were asked whether they gained weight without trying to. If respondents reported gaining weight without trying to, the CAI program asked follow-up questions to rule out weight gains due to growth (for respondents aged 21 or younger) or pregnancy (for females); respondents who indicated that they gained weight because they were growing or because they were pregnant were not asked to report how many pounds they gained. Thus, editing of the adult and adolescent depression variables related to changes in appetite or weight involved assignment of legitimate skip codes to these variables based on the routing logic for the corresponding questions.

As noted in Section 2.4.3, legitimate skip codes generally were not assigned if a lead question was answered as "don't know" or "refused." However, important exceptions to this

principle were made in editing of the adult and adolescent depression variables because of consideration of other aspects of the routing logic in these sections. In particular, the lead screening questions ASC21 through ASC23 (corresponding to edited variables ADDPREV, ADDSCEV, and ADLOSEV) and YDS21 through YDS23 (corresponding to edited variables YODPREV, YODSCEV, and YOLOSEV) at the beginning of these sections had a special skip logic. If a particular lead question was not answered affirmatively, this logic routed respondents into follow-up questions that could screen respondents into further questions about depression.

In the adult depression section, for example, if question ASC21¹⁶⁰ was answered as "don't know" or "refused," the subsequent question AD01¹⁶¹ was skipped. In this situation, however, respondents were routed to follow-up question ASC22.¹⁶² Thus, if respondents answered question ASC21 "don't know" or "refused" but they answered question ASC22 as "yes," they were still eligible to be administered the remainder of the adult depression section, depending on how they answered subsequent questions.

Therefore, legitimate skip codes were assigned (where relevant) to the variables corresponding to questions AD01 through AD09 in the adult depression section (edited variables ADDPDISC, ADDPLSIN, ADDSLSIN, and ADLSI2WK) if at least one item from questions ASC21, ASC22, or ASC23 was answered as "yes" or "no." For example, if question ASC21 was answered as "don't know," and question ASC22 or ASC23 was answered as "yes" or "no," the editing procedures assigned a legitimate skip code to ADDPDISC, corresponding to question AD01. Similarly, if question ASC22 was answered as "don't know" but ASC23 was answered as "yes" or "no," the editing procedures assigned a legitimate skip code to ADDSLSIN, corresponding to question AD02. In turn, if ADDPDISC, ADDPLSIN, and ADDSLSIN all were answered as "no" or had legitimate skip codes after the above edits, then ADLSI2WK (corresponding to AD09) was assigned a legitimate skip code. The values in ADDPDISC, ADDPLSIN, ADDSLSIN, and ADLSI2WK determined whether subsequent variables were assigned legitimate skip codes. The logic provided in this example for adult depression also was applied to the corresponding variables in the adolescent depression section.

As discussed in Section 2.4.3, if a lead question that governs a skip pattern was refused, the editing procedures typically "propagated" that refusal from the lead question to the variables that had been skipped. For most of the adult and adolescent depression variables, however, this refusal propagation was not performed. The CAI program contained routines for scoring the symptom indicators for MDE. The CAI program coded the symptom score variable DSMMDA2 as 2 if the sum of the numbers of codes of 1 (i.e., has symptom), "don't know," or "refused" in the individual symptom indicators was less than 5. Therefore, not propagating refusals helped to avoid situations in which a different overall score might be obtained if analysts were to calculate DSMMDA2 based on edited variables.

¹⁶⁰ Question ASC21 asked, "Have you ever in your life had a period of time lasting several days or longer when most of the day you felt sad, empty, or depressed?"

¹⁶¹ If respondents answered question ASC21 as "yes," they were asked question AD01. Question AD01 asked, "During times when you felt sad, empty, or depressed most of the day, did you ever feel discouraged about how things were going in your life?"

¹⁶² Question ASC22 asked, "Have you ever had a period of time lasting several days or longer when most of the day you were very discouraged about how things were going in your life?"

The exception to this rule of not propagating refusals in the adult and adolescent depression sections concerns the final questions regarding receipt of counseling from a medical doctor or other professional about the respondents' symptoms of depression (e.g., questions AD86 through AD86F in the adult depression section). For example, if question AD86 (edited variable ADSEEDOC) was refused, the editing procedures still propagated that refusal code to the skipped variables that were dependent on AD86.

In addition, the CAI program created MDE symptom variables and overall MDE symptom scores for adults and adolescents. [Table 10.2](#) lists the final, edited variables that were created from these symptom variables and overall symptom scores. For each variable, explanation of the meaning of that variable also is provided. The only editing that was done to these variables in [Table 10.2](#) was to assign legitimate skip codes based on the respondent's age (i.e., 12 to 17 or 18 or older). Thus, the values that were created by the CAI program were preserved in the variables listed in [Table 10.2](#).

Table 10.2 Depression Symptom and Score Variables

Adult Depression Variable	Adolescent Depression Variable	Explanation
AD_MDEA1	YO_MDEA1	Respondent (R) felt sad, empty, depressed, or discouraged most of the day.
AD_MDEA2	YO_MDEA2	R lost interest or pleasure in most things.
AD_MDEA3	YO_MDEA3	R had changes in appetite or weight (not due to growth, pregnancy, illness, or dieting).
AD_MDEA4	YO_MDEA4	R had sleep problems.
AD_MDEA5	YO_MDEA5	Others noticed that the R was restless or lethargic.
AD_MDEA6	YO_MDEA6	R felt tired or low on energy nearly every day.
AD_MDEA7	YO_MDEA7	R felt worthless nearly every day.
AD_MDEA8	YO_MDEA8	R was unable to concentrate or make decisions.
AD_MDEA9	YO_MDEA9	R was suicidal (had thoughts of suicide, made plans, or made an attempt).
ADSMMDA	YODSMMDA	Score of symptom indicators 1 through 9 from above.

Relatively little additional editing was done to the adult and adolescent depression variables, aside from assigning legitimate skip codes. Additional editing issues that were relevant to these sections are described in the remainder of this section.

If respondents reported a period of time when their symptoms or problems were the worst, they were asked to report how old they were when this time started (edited variables ADWRAGE for adults and YOWRAGE for adolescents). In addition, if respondents scored as positive for MDE (edited variables ADSMMDEA and YODSMMDA) and they reported that these problems caused some interference with their work, social life, or relationships, they were asked to report the age at which these problems first occurred (edited variables ADPBAGE and YOPBAGE). If respondents reported an age of onset in any of these variables that was greater than their current age, these variables were set to bad data. For adults, if ADWRAGE had been set to bad data and the respondent's original answer was age 22 or greater, the respondent was

skipped out of the question that asked about weight gain because the respondent was growing (edited variable ADWRGROW). Therefore, if ADWRAGE had been set to bad data, ADWRGROW retained a code of blank.

As noted previously, if respondents reported gaining or losing weight and these gains or losses could not be attributed to factors other than depression (e.g., growth, pregnancy, dieting), respondents were asked to report the number of pounds they gained or lost. In particular, respondents were allowed to report that they gained or lost 0 pounds. No editing was done to the variables ADWRGNLB, ADWRLSLB, YOWRGNLB, or YOWRLSLB when this response of 0 pounds occurred because respondents did not have an opportunity to report gains or losses of less than 1 pound. Furthermore, only gains or losses of 10 or more pounds resulted in respondents being scored as having the symptom of changes in appetite or weight.

In addition, a feature of the logic for asking respondents about gains or losses in weight was that if respondents were asked the question about gaining weight without trying and they answered it as "don't know" or "refused," they had an additional opportunity to be asked questions about losing weight without trying. That is, the program was looking for the first affirmative set of answers that would allow a determination to be made of whether respondents gained or lost enough pounds to qualify for being depressed. Consequently, no editing was done if respondents originally gave an answer in the questions corresponding to ADWREMOR or YOWREMOR (i.e., having a much larger appetite than usual), they answered the weight gain question (corresponding to ADWRGAIN or YOWRGAIN) as "don't know" or "refused," and then they were routed into the questions about weight loss.

If respondents reported that they talked to a medical doctor or other professional in the past 12 months about the problems they were experiencing related to depression, they were asked to report which professionals they saw or talked to. In the question pertaining to the specific professionals that they saw or talked to, respondents were allowed to enter more than one type of professional from the list they were presented. As in other sections, the "enter all that apply" variables in the adult and adolescent depression sections were coded as 1 if the response was entered and as 6 if the response was not entered (Section 2.4.4).

Codes of 94 and 97 (for "don't know" and "refused," respectively) were assigned to an entire list of variables if respondents did not know or refused to report the specific professionals that they saw or talked to about their problems. If the entire list was blank but respondents had previously reported that they saw or talked to a professional about their problems, then the specific variables corresponding to categories of helping professionals retained a code of 98 (i.e., "blank").

Respondents could report that they saw or talked to "another type of helping professional" and then specify a helping professional that they had already been asked about, such as a psychiatrist. Thus, for example, if the edited variable ADPSYMD, pertaining to services from a psychiatrist, was not coded as 1 and respondents specified that they saw or talked to a psychiatrist, ADPSYMD was assigned a code of 3, where 3 = Response entered LOGICALLY ASSIGNED.

10.6 Youth Mental Health Service Utilization

The section on youth mental health service utilization asked respondents aged 12 to 17 about their receipt of specific sources of inpatient, foster care, outpatient, or school-based mental health services in the past 12 months; the number of nights that respondents spent in specific inpatient or foster care mental health settings; the number of times they visited specific types of outpatient or school-based mental health providers; and the reasons for receiving inpatient, foster care, outpatient, or school-based services for mental health problems the last time they received such services. Specific sources of mental health services that respondents were asked about included (1) any type of hospital, (2) a residential treatment center, (3) foster care or a therapeutic foster home, (4) a partial day hospital or day treatment program, (5) a mental health clinic or center, (6) a private therapist, (7) an in-home therapist, (8) a pediatrician or other family doctor, (9) special education services, and (10) in-school counseling, such as from school counselors or school psychologists.

In 2009, a set of questions from prior years was deleted from the youth mental health utilization section; the deleted questions asked about receipt of special education services and receipt of school counseling in the past 12 months, and whether youths were ever in jail or foster care. In their place, a set of questions has been included since 2009 that measure the same constructs as the questions from prior years but with different wordings. For youths' experiences with the juvenile justice system, the questions since 2009 also focus on the past 12 months rather than the lifetime period and collect additional information about these encounters with the juvenile justice system. The revised questions that have been included since 2009 asked about the following for the past 12 months: (1) receipt of counseling or treatment for mental health problems from social workers, school psychologists, or school counselors; (2) the reason for the last talk with a social worker, school psychologist, or school counselor (if applicable); (3) attendance at a school for students with emotional or behavioral problems or participation in a special program at a regular school for students with emotional or behavioral problems; (4) time spent overnight or longer in a juvenile detention center; (5) the number of nights spent in a juvenile detention center (if applicable); and (6) receipt of treatment or counseling for emotional or behavioral problems in a juvenile detention center (if applicable).

An important aspect of processing the variables in this section involved assignment of legitimate skip codes, where relevant. That included (1) assignment of legitimate skip codes to variables in the entire section for respondents who were aged 18 or older and (2) assignment of legitimate skip codes to youths' data based on routing logic within the youth mental health service utilization section. For example, if respondents reported that they did not stay overnight or longer in a hospital to receive mental health counseling in the past 12 months (YUHOSPYR = 2), all subsequent variables pertaining to mental health services in a hospital were assigned legitimate skip codes. That included the number of nights that respondents stayed in a hospital and the reasons that they were hospitalized the last time.

Although respondents in the youth experiences section who reported that they were not enrolled in school in the past 12 months were asked whether they were homeschooled during this period, the youth experiences variable pertaining to homeschooling (YEHMSLYR, corresponding to question YE09A) was not used to edit youth mental health service utilization variables pertaining to receipt of school-based mental health services. Only the youth

experiences variable pertaining to school enrollment in the past 12 months (YEATNDYR, corresponding to question YE09) was used to edit these school-based service variables.

For each type or location of mental health treatment or counseling that respondents were asked about, they could report that they received services the last time at that particular location for any of the following reasons: (1) they thought about or tried to kill themselves, (2) they felt depressed, (3) they felt very afraid or tense, (4) they were breaking rules or "acting out," (5) they had eating problems, (6) they had trouble controlling their anger, (7) they had gotten into physical fights, (8) they had problems at home or in their families, (9) they had problems with their friends, (10) they had problems with people other than friends or family, (11) they had problems at school, or (12) some other reason. The reasons pertaining to trouble controlling anger through problems at school have been included since 2005 based on identification of commonly reported "other" reasons that respondents specified prior to 2005 for why they received treatment.

For each mental health service location where youths received services, information on these reasons for receiving services subsequently was captured as a discrete variable. For example, if respondents reported receiving mental health counseling from a pediatrician or family doctor, information about why they received counseling the last time was captured in the variables YUFDSUIC (suicidal), YUFDDEPR (depressed), YUFDPEAR (afraid and tense), YUFD BKRU (breaking rules), YUFD EATP (eating problems), YUFD ANGR (anger), YUFD FITE (physical fights), YUFD FMLY (problems in family), YUFD FRND (problems with friends), YUFD OTTP (problems with people other than family or friends), YUFD SCHL (problems at school), and YUFD SOR (some other reason). The individual edited variables for these "enter all that apply" variables were coded as 1 or 6 to indicate that the response was entered or not entered, respectively (Section 2.4.4). No further editing was done if respondents endorsed every single reason on a list as pertaining to why they received mental health services at a given location in the past 12 months.

Codes of 94 and 97 (for "don't know" and "refused," respectively) were assigned to an entire list of variables if respondents did not know or refused to report why they received counseling at a specific location in the past 12 months; this applied as well to the items for the additional reasons why respondents received counseling that were added in 2005. If an entire list of reasons was blank but respondents had previously reported receiving services at a given location (e.g., if respondents broke off the interview), then the list of reasons for receiving services at that location retained a code of 98 (i.e., "blank").

For purposes of illustration, if youths reported in question YSU22 that they had received treatment or counseling from a pediatrician or other family doctor in the past 12 months (i.e., for emotional or behavioral problems that were not caused by alcohol or drugs), they were routed first to question YSU23 regarding the number of times they received treatment from a family doctor, and then they were routed to question YSU24, regarding the reason(s) for their visit; question YSU24 has been present in the section ever since the section was added to the

survey.¹⁶³ As in prior years, YSU24 was an "enter all that apply" item, and individual variables were created corresponding to the specific reasons why respondents got treatment.

Respondents could report up to six reasons in YSU24 for why they received treatment from a family doctor, including "some other reason." If respondents chose the "some other reason" option in question YSU24, they were routed next to question YSU24A, where they could choose the additional reasons for receiving treatment that were noted above, such as difficulty controlling anger; again, respondents were given the option in question YSU24A to report "some other reason." If respondents chose this "other" response category in YSU24A, they were asked to specify the *most important* other reason why they did get treatment in this location; this request for the most important other reason was a noteworthy change relative to prior years (see below).

Based on the logic noted above, the sixth response option in question YSU24 (some other reason) was principally considered to be a "toggle" to question YSU24A. Therefore, a separate "some other reason" variable was not created to correspond to the last response category in YSU24.

The edits described below have been implemented since 2005 to take into account the items on additional reasons why youths received treatment (e.g., YSU24A).

- If respondents chose the sixth response option in YSU24, any response that was entered from the YSU24A series was coded as 1, and anything that respondents did not choose from the YSU24A list was coded as 6.
- If respondents chose a response from YSU24 but did not choose the sixth response category in YSU24, the variables corresponding to the response categories in YSU24A (i.e., YUFDANGR through YUFDSOR) all were given a code of 6 (Response not entered), rather than being assigned "legitimate skip" codes. That is, YSU24 and YSU24A were considered together to be one big series of reasons.
- If respondents chose the sixth category in question YSU24, chose at least one reason from YSU24A, but did not choose category 7 in YSU24A (some other reason), the edited variable YUFDSOR (some other reason for receiving treatment) was assigned a code of 6. That is, it was inferred in this situation that the list of specific reasons in YSU24A was adequate for capturing why respondents saw a family doctor about emotional or behavioral problems in the past 12 months. For example, if a respondent chose category 6 in YSU24 and then chose only category 3 in YSU24A ("You had problems at home or in your family"), it would be reasonable to infer that this response in YSU24A was the only other reason why the respondent visited a family doctor about emotional or behavioral problems.
- If respondents chose the sixth response category in question YSU24, it was possible for them to answer YSU24A as "don't know" or "refused" (i.e., did not know or refused to report what the other reasons were). When this occurred, the "some other reason" variable YUFDSOR was set to 1 (Response entered) in order to retain

¹⁶³ Questions in this section for treatment that youths received in other settings were structured in the same manner as in this example for treatment from a family doctor. The changes to the questions for additional reasons why youths received treatment from a family doctor also applied to the other sources of treatment in this section.

information that the respondent chose "some other reasons" somewhere in the series. Remaining variables corresponding to the YSU24A series retained codes of 94 ("don't know") or 97 ("refused").

- If respondents answered question YSU24 as "don't know" or "refused," question YSU24A was skipped. Therefore, the relevant code of 94 or 97 was propagated onto the variables corresponding to the YSU24A list.
- If YUFDSOR had a value of 6 (see above), the "OTHER, Specify" variable YUFDIMPR (i.e., the most important other reason why the respondent received treatment from a family doctor) was assigned a legitimate skip code. If YUFDSOR had a refusal code, that refusal was propagated onto YUFDIMPR.
- If YUFDSOR had a code of 1 when the respondent answered "don't know" or "refused" to the YSU24a series, the "OTHER, Specify" variable YUFDIMPR (which has been three digits in length since 2005) retained a code of 998 (blank).

Consistent with general editing procedures, if respondents reported a reason that corresponded to a reason in the lists for YSU24/YSU24A, that reason was logically inferred to have been chosen in the relevant edited variable. Suppose, for example, that the most important other reason that a respondent reported for receiving treatment from a family doctor indicated that the respondent was breaking rules or "acting out." If the respondent had not chosen this response in YSU24, the edited variable YUFDBKRU was assigned a code of 3 (Response entered LOGICALLY ASSIGNED). Similarly, if the respondent specified that a problem in the respondent's family was the most important other reason for receiving treatment from a family doctor and the respondent had not chosen this response in YSU24A, the edited variable YUFDFMLY was assigned a code of 3. Conversely, if respondents did not report "some other reason" why they received treatment in the past 12 months from a family doctor (edited variable YUFDSOR = 6, corresponding to response category 6 in question YSU24 not being chosen or response category 6 in question YSU24A not being chosen), legitimate skip codes were assigned to the edited "OTHER, Specify" variable YUFDIMPR (corresponding to question YSU24SP).

As noted previously, the "OTHER, Specify" items in this section underwent an important change in 2005. Prior to 2005, youths were asked to specify "the other reason" why they received services in a particular location or from a given type of provider. Some respondents in these prior years gave a considerable amount of information in the space that was allotted to them to specify their other reason(s) for receiving services. Often, multiple reasons were reported. Therefore, prior to 2005, up to five separate "OTHER, Specify" codes were assigned for a given treatment location or provider based on respondents' explanations regarding why they received services.

Because youths since 2005 have been asked to specify the "most important" other reason why they received services in a particular location, only one reason was captured in the "OTHER, Specify" variables. If respondents specified more than one reason as the "most important" other reason why they received treatment in a given location, only the first reason that respondents specified was coded as a general rule. The exception to this rule was that any reports of respondents thinking about or trying to kill themselves were given precedence in coding, regardless of whether this reason was specified first. In addition, the change in the "OTHER, Specify" variables for this section was used as an opportunity to revise the "OTHER, Specify" codes to regroup related reasons together, such as reasons that indicated that respondents had a

diagnosed condition (e.g., attention-deficit/hyperactivity disorder, or ADHD). For these reasons, the "OTHER, Specify" data from 2005 onward are not comparable with data prior to 2005.

In a relatively rare number of situations, youths denied receiving mental health services as part of an "OTHER, Specify" response. In these situations, the "OTHER, Specify" response was assigned a bad data code. Data were retained that indicated that the youths received mental health services in a given location in the past 12 months.

Respondents could report that the number of nights they stayed overnight in a hospital or residential treatment program in the past 12 months (or the sum of the two, if respondents reported staying in both settings) was greater than or equal to 365 nights. In these situations, no editing was done to the data. If respondents reported spending 366 nights in an inpatient/residential setting or in foster care for treatment or counseling for emotional or behavioral problems, however, the number of nights in the corresponding edited variable was "trimmed" back to 365. For example, if respondents reported staying overnight in a hospital for 366 nights for treatment or counseling for emotional or behavioral problems, the corresponding edited variable YUHOSPNM was set to a value of 365.

11. Editing and Imputation for the NSDUH Roster Variables

11.1 Introduction

This chapter describes the techniques used to edit and impute variables associated with the household roster for the 2016 National Survey on Drug Use and Health (NSDUH). The variables described in this chapter can be divided into three groups:

- respondent-level detailed roster variables,
- roster-derived household composition variables, and
- proxy variables.

The respondent-level detailed roster variables included the age, gender, and relationship to the respondent for each household member. The introductory question for the household roster portion of the questionnaire (QD54) was interviewer administered. This question asked the respondent how many individuals lived in the household. The computer-assisted interviewing (CAI) instrument was set up to be able to collect data on up to 25 household members. If only one person lived in the household or the respondent did not know or refused to answer, then the household composition (roster) section was skipped. Otherwise, the respondent was asked questions about the age, gender, and relationship to the respondent of every member of the household, starting with the household's oldest member and including the respondent.

If a pair of respondents was selected in a household, the interviewer indicated which member of a respondent's household roster corresponded to the other selected pair member. The roster entry for the respondent was referred to as the "self" entry. In effect, the respondent completed a grid with the number of rows corresponding to the value entered in QD54. [Table 11.1](#) shows an example grid where the number of individuals in the household is four. In this example, the roster of the respondent is shown, and the indicator variable shows that the respondent's son was selected as the other pair member. The possible relationship codes between pair members are listed in [Tables 11.2](#) and [11.3](#).

In 2016, the questionnaire levels describing the relationships of household roster members to the respondent were updated and expanded to a more detailed list of relationships. The relationships captured in the 2015 survey year and prior years are shown in [Table 11.2](#). The relationship codes new to the 2016 survey year are shown in [Table 11.3](#) (i.e., relationship levels for sibling-in-law, aunt or uncle, niece or nephew, cousin, and ex-spouse). The new detailed relationship levels were captured by the new detailed roster relationship variables ROSRLTD1-ROSRLTD25 (detailed relationship to respondent). To maintain comparability with previous survey years, ROSRLT1-ROSRLT25 (relationship to respondent) were created by recoding the new relationship level to match the original 14 relationship levels captured in 2015 and prior survey years.

An additional edited proxy variable, PRXRELATD, was also created in 2016 to capture the newly available detailed relationship codes when identifying the relationship of the proxy to the respondent. Again, the edited proxy variable from previous survey years, PRXRELAT, was also created to maintain comparability across the years.

Table 11.1 Roster Grid Example Where Number of Individuals in Household (QD54) Equals 4

Individual #	Relationship to Respondent	Age in Years	Other Member Selected for Pair ¹
1	Self	44	0 (No [Impossible])
2	Husband	42	0 (No)
3	Son	16	1 (Yes)
4	Boarder/roomer	16	0 (No)

¹ This indicator variable applied only to respondents who were part of a pair selection. The other member selected could not have been the self because respondents were not interviewed twice. The other member selected was the roster member who had a value of "1" for this variable.

Table 11.2 Roster Relationship Codes Prior to 2016

Level Value	Relationship to Respondent
1	Self
2	Parent
3	Child
4	Sibling
5	Spouse
6	Unmarried partner
7	Housemate or roommate
8	Child-in-law
9	Grandchild
10	Parent-in-law
11	Grandparent
12	Boarder or roomer
13	Other relative
14	Other nonrelative

Table 11.3 Detailed Roster Relationship Codes Beginning in 2016

Level Value	Relationship to Respondent
1	Self
2	Spouse
3	Child
4	Child-in-law
5	Sibling
6	Sibling-in-law
7	Parent
8	Parent-in-law

Table 11.3 Detailed Roster Relationship Codes Beginning in 2016 (continued)

Level Value	Relationship to Respondent
9	Aunt or uncle
10	Niece or nephew
11	Grandparent
12	Grandchild
13	Cousin
14	Ex-spouse
15	Unmarried partner
16	Housemate or roommate
17	Tenant, boarder, or exchange student
18	Other relative
19	Other nonrelative

The second group of variables, the roster-derived household composition variables, was derived from the respondent-level detailed roster variables. These were mostly count variables reporting the number of individuals in the roster with various characteristics. These variables are listed in [Table 11.4](#), which also shows that some of these variables underwent imputation. Among the three groups of variables described in this chapter, only the roster-derived household composition variables underwent imputation of any kind.

Table 11.4 Roster-Derived Household Composition Variables

Variable Description	Edited Variable Name	Imputation-Revised Variable Name
Total number of rostered individuals	TOTPEOP	IRHHSIZE
Number of individuals in household aged 17 or younger	KID17	IRKID17
Number of individuals in household aged 65 or older	HH65	IRHH65
Indicator of whether the respondent had family members in household	FAMSKIP	IRFAMSKIP
Number of respondent's family members in household (includes foster relationships)	FMLYSIZE	IRFMLYSZ
Number of respondent's family members in household aged 17 or younger (includes foster relationships)	KIDFMLY	IRKDFMLY
Number of respondent's children in household aged 2 or younger	NRBABIES	N/A
Number of respondent's children in household aged 3 to 5 years old	NRPRESCH	N/A
Number of respondent's children in household aged 6 to 11 years old	NRYUNGCH	N/A
Number of respondent's children in household aged 12 to 17 years old	NRTEENS	N/A
Number of respondent's children in household aged 17 or younger	NRCH0_17	N/A
Number of respondent's children in household aged 18 to 20 years old	NROLDRCH	N/A

Table 11.4 Roster-Derived Household Composition Variables (continued)

Variable Description	Edited Variable Name	Imputation-Revised Variable Name
Number of respondent's children in household aged 21 or older	NROLDCH	N/A
Number of roommates/housemates in household	NROOMATE	N/A
Indicator of presence of mother in household (12- to 17-year-olds) ¹	IMOTHER	N/A
Indicator of presence of father in household (12- to 17-year-olds) ¹	IFATHER	N/A

¹ The IMOTHER and IFATHER indicators were not 0/1 indicators because levels were provided for "unknown" and "18 or older."

The third group of variables, the proxy variables, allowed for the selection and identification of a relative of the respondent who lived in the respondent's household (according to the household roster), was aged 18 years or older, and answered the health insurance coverage and income questions for the respondent. The edited versions of these variables and the questions to which they map are shown in [Table 11.5](#). In 2016, an additional edited variable called PRXRELATD was created to capture the detailed relationship response levels available for QP02. The previously existing edited variable, PRXRELAT, was also created to maintain comparability with previous survey years.

Table 11.5 Proxy Variables

Unedited Variable	Text of Survey Question Associated with Unedited Variable	Edited Variable
QP01	Is there anyone else who lives here who is 18 or older who would be better able to give me the correct information about your health insurance coverage and the kinds of income you receive?	PRXABLE2
QP02	Who is the person you think can help us get the correct information for these questions?	PRXRELAT PRXRELATD
QP03	Is your [QP02 fill] available right now?	PRXHOME2
QP04	Would you ask your [QP02 fill] to join us to help with these last questions about health insurance and income?	PRXJOIN2
HASJOIN	Has the person's [QP02 fill] joined R?	PRXYANS2
QP03a	Is there any other adult family member available who might be able to answer these questions?	PRXRETRY

11.2 Editing the Respondent-Level Detailed Roster Variables

This section describes the methods used to create edited versions of the respondent-level detailed roster variables: ROSAGE1-ROSAGE25 (roster age), ROSSEX1-ROSSEX25 (roster gender), ROSRLT1-ROSRLT25 (relationship to respondent), ROSRLTD1-ROSRLTD25 (detailed relationship to respondent), ROSMSL1-ROSMSL25 (0/1 indicator: other member selected, pair members only), and TWNTYP1-TWNTYP25 (type of twin: identical, fraternal, or neither). These variables describe up to 25 members of the household. The editing procedures for the respondent-level detailed roster variables began with consistency checks included in the Blaise program code, which were implemented to reduce the amount of editing required at the

data processing stage. The consistency checks in the questionnaire were supplemented with other edits involving the respondent-level detailed roster variables outside the CAI instrument. These involved resolving cases where it was unclear which roster member was the self and cases where relationship codes were impossible (or very unlikely) given the age and gender in relation to the self.

Section 11.2.1 describes the consistency checks programmed into the questionnaire. Section 11.2.2 describes the creation of a roster-level dataset for further processing. Section 11.2.3 describes roster edits involving the self. Section 11.2.4 describes roster edits for other household members, after the self has been established. Finally, Section 11.2.5 describes the creation of the final edited respondent-level detailed roster variables. In 2016, the roster edits were performed using the relationship codes described in [Table 11.3](#). Although the detailed codes introduced in 2016 allow for the possibility of further developing the household roster editing checks, this will be introduced in 2017 to have a full year of survey data to evaluate changes.

11.2.1 Roster Consistency Checks

Two types of consistency checks were employed in the CAI instrument for the household roster section of the questionnaire. These checks (1) compared the roster entry corresponding to the respondent with previously entered questionnaire information and (2) compared a roster entry against other roster entries for internal consistency. With the exception of the check against the previously entered respondent's gender, the interviewer could override the consistency checks and explain why the response given was correct. Interviewers' explanations for overrides to consistency checks and evaluations of their validity are provided in Appendix G.

11.2.1.1 Comparisons with Previously Entered Questionnaire Information

Gender and age were the two consistency checks built into the household roster section of the CAI instrument that compared the roster entry with the previously entered questionnaire information. The check for gender was added in 2001 and was triggered if the respondent in the household roster entered a gender that was different from the one previously recorded in the interview (question QD01). If the gender did not match, the interviewer was required to change either the roster entry or the gender that had been entered at the beginning of the interview.

The consistency check for age was added in 2002 to compare the respondent's age in the roster with the age previously entered in the questionnaire (the Blaise variable CURNTAGE). The interviewer could either change the respondent's age entered in the roster or override the consistency check. If the interviewer chose to override the consistency check, then he or she provided an explanation as to why the roster age did not match CURNTAGE. Explanations given by the interviewer for overriding this particular consistency check were carefully reviewed. Interviewers' explanations for overrides to consistency checks and evaluations of their validity are provided in Appendix G. In rare cases, the final value for age (AGE) was set to the age of the self in the household roster (the "roster age") based on these explanations as well as other evidence. Additional details about how roster age was used for creating AGE are described in Chapter 4. Strategies for the more common situation, where the original value for AGE was not set to the roster age, are discussed in Section 11.2.3.

11.2.1.2 Internal Consistency Checks

Since the 2002 survey, internal consistency checks have been implemented in the CAI instrument to compare one roster entry with another. These checks were triggered if any of the following conditions occurred:

1. The interviewer reported that the respondent had more than one spouse or unmarried partner or reported a spouse and an unmarried partner.
2. The interviewer reported that a household member was a grandparent of the respondent and the respondent was older than or the same age as the household member.
3. The interviewer reported that a household member was a grandchild of the respondent and the respondent was younger than or the same age as the household member.
4. The interviewer reported that a household member was a spouse, an unmarried partner, or an ex-spouse of the respondent and the household member was 16 years old or younger. In 2016, ex-spouse was added to this check in the CAI.
5. The interviewer reported that the respondent had a spouse, unmarried partner, or ex-spouse and the respondent was 16 years old or younger. In 2016, ex-spouse was added to this check in the CAI.
6. The interviewer reported that the respondent was either a child-in-law or a parent-in-law and the respondent was 16 years old or younger.
7. The interviewer reported that a household member was a child-in-law of the respondent and the household member was the same age or older than the respondent.
8. The interviewer reported that a household member was a parent-in-law of the respondent and the household member was the same age or younger than the respondent.
9. The interviewer reported that a household member was a grandparent or grandchild of the respondent and the age difference was less than 30 years.
10. The interviewer reported that only one member of a pair was available for interviewing, potentially due to a change in household size.

In most cases, if a consistency check was triggered, the interviewer changed either an age code or a relationship code in the roster to a more appropriate value. Any edit that was invoked because of an override to a consistency check was carefully scrutinized during the data processing stage. The relevant household roster, as well as the explanation given by the interviewer for the override, was carefully examined to determine whether the override was legitimate. If the override was deemed legitimate (e.g., a father marries a woman, listed as mother, who is younger than the respondent), the original answer was allowed to remain and no edit was applied. If the interviewer's explanation was not considered legitimate, then an edit was applied. More details about roster edits are provided in Section 11.2.4. Explanations given by the interviewers for the overrides that were not considered legitimate are provided in Appendix G.

11.2.2 Roster-Level Dataset

To facilitate processing of the roster variables, a roster-level dataset was created in which the number of records per respondent was given by the household size in question QD54. For example, if a respondent indicated a household of three consisting of himself or herself, a mother, and a father, then there would be three records on the dataset associated with this respondent: one for the self, one for the mother, and one for the father. Even if the respondent did not start or complete the roster questions, records were created for each household member.

11.2.3 Roster Edits Involving the Self

The Blaise program code required the interviewer to identify exactly one self and a corresponding age and gender in the household roster. Moreover, the interviewer was required to confirm with the respondent that the respondent was in fact the identified self. Because the check involving gender was not allowed to be overridden, the gender for the self in the roster always matched QD01, which was equivalent to IRSEX (see Chapter 4). However, it was possible to have problems matching AGE (see Chapter 4 for a description of the methodology used to create AGE) with the age of the self in the roster, despite the consistency check comparing the respondent's roster age against CURNTAGE.

The interviewer was able to override the consistency check for age of the self for one of two reasons: (1) the self was misidentified and another roster member was the true self but the interviewer decided not to change the entries; or (2) the interviewer correctly identified the self but indicated that the correct age for the respondent was different than CURNTAGE, and other evidence did not support this claim (AGE was not set to the roster age, as discussed in Section 11.2.1.1). In the case of a misidentified self, a second roster member in the household was selected whose gender matched IRSEX and whose age was within 1 year of AGE. The second roster member who replaced the original self had a gender and age that matched IRSEX and AGE, respectively.

If the consistency check was overridden, a misidentified self was diagnosed if (1) the roster age of the self differed from AGE by more than 1 year and (2) another roster member of the same gender as QD01 (and IRSEX) had a roster age within 1 year of AGE.¹⁶⁴ If a misidentified self was diagnosed, it was assumed that the interviewer used the roster member identified as the self, rather than the respondent, as the point of reference. Using the example shown in Table 11.1, if the respondent's son was used as the reference point, the relationship for the respondent became "mother" instead of "self," and the "husband" became "father." Under these circumstances, the self code was set to missing, and the respondent's roster entries did not include a self. The remaining relationship codes in the roster also were set to missing. In some cases, the original relationship codes were salvaged, depending upon the roster member who was used as a reference point.

¹⁶⁴ A 1-year difference was allowed because the respondent's age might have changed during the interview. In this instance, the values of AGE and CURNTAGE may have differed by 1 year.

11.2.3.1 Original Self Misidentified: Identifying the Real Self

If the self was misidentified in the roster, an attempt was made to identify a self among the roster members corresponding to the respondent. A roster member was selected as the self under one of two possible circumstances: (1) the roster member's age, gender, and relationship data were missing; or (2) the roster member was of the respondent's gender and was within 1 year of the respondent in age. If more than one roster member met the above criteria, the roster members who met the criteria but were not assigned the self code, were given a bad data code; that is, the original relationship code would no longer make sense because the person of reference in the household had been changed.

11.2.3.2 Original Self Misidentified: Salvaging Relationship Codes

As stated earlier, if the self was misidentified, all other relationship codes were set to missing because the reference person was someone other than the respondent. In some cases, however, the original relationship codes were salvaged, depending upon the roster member who was used as a reference point. Relationship codes were salvaged under the following circumstances:

1. If the reference person was the respondent's sibling, the roster member listed as "self" was actually a sibling, and all other relationship codes were salvaged. (Generally, relationships between the respondent and other household members would be the same with a sibling. For example, the respondent's parents are also the respondent's sibling's parents.)
2. If the reference person was the respondent's spouse or unmarried partner, the roster member listed as "self" was actually a spouse or unmarried partner, and the children relationship codes were salvaged.
3. If all the roster members other than the misidentified self were either roommates, boarders, or other nonrelatives, then the reference person was the respondent's roommate, boarder, or other nonrelative. All other relationship codes were salvaged.

11.2.4 Roster Edits for Other Household Members

Relationship codes were edited if the relationship of the roster member to the self was logically impossible based on age and gender. Edits of household roster ages, genders, and/or relationship codes were performed that either changed the reported value to another value or changed the reported value to bad data. It is important to note that in some cases, two members were selected in a household, which greatly increased the ability to edit the roster for those respondents. Some edits were associated with consistency checks, and interviewers' explanations for overrides to these consistency checks were carefully examined to assess the legitimacy of the override as explained in Section 11.2.1. Some edits were "automatic" in the programming code, which meant that the interviewer was assumed to have been incorrect when the override was implemented. These edits were undone if the interviewer's explanation for the override was considered legitimate. In other situations, the default strategy was to assume that the override of the consistency check was correct and, therefore, that the edit was applied only if the interviewer's explanation appeared incorrect.

Since 2015, another editing method was implemented to consider logical family relationships and age ranges without regard to relationship type (e.g., biological, step, foster, adopted). Most of these edits did not result in an "automatic" correction in the programming code; they relied on a review of the entire household roster and corresponding pair member's roster (if present) to apply an edit if supported by other evidence. In 2016, both the new detailed relationship codes shown in [Table 11.3](#) and the 14-level relationship codes shown in [Table 11.2](#) were edited for these cases to maintain consistency between ROSRLT1-ROSRLT25 and ROSRLTD1-ROSRLTD25. This ultimately resulted in a better edit for cases where "other relative" would have previously been the only available option. For example, a roster may have been flagged for multiple parents present, and the corresponding pair member's roster identified one of the extra parents as an aunt or uncle. This person would still be identified as an "other relative" in ROSRLT1-ROSRLT25 but would be identified as an aunt or uncle in ROSRLTD1-ROSRLTD25.

11.2.4.1 Edits to Roster Age, Gender, and Relationship Codes: Changes to Different Values (Correct Person of Reference)

The following edits were performed on the roster age, gender, and relationship code values when the recorded age, gender, and/or relationship code was either missing or internally inconsistent and replaced by internally consistent values. When typing on a computer keyboard, it was possible for a double-digit age to have been entered as a single-digit age ("5" instead of "55"), or vice versa ("55" instead of "5"). If the relationship code was still believable, even with the incorrectly entered age (e.g., "other relative"), this type of error was difficult to detect because no inconsistency check was triggered. On the other hand, if an age entered this way triggered one of the consistency checks discussed in Section 11.2.1.2, the interviewer had an opportunity to correct the entry error. On those occasions where the age did not trigger a consistency check, detection of the error was still possible among selected pairs by examining the roster entries of the other pair member. For example, if one pair member had an x-year-old and no xx-year-olds, and the other had an xx-year-old and no x-year-old, where x denoted a single-digit number, it was highly probable that an error had occurred. By comparing the number of children younger than 12 years old in each roster with the number of children on the screener roster, it was apparent how a correction should be made. In this instance, the incorrectly entered age was replaced with the value given by the pair member whose roster age and screener age agreed.

1. If two members were selected in a household, the roster age for the other member selected was commonly not the same as the questionnaire-edited age (AGE, defined in Chapter 4) of the other pair member. In this case, the roster age for the other member selected was changed to this questionnaire-edited age value.
2. If two members were selected in a household, the gender that one member selected for the other on the household roster was often not the same as the gender (IRSEX, defined in Chapter 4) reported by that other pair member in his or her interview. In this case, the roster gender was changed to match the gender value the other pair member reported in his or her interview.
3. In previous survey years, the relationship codes for grandchild (9) and grandparent (11) were commonly confused. The following edit, which was used in previous

survey years, was maintained in case of overrides: If the age of the respondent was at least 20 years older than that of the roster member, but the roster member was identified as a grandparent, the relationship code was changed to grandchild. Conversely, if the age of the respondent was at least 20 years younger than that of the roster member, but the roster member was identified as a grandchild, then the relationship code was changed to grandparent.

11.2.4.2 Edits to Relationship Codes: Changes to Missing Codes

The following edits were performed on the roster relationship code values, where the relationship code given was internally inconsistent and no internally consistent value could be used to replace it. These edits were performed before the edits listed in Section 11.2.4.1 were completed. For respondents who had changes to their rosters that were due to the edits described below, the changes to age and gender that were due to the edits described in Section 11.2.4.1 were checked to make sure that they did not impact the decision to implement the edits below. The relationship code in these instances was set to a bad data code. During editing in 2016, the new detailed relationship codes were output in addition to the original relationship codes in cases that were flagged to gain more insight into the household composition.

1. More than one roster member aged 15 years or older was listed as the respondent's unmarried partner or as the respondent's spouse. This situation should have been covered by consistency check #1 listed in Section 11.2.1.2.
2. A roster member aged 15 years or older was identified as a spouse and another was identified as an unmarried partner. In this case, the spouse code was maintained and the unmarried partner code was set to bad data. This situation should have been covered by consistency check #1 listed in Section 11.2.1.2.
3. The roster member was the respondent's parent, but was younger than the respondent. This situation should have been covered by consistency check #2 listed in Section 11.2.1.2. This edit would have been automatic for respondents younger than 15 years old.
4. The roster member was the respondent's child, but was older than the respondent. This situation should have been covered by consistency check #3 listed in Section 11.2.1.2. This edit would have been automatic for respondents younger than 15.
5. The roster member was the respondent's parent, but was less than 12 years older than the respondent.
6. The roster member was the respondent's mother, but was more than 60 years older than the respondent.
7. The roster member was the respondent's child, but was less than 12 years younger than the respondent.
8. A respondent had a sibling older than the youngest parent. If this situation occurred, the relationship code of the "sibling" was set to bad data.
9. A respondent had a parent younger than the oldest sibling in a household. If this situation occurred, the relationship code of the "parent" was set to bad data.

10. The roster member was the respondent's child-in-law, but was at least 10 years older than the respondent. This situation should have been covered by consistency check #7 listed in Section 11.2.1.2.
11. The roster member was the respondent's parent-in-law, but was at least 10 years younger than the respondent. This situation should have been covered by consistency check #8 listed in Section 11.2.1.2.
12. The roster member was the respondent's parent-in-law or child-in-law, but either the roster member or the respondent was younger than 15 years old. This situation should have been covered by consistency check #6 listed in Section 11.2.1.2.
13. The respondent had two or more children-in-law, but had no children in the household. The in-law codes were all set to missing.
14. The roster member was the respondent's grandchild, but the respondent or respondent's spouse (if applicable) was 25 years old or younger. This situation should have been covered by consistency check #9 listed in Section 11.2.1.2.
15. The roster member was the respondent's grandchild, but the respondent's parents lived in the household. Also, the respondent had no children in the household and was less than 24 years older than the roster member. As with the previous edit, if the grandchild was in fact older than the respondent, this error should have been covered by consistency check #3 listed in Section 11.2.1.2.
16. The roster member was the respondent's sibling and there was a 20-year age difference between the roster member and respondent, but the roster member's age was within 4 years of the age of the oldest parent.
17. The roster member was the respondent's grandparent or grandchild, but the age difference between the respondent or the respondent's spouse (if applicable) and the roster member was less than 20 years. If the roster member was a "grandchild" who was older than the respondent, then this situation was covered by consistency check #3 listed in Section 11.2.1.2. Similarly, if the roster member was a "grandparent" who was younger than the respondent, then this situation was covered by consistency check #2 listed in Section 11.2.1.2. If the age difference was less than 30 years, this was covered by consistency check #9 in Section 11.2.1.2.
18. If the respondent had more than two parents, the roster relationships of all parental roster members to the respondent were set to bad data.

11.2.4.3 Edits to Relationship Codes: Changes to Different Values (Incorrect Person of Reference: Illogical Child Code)

In Section 11.2.4.2, illogical relationship codes were set to bad data. Often, this occurred because the interviewer used someone other than the respondent as the person of reference for one or more roster members. In some of these cases, the structure of the roster could have been used to determine the appropriate relationship code for that individual. Edits where the illogical code was "child" are listed below.

1. The interviewer might have put a roster member after the respondent's parent in the household roster. If the relationship code for that roster member was given as "child," the relationship code was illogical if the age made it impossible for the roster member to be the respondent's child (see #4 in Section 11.2.4.2). However, it was likely that the interviewer was making the reference to the respondent's parent rather than the respondent. In this case, if the age difference between the respondent's parent and the "child" was at least 12 years, then the relationship code was changed to sibling.
2. In some cases, the interviewer's entry for a roster member listed as "child" might simply be a typographical error, for example, where the "3" (child) should be a "2" (parent) (see [Tables 11.2](#) and [11.3](#) for all the relationship codes).

If the respondent was living with parent(s) and unmarried and not living with an unmarried partner, and the roster member was not 12 or more years younger than the respondent, then the relationship code was changed to sibling.

3. Both sides in a selected pair¹⁶⁵ were respondents aged 18 or younger, both sides identified parents in the household, and one side had an illogical child code. When the number of illogical child codes was added to the number of siblings on one side, the sum was equal to the number of siblings on the other side. If the age of the roster member was younger than 25 years, then the relationship code was changed to sibling.
4. A roster member was listed as the respondent's child who was not more than 12 years younger than the respondent and the respondent was 25 or younger. The previous roster member was listed as "grandparent." The "child" was in reference to the respondent's grandparent and was considered either the respondent's parent or the respondent's aunt or uncle. If the roster member's age was at least 12 years older than the respondent and there were no nonimmediate family codes (7, 12, 13, or 14 as described in [Table 11.2](#)), then no aunts or uncles lived in the household. If a pair was selected and no nonimmediate family codes were found in either pair member's roster, then in either of these cases the relationship code was set to parent. Otherwise, the relationship code was set to missing.

11.2.4.4 Edits to Relationship Codes: Changes to Different Values (Incorrect Person of Reference: Illogical Spouse Code)

The interviewer also could have used an incorrect person of reference with illogical spouse codes. This error occurred most frequently when a selected child had a parent with a spouse (the other parent) or unmarried partner. Rather than identifying this individual as a "parent" or "other nonrelative," the interviewer identified the roster member as a spouse or unmarried partner of the child, even though the interviewer intended that the point of reference be the child's parent rather than the child. This manifestation of the illogical spouse code, along with others, is described below. Many of these edits were covered by consistency checks #4 and

¹⁶⁵ A selected pair has two rosters where each respondent is from the same household. A "side" refers to one of the two rosters that make up a selected pair.

#5 listed in Section 11.2.1.2, provided either the respondent or the roster member was 16 or younger.

1. Both sides in a selected pair identified that they had a spouse or unmarried partner, but the two respondents were not part of a spouse-spouse pair. This legitimately could have occurred only if there were multiple spouse-spouse pairs in the household. In this edit, an attempt was made to identify cases with a single spouse-spouse pair in the household, where one pair member had a correctly identified spouse or unmarried partner and the other pair member had an incorrectly identified spouse or unmarried partner. If the younger respondent, who was 21 years old or younger and at least 10 years younger than the older respondent, indicated a parent, and the older respondent indicated neither parents nor parents-in-law, then the older respondent should be considered either the younger respondent's parent or the parent's spouse or unmarried partner. If the misidentified code was "spouse," then the code was changed to "parent." However, if the misidentified code was "unmarried partner," then the roster member may or may not be considered the parent of the respondent. In most cases where the misidentified unmarried partner was the respondent's parent's unmarried partner, the code was changed to parent. The exception occurred when (1) the unmarried partner of this respondent's parent was the other respondent selected in a pair and (2) the unmarried partner did not indicate that the other pair member selected was his or her child in the parenting experiences question, FIPE3. In this instance, the relationship code was changed to indicate that the roster member was an unmarried partner of the respondent's parent.
2. As in the previous edit, both sides in a selected pair identified a spouse or unmarried partner, but were not part of a spouse-spouse pair, and there was only a single spouse-spouse pair in the household. In this edit, both sides incorrectly identified the spouse or unmarried partner. In most cases, the pair was a sibling-sibling pair. If both respondents were younger than 21, both indicated a parent in the household, and the age difference between the respondents and their respective "spouse or unmarried partner" was unusually large, then on each side the misidentified spouse or unmarried partner should have been considered a spouse or unmarried partner of the respondent's parent. If both misidentified codes were "spouse," then both codes were changed to "parent." As stated in the previous edit, if both misidentified codes were "unmarried partner," then it was not clear whether each misidentified code should have been "parent." The rules used to determine whether the roster member was the respondent's parent were the same as in edit #1. Hence, the incorrectly identified "spouse or unmarried partner" code was changed for each respondent in the pair to either "parent" or another applicable code.
3. In this edit, only one side in a selected pair identified a spouse (not unmarried partner), but the spouse was identified even though (1) the respondent was younger than 15, (2) the spouse was younger than 15 and the other pair member did not have a spouse, or (3) the respondent was younger than 18 but responded that he or she was "never married" in the marital status part of the questionnaire, and the respondent did not have any parents-in-law in the household. If the respondent listed one parent, but the other pair member listed two parents, then the pair was a sibling-sibling pair and the relationship code was in reference to the parent. If the respondent listed one fewer

- sibling than the other pair member, then the pair was a sibling-sibling pair and the spouse code was a typographical error (meant to be a sibling, with a code of "4" instead of "5").
4. Only one side in a selected pair identified an unmarried partner, but the unmarried partner was identified even though (1) the respondent was younger than 15 or (2) the unmarried partner was younger than 15. If the respondent listed one parent, but the other pair member listed two parents, then the pair was a sibling-sibling pair and the relationship code was in reference to the parent's unmarried partner. In this case, the relationship code was changed to parent. If the respondent listed one fewer sibling than the other pair member and the age difference between the respondent and the roster member identified as the unmarried partner was less than 15 years, then the pair was a sibling-sibling pair and the unmarried partner code was changed to sibling.
 5. Both sides in a pair identified the same household member as spouse or unmarried partner. If the previous roster member on one of the sides was a sibling, then the spouse or unmarried partner should be considered the sibling's spouse or unmarried partner. The spouse or unmarried partner relationship code was changed to bad data. If both sides had a previous roster member who was a sibling, then it was not clear to which pair member the spouse or unmarried partner belonged. To maintain proper counts, the spouse or unmarried partner code for the youngest pair member was changed.
 6. A spouse or unmarried partner was identified even though (1) the respondent had one parent in the household who was the roster member listed before the spouse or unmarried partner, (2) the respondent either was younger than 17 years old or was between 17 and 20 years old and the spouse or unmarried partner was older than the respondent's parent, and (3) the respondent was more than 15 years younger than the spouse or unmarried partner. In the case of the misidentified spouse, the "spouse" of the respondent was considered the respondent's other parent. In the case of the misidentified unmarried partner, the "partner" of the respondent was considered the unmarried partner of the respondent's parent. The code was changed to "parent." For a household member with a spouse code who was aged 16 years or younger, this edit should have been covered by consistency check #4 listed in Section 11.2.1.2.
 7. In cases where the respondent was younger than 15 years old, he or she identified a spouse or unmarried partner, and the above edits did not apply, the relationship code was set to bad data. In cases where the roster member was younger than 15, the roster member was identified as a spouse or unmarried partner, and the above edits did not apply, the relationship code and roster member's age were set to bad data. This should have been covered by consistency checks #4 and #5 listed in Section 11.2.1.2.

11.2.4.5 Edits to Relationship Codes: Changes to Different Values (Incorrect Person of Reference: Illogical Sibling Codes)

If the relationship code indicated that one of the other roster members was the respondent's sibling, but the age difference between the sibling and the respondent was at least 20 years, then the sibling relationship code was suspicious. If the age difference between the sibling and respondent was 25 years or greater and a pair member roster was present, the roster

was output for review. If the age difference between a reported sibling and the respondent was 20 years or more and the respondent's age was 25 years or older, and either the respondent did not have parents in the household or the parent was a mother and the age difference between the mother and the sibling was 50 years or greater, then the sibling relationship codes were referencing the respondent's children's relationships to each other. The relationship codes were therefore changed to "child."

11.2.4.6 Edits to Relationship Codes: Changes to Different Values (Incorrect Person of Reference: Illogical Grandchild Codes)

If the relationship code indicated that one of the other roster members was the respondent's grandchild, but the respondent was too young to have a grandchild (25 or younger), it was possible that the roster member was a grandchild of a previous roster member. If two young respondents were selected where both identified the same grandparents and the same parents, and the respondent on the other side had siblings, then the grandchild should be considered the respondent's sibling. If this was not established, then the roster member could be the respondent's sibling or the respondent's cousin, and the code was set to bad data. If the grandchild was older than the respondent, then this edit would have been covered by consistency check #3 listed in Section 11.2.1.2. If the age difference between the grandchild and the respondent was less than 30 years, then this edit would have been covered by consistency check #9 listed in Section 11.2.1.2.

11.2.4.7 Edits to Relationship Codes: Changes to Different Values (Incorrect Person of Reference: Illogical In-Law Codes)

In some situations, the in-law code was incorrectly used because the respondent was not using himself or herself as the person of reference. In such cases, either the child-in-law was the child of someone else in the roster other than the respondent or the respondent was referring to himself or herself as the parent-in-law of the roster member. An in-law code was deemed incorrect if a roster member was listed as the respondent's child-in-law who was not more than 12 years younger than the respondent and the respondent was 25 or younger. If the relationship code was listed as child-in-law, and the previous roster member was listed as grandparent, then the child-in-law was in reference to the respondent's grandparent and should have been considered either the respondent's parent or the respondent's aunt or uncle. If the roster member's age was at least 12 years older than the respondent and there were no nonimmediate family codes (7, 12, 13, or 14 as described in [Table 11.2](#)), then no aunts or uncles lived in the household. If a pair was selected, no nonimmediate family codes were found in either pair member's roster. In either of these cases, the relationship code was set to parent. Otherwise, no certainty was associated with the relationship code, and this code was set to missing.

11.2.5 Final Edited Respondent-Level Detailed Roster Variables

The unedited roster variables contained information for each roster member: age, gender, relationship to respondent, and a 0/1 variable that indicated whether the roster member was the other member selected in a pair (see [Table 11.1](#) for an example). This information could be captured for up to 25 members of a household. Within the CAI instrument, separate variables were created to collect this information for male and female household members and for

household members with ages reported in years as opposed to months. When the edited versions of these variables were created, this information was combined for each household member into four variables, one for each attribute (i.e., age, gender, relationship to respondent, and pair status). The edits listed in Section 11.2 were incorporated into the values of the detailed roster variables: ROSAGE1-ROSAGE25 (roster age), ROSSEX1-ROSSEX25 (roster gender), ROSRLT1-ROSRLT25 (relationship to respondent), ROSRLTD1-ROSRLTD25 (detailed relationship to respondent), and ROSMSL1-ROSMSL25 (0/1 indicator: other member selected, pair members only). Additional variables created are TWNTYP1-TWNTYP25 (type of twin: identical, fraternal, or neither).

Final edited versions of the respondent-level detailed roster variables were used to derive (or, at a minimum, to calculate bounds when data were missing) the household composition variables described in Section 11.3. In 2016, the variables ROSRLT1-ROSRLT25 were still used to create the relevant edited household composition variables because these variables rely on household roster relationships that exist in both ROSRLT1-ROSRLT25 and ROSRLTD1-ROSRLTD25.

11.3 Editing and Imputation Procedures for Roster-Derived Household Composition Variables

11.3.1 Creation of Edited Roster-Derived Household Composition Variables

This section discusses the creation of edited versions of the roster-derived household composition variables. After replacing apparently erroneous information in the roster with missing values, the number of individuals with various characteristics in each roster was determined. These counts were recorded in the edited roster-derived household variables shown in [Table 11.4](#). If any information in the roster was missing, the roster-derived variable was set to missing. However, if some of the roster records for a respondent's household had missing data, then roster records with nonmissing data for that household were used to limit the possible values to which the missing roster-derived variable could have been imputed. Details on the imputation of the roster-derived household variables are provided in Section 11.3.2. If two respondents were selected in a single household as part of a pair, then the information from one pair member was not used to directly edit the household composition variables of the other pair member.

The respondent's household size was assumed to equal the total number of rostered individuals in the household, TOTPEOP, as shown in [Table 11.4](#). The value of TOTPEOP was expected to equal the value of QD54 in most cases. However, in some cases, the original self was misidentified and no other roster members were close to matching the respondent's age and gender. In these cases, an extra roster member was added to correspond to the respondent (the self) so that the value of TOTPEOP was 1 greater than the value of QD54. For other cases, the respondent did not enter a value for QD54, and thus TOTPEOP and all the roster-derived variables were missing. Finally, it was possible that duplicate entries were put into the household roster so that the value of TOTPEOP would be determined by excluding the duplicates from the roster. This latter situation was usually impossible to detect, unless the respondent had two fathers or two mothers of exactly the same age. In this instance, the extra parent of the same

gender was dropped from the roster, and the value of TOTPEOP was reduced to 1 less than the value of QD54.

The variables KID17 (number of individuals in the household aged 17 or younger) and HH65 (number of individuals in the household aged 65 or older) were simple counts based on the roster ages and did not account for the relationships of the individuals to the respondent. If some of the roster members had missing ages, the values of KID17 and HH65 also were missing, regardless of whether some of the roster members were eligible to be part of the count. In these instances, the imputed values for KID17 and HH65 were restricted based on the nonmissing information available in the roster, as explained in Section 11.3.2.2. However, if the roster member was missing a relationship code, but not an age, then that roster member was still eligible to be included in these variables.

The variable FAMSKIP was an indicator of whether the respondent's household contained any other family members. It was created based on the relationship codes of the roster members. If one or more of the roster members had a missing relationship code, and no other family members were in the respondent's household, then the value of FAMSKIP was set to missing. However, if one of the nonmissing roster member's relationship codes indicated that the household contained one of the respondent's family members, then the value of FAMSKIP was not missing, even if other roster members had missing relationship codes.

The variables FMLYSIZE (number of respondent's family members in the household, including foster relationships) and KIDFMLY (number of respondent's family members in the household aged 17 or younger, including foster relationships) were simple counts based on the relationships of the individuals to the respondent and the ages in the respondent's household roster. FMLYSIZE and KIDFMLY were created to determine appropriate measures of poverty levels, using Federal poverty definitions starting in 2006. If some of the roster members had missing ages or missing relationship codes, the values of FMLYSIZE and KIDFMLY were set to missing, even though some of the roster members might have been eligible to be part of the count. In these instances, the imputed values were restricted based on the nonmissing information available in the roster, as explained in Section 11.3.2.2.

Ten other roster-derived variables were created that used both the age and relationship codes of the roster members. All of the roster-derived variables and their definitions are summarized in [Table 11.4](#). Except for FAMSKIP, each of these variables was missing if the age or relationship codes for at least one roster member in a respondent's household were missing. FAMSKIP could be coded despite missing values if there was at least one nonmissing family relationship code in the roster. Edited versions of the roster-derived variables were also used in the editing procedures applied to the creation of the edited proxy variables, described in Section 11.4.

11.3.2 Imputation for Roster-Derived Household Composition Variables

Of the three groups of edited roster variables described in the introduction to this chapter, the only group that underwent any imputation at all was a subset of the roster-derived household composition variables. Each of the eight variables in this subset formed its own single-member

imputation set¹⁶⁶ and tended to have item response rates of more than 99 percent. Table A.27 in Appendix A has details on the rates of missingness for the variables that were imputed. The single RP/single PRD type of PMN, described in Section 3.4.1, was used to impute nonmissing values among these eight variables in the order shown in Table 11.4. The order was important, as imputation-revised variables from earlier in the sequence were frequently used to assist with imputation-revised variables later in the sequence.

Section 11.3.2.1 describes the imputation process applied to the first edited variable, TOTPEOP. Section 11.3.2.2 summarizes the imputation processes applied to the other seven variables. Since the processes applied to the other seven variables are very similar to the process applied to TOTPEOP, Section 11.3.2.2 will only list divergences from the process that was applied to the TOTPEOP variable.

In 2016, there was a change made to the imputation procedures to account for a donorless case for the imputation-revised measure of the total number of children younger than 18 in a household, IRKID17. The previously imputed values for IRHH65 (imputation-revised adults in the household aged 65 or older) for the donorless case created a bounding inconsistency in the pre-hot-deck step for IRKID17. In 2015 and prior years, if the upper bound derived from the edited variable KID17 was missing, then the upper bound for IRKID17 was set equal to the imputation-revised household size, IRHHSIZE. This bounding did not consider the number of adults or adults aged 65 or older in the household. To remedy this situation and to prevent future bounding inconsistencies, the upper bound for IRKID17 was modified to account for the minimum number of adults that are known to be in the household in addition to the total household size.

11.3.2.1 Imputation for TOTPEOP (Imputation Set 1)

The first imputation set included a single variable, TOTPEOP. The analogous imputation-revised variable IRHHSIZE was created using the single RP/single PRD type. There were no noteworthy deviations from this general approach. Section 11.3.2.1.1 describes the RP step, Section 11.3.2.1.2 describes the PRD step, and Section 11.3.2.1.3 describes the hot-deck step. All the roster-derived household composition variables that underwent imputation had a very high item response rate (more than 99 percent).

11.3.2.1.1 Response Propensity Step

The response propensity model for imputation set 1 utilized the preliminary analysis weight, PANALWT. All respondents were in the domain for the TOTPEOP variable. A domain member was considered an item respondent if and only if TOTPEOP was nonmissing. See Tables D.64, D.65, D.66, and D.67 in Appendix D for details of the covariates used in the RP models for this variable.

¹⁶⁶ An imputation set is a set of variables for which a single donor is used in the final hot-deck step. Chapter 3 describes this concept more fully.

11.3.2.1.2 Prediction Step

TOTPEOP was a count variable. It was assumed to have a Poisson distribution, and the parameters for the models were estimated using the adjusted weights that are outputs of the RP step and using Poisson regression as implemented by the LOGLINK procedure in SUDAAN software.¹⁶⁷ The single predicted mean used in the subsequent hot-deck step was the predicted number of people in the household.

11.3.2.1.3 Hot-Deck Step

The hot-deck step for the TOTPEOP variable was the simplest one used in the NSDUH. There were no logical constraints, and the only likeness constraint was the delta constraint. The predictive mean vector was actually a scalar. Every item nonrespondent was handled on the first attempt to find a donor. Additional details on the hot-deck step for TOTPEOP are available in Appendix E.

11.3.2.2 Imputation for Other Roster-Derived Household Composition Variables That Underwent Imputation (Imputation Sets 2 through 8)

Like TOTPEOP, the remaining seven roster-derived household composition variables that underwent imputation (from [Table 11.4](#): KID17, HH65, FAMSKIP, FMLYSIZE, and KIDFMLY) were handled separately using the single model type of PMN and utilized the preliminary analysis weight, PANALWT, in the RP step. The methods were very similar, with only a few exceptions, as follows:

- FAMSKIP was a dichotomous variable, not a count variable. Therefore, its PRD model was a logistic regression model as implemented by the RLOGIST procedure in SUDAAN. The single predicted mean used in the later hot-deck step was the predicted probability that the respondent did not have any other family members in his or her household.
- Bounds were determined for every other variable except FAMSKIP. These bounds were based both on nonmissing information in the roster and on previously imputed variables. For each of these variables, a single logical constraint was used in the hot-deck step, which required the donor to have a value within the bounds.
- Previously imputed roster-derived household composition variables were frequently used in likeness constraints in the hot-deck steps.

[Tables D.64](#), [D.65](#), [D.66](#), and [D.67](#) in Appendix D provide details of the covariates used in the RP models for these variables.

11.4 Editing the Proxy Variables

This section describes the creation of edited proxy variables, as listed in [Table 11.5](#). Section 11.4.1 describes the creation of an indicator variable, EDFAM18, which was used to

¹⁶⁷ Details about the LOGLINK procedure and additional references are provided in the *SUDAAN® Language Manual, Release 11.0.1* (RTI International, 2013).

determine skip patterns and missing codes for the five edited proxy variables. Sections 11.4.2 and 11.4.3 describe the editing processes for each value of EDFAM18.

All survey respondents were allowed to choose someone from the household to be their proxy as long as the following conditions were met:

1. There was more than one person in the household.
2. The eligible person was a relative (not a boarder, roommate, or some other nonrelative).
3. The eligible person was aged 18 or older.

The correspondence between the six questionnaire items in the proxy section of the questionnaire and the corresponding edited variables are shown in [Table 11.5](#). Except for QP02 and its edited variables PRXRELAT and PRXRELATD, the valid questionnaire responses were "1 = Yes" and "2 = No." QP02 and PRXRELATD had multiple responses ranging from 1 to 31, with each level representing the relationship of the proxy to the respondent. PRXRELATD was introduced in 2016 to capture the detailed relationship codes available as responses to QP02. PRXRELAT is still created for consistency between survey years by recoding the 31 values available from PRXRELATD to the 21 levels available prior to 2016. The levels of PRXRELAT are shown in [Table 11.6](#), and the levels of PRXRELATD are shown in [Table 11.7](#). Depending on household size and composition, QP02 gave respondents a maximum of nine possible proxy choices out of the 31 relationship types detailed by PRXRELATD. In addition to the nine possible answer choices presented by QP02 based on reported household size and composition, respondents were also given a 10th and final choice of "other adult relative" to select a proxy who is related to the respondent but may not live in the household. Among the detailed relationship codes introduced in 2016, roster members reported as sibling-in-law, uncle/aunt, nephew/niece, cousin, or ex-spouse were considered relatives and thus were eligible for selection as a proxy.

Table 11.6 Levels of PRXRELAT

PRXRELAT	Relationship of Proxy Member	Gender of Proxy Member
1 = Father	Parent	Male
2 = Mother	Parent	Female
3 = Son	Child	Male
4 = Daughter	Child	Female
5 = Brother	Sibling	Male
6 = Sister	Sibling	Female
7 = Husband	Spouse	Male
8 = Wife	Spouse	Female
9 = Male Unmarried Partner	Unmarried partner	Male
10 = Female Unmarried Partner	Unmarried partner	Female
11 = Son-in-law	Child-in-law	Male
12 = Daughter-in-law	Child-in-law	Female

Table 11.6 Levels of PRXRELAT (continued)

PRXRELAT	Relationship of Proxy Member	Gender of Proxy Member
13 = Grandson	Grandchild	Male
14 = Granddaughter	Grandchild	Female
15 = Father-in-law	Parent-in-law	Male
16 = Mother-in-law	Parent-in-law	Female
17 = Grandfather	Grandparent	Male
18 = Grandmother	Grandparent	Female
19 = Other Male Relative	Other relative	Male
20 = Other Female Relative	Other relative	Female
21 = Other Adult Relative	Other relative	Male or Female

Table 11.7 Levels of PRXRELATD

PRXRELATD	Relationship of Proxy Member	Gender of Proxy Member
1 = Husband	Parent	Male
2 = Wife	Parent	Female
3 = Son	Child	Male
4 = Daughter	Child	Female
5 = Son-in-law	Child-in-law	Male
6 = Daughter-in-law	Child-in-law	Female
7 = Brother	Sibling	Male
8 = Sister	Sibling	Female
9 = Brother-in-law	Sibling-in-law	Male
10 = Sister-in-law	Sibling-in-law	Female
11 = Father	Parent	Male
12 = Mother	Parent	Female
13 = Father-in-law	Parent-in-law	Male
14 = Mother-in-law	Parent-in-law	Female
15 = Uncle	Uncle/aunt	Male
16 = Aunt	Uncle/aunt	Female
17 = Nephew	Nephew/niece	Male
18 = Niece	Nephew/niece	Female
19 = Grandfather	Grandparent	Male
20 = Grandmother	Grandparent	Female
21 = Grandson	Grandchild	Male
22 = Granddaughter	Grandchild	Female
23 = Male cousin	Cousin	Male
24 = Female cousin	Cousin	Female
25 = Ex-husband	Ex-spouse	Male
26 = Ex-wife	Ex-spouse	Female

Table 11.7 Levels of PRXRELATD (continued)

PRXRELATD	Relationship of Proxy Member	Gender of Proxy Member
27 = Male unmarried partner	Unmarried partner	Male
28 = Female unmarried partner	Unmarried partner	Female
29 = Other Male Relative	Other relative	Male
30 = Other Female Relative	Other relative	Female
31 = Other Adult Relative	Other relative	Male or Female

11.4.1 Edited Indicator of Potential Proxies in Household (EDFAM18)

As described in Section 11.3.1, a binary variable (FAMSKIP) was created that indicated whether the respondent's household roster included other family members. If the presence or absence of other family members was unknown because of a missing household size or missing values in the roster, FAMSKIP could not be determined. A similar variable was created to identify households where the respondent's household roster included other family members aged 18 years or older ("adult" family members), any one of whom could potentially serve as a proxy for the respondent. The edited indicator was called EDFAM18, where "1" indicated that no potential proxy existed in the respondent's household, "0" indicated otherwise, and "98" indicated unknown.

11.4.2 Editing the Proxy Variables when EDFAM18 = 1

In most cases, a value of EDFAM18 = 1 implied that the respondent was skipped out of the proxy questions because no potential proxy existed in the household. In these cases, all of the proxy variables were given a legitimate skip code (99). Two situations could occur, however, where adult family members were incorrectly identified in the household roster: (1) the respondent had not identified any adult family members in the household but had nonfamily members in the household whose ages were not known; and (2) the unedited household roster indicated that one potential proxy existed in the household but editing changed the age of this single potential proxy to younger than 18. In these cases, the respondent was allowed to answer the proxy questions even though the value of EDFAM18 was 1 (i.e., the final edited household roster indicated that no potential proxy existed in his or her household). Moreover, in these situations, the interviewer indicated that none of these household members who were incorrectly identified as adult family members were proxies. However, the "no" value in the first unedited proxy variable (QP01) was replaced by a logically assigned legitimate skip (89) in the corresponding edited variable (PRXABLE2). For cases where PRXABLE2 was set to 89, all of the edited proxy variables corresponding to the unedited proxy variables, which followed QP01, were given legitimate skip codes (99). These were cases in which the respondent answered the proxy section (questions about a proxy) but the interviewer indicated that they were not proxies, so no proxy should have actually answered the health insurance and income sections.

11.4.3 Editing the Proxy Variables when EDFAM18 = 0

If EDFAM18 was 0, the proxy variables were edited as follows:

1. If the unedited proxy variables had legitimate nonmissing values (i.e., not replaced by a logically assigned legitimate skip), the edited proxy variables (except PRXRELAT and QP03A) were set to those nonmissing values.
2. If any of the unedited proxy variables (except PRXRELAT and QP03A) had a value of 2 ("no"), then all of the variables that followed were edited to legitimate skips.
3. If any of the unedited proxy variables (except QP03A) had a value of "don't know" or "refused," then the corresponding edited variable and all the edited variables that followed were given a "don't know" or "refused" code (94 or 97).
4. If any of the unedited proxy variables did not have a value and a legitimate skip code could not be applied, then the corresponding edited variable and all the variables that followed were given a "no answer" code (98).

11.4.4 Additional Editing for PRXRELAT

In addition to these, more detailed rules were used to assign values to PRXRELAT. The value of QP02, which identified the proxy for the respondent, was chosen directly from the respondent's household roster with one caveat. A list of adult family members (a proxy roster) was shown to the respondent, and the respondent was asked to select the family member who could best answer the health insurance and income sections. In the cases where the proxy roster included a large number, only the first nine adult family members were listed. Regardless of roster size and composition, respondents were given the final choice to select an "other adult relative" as a proxy, even if that person was not part of the household roster. Once the proxy roster was established, the number selected in QP02 was matched to the corresponding person in the proxy roster.

11.4.5 Additional Editing for QP03A

Item QP03A allowed respondents to select another proxy from the list of potential proxies in QP02 if the previously selected proxy was not available to join the interview. This question was added in 2014. Respondents were redirected to question QP03A if they answered "no" to any of the unedited proxy variables; otherwise, QP03A was edited to a skip code. The edited variable of QP03A is PRXRETRY.

This page intentionally left blank

12. Editing and Imputation for the NSDUH Income Variables

12.1 Introduction

This chapter describes procedures for editing and imputing the income variables in the 2016 National Survey on Drug Use and Health (NSDUH) interview. The income section is interviewer administered. At the end of the income section, questions also are included about cell phone coverage in households. These cell phone coverage variables are edited but not imputed.

Unlike other sections of the interview that asked about behaviors or situations in the 12-month period prior to the interview, the reference period for the income section in the 2016 NSDUH was the previous calendar year (i.e., January 1, 2015, through December 31, 2015). Respondents (or other household members serving as proxies¹⁶⁸) were asked whether they (or other family members living in the dwelling unit, if applicable) received income or benefits from specific sources, such as social security during the previous calendar year. These questions are subsequently referred to as "source-of-income" questions in this chapter. In addition, the income section of the interview included questions about the dollar amounts of total personal income and total family income (if applicable) during the previous calendar year.

Several changes to the income section of the interview have occurred over the years. From 1999 to 2007, respondents were asked about personal and family-level income from specific sources as well as binary (i.e., \$20,000 or more or less than \$20,000) and finer categories of total annual income at both the personal and family levels. Beginning in 2008, respondents were no longer asked about personal income from specific sources, nor were they asked about income from child support, interest/investment income, or other income. Beginning in 2015, respondents were no longer asked the question about the receipt of wages or pay that respondents earned while working at a job or business. A comparison of the 1999-2007, 2008-2014, and 2015-2016 sets of income questions is shown in [Table 12.1](#). See Section 3.4 of the 2008 imputation report (Ault et al., 2010) for a more detailed explanation of the changes to the income questions over the years prior to 2015.

Consistent with most of the imputation-revised variables that were discussed in the previous chapters of this report, imputations for the income variables in the 2016 NSDUH were performed using the predictive mean neighborhood (PMN) methodology detailed in Chapter 3. The edits that were applied to the income variables prior to imputation are described in Section 12.2.

¹⁶⁸ Reference is made only to "respondents" in the remainder of this chapter. However, readers are advised that the income information for a respondent may have been provided by another adult household member who was serving as a proxy for the respondent because the proxy was considered to be better able to answer the income questions for the respondent. See Chapter 11 for more information about proxy variables.

Table 12.1 Comparison between 1999-2007, 2008-2014, and 2015-2016 Sets of Income Questions

Income Questions Included in NSDUH	1999-2007 Surveys ¹			2008-2014 Surveys		2015-2016 Surveys	
	Personal Level	Other Family Member Level	Family Level	Personal Level	Family Level	Personal Level	Family Level
Social Security	Yes	Yes	No	No	Yes	No	Yes
Supplemental Security Income	Yes	Yes	No	No	Yes	No	Yes
Welfare Payments	Yes	Yes	No	No	Yes	No	Yes
Other Welfare Services	Yes	Yes	No	No	Yes	No	Yes
Investment Income	Yes	Yes	No	No	No	No	No
Child Support Payments	Yes	Yes	No	No	No	No	No
Wages	Yes	Yes	No	No	Yes	No	No
Other Income	Yes	Yes	No	No	No	No	No
Food Stamps/SNAP Benefits ²	No	No	Yes	No	Yes	No	Yes
Months on Welfare	No	No	Yes	No	Yes	No	Yes
Binary Total Income	Yes	No	Yes	Yes	Yes	Yes	Yes
Finer Category Total Income	Yes	No	Yes	Yes	Yes	Yes	Yes

¹ In the 2006-2007 surveys, 5 percent of the selected sample was asked a reduced set of income questions, which are identical to the set of questions asked in the 2008-2014 surveys. For detailed explanations of these subsamples, consult the 2006 and 2007 NSDUH methodological resource books.

² In the 2015 and 2016 surveys, respondents were asked whether any family members received benefits from the Supplemental Nutrition Assistance Program (SNAP) instead of being asked about the receipt of food stamps. However, the question indicated that SNAP was formerly known as food stamps.

12.2 Editing for Binary and Finer Category Income Variables

12.2.1 Binary Income Variables

The source-of-income variables recorded binary (i.e., yes/no) responses. Information for most of these variables was captured through single questions about a given source of income. One exception to this single-question format was for measurement of the number of months on welfare (questions QI12AN and QI12BN and the edited variable WELMOS). Question QI12AN was asked if respondents did not report any family members' receipt of Supplemental Nutrition Assistance Program (SNAP) benefits (food stamps prior to 2015) in question QI07N; otherwise, question QI12BN was asked. In asking about the number of months on welfare, question QI12BN included the additional phrase, "not including SNAP benefits."

Creation of WELMOS also was dependent on respondents' answers in the questions corresponding to the relevant edited variables for family members' receipt of cash assistance from a state or county welfare program (FAMPMT) and receipt of other (noncash) welfare or public assistance (FAMSVC). If FAMPMT or FAMSVC indicated that respondents received welfare payments, WELMOS was assigned the values from questions QI12AN and QI12BN. Otherwise, if family members in the household (including the respondent) definitely did not

receive welfare payments or noncash benefits (e.g., job training, help with child care), WELMOS was assigned a legitimate skip code.

In addition to the source-of-income variables, binary variables were created from a pair of questions that asked whether the respondent's personal total income or the respondent's total family income was \$20,000 or more. For this pair of questions (QI20N and QI22, corresponding to edited variables PINC1 and FINC1), the second question in the pair applied to the entire family. For answering these questions about total personal or family income, respondents were shown a list of other sources of income that they had not been previously asked about (e.g., child support payments). Respondents were instructed in answering these questions about total income to consider both these other sources of income and the sources of income they had been asked about previously.

If a respondent had no other family members living in his or her dwelling unit, FINC1 was assigned a legitimate skip code (Section 2.4.2). QI22 also was skipped and FINC1 was assigned a legitimate skip code if respondents reported in QI20N that they had personal incomes of \$20,000 or more. A third binary total family income variable, FAMINC1, was created and was equal to either PINC1 or FINC1, depending on whether other family members were present in the household.

12.2.2 Finer Category Income Variables

Respondents also were asked to identify, both for themselves and for their families, finer categories of income within the two general categories of "\$20,000 or more" or "Less than \$20,000." If respondents answered binary total income questions as "Less than \$20,000," they were asked to report a finer category of income from incomes less than \$1,000 (including a net loss of income) up to \$20,000 in increments of \$1,000. Similarly, respondents who answered binary total income questions as "\$20,000 or more" were asked to report a finer category of income from \$20,000 up to \$50,000 in increments of \$5,000. If respondents' or their families' incomes were greater than \$50,000, they could select the following additional categories: "\$50,000 to \$74,999," "\$75,000 to \$99,999," "\$100,000 to \$149,999," or "\$150,000 or more." (In 2004-2014, the maximum income category was "\$100,000 or more." Prior to 2004, the maximum income category was "\$75,000 or more.")

Since 2015 new income variables have been created based on a maximum income category of \$150,000 or more. Specifically, the variable PRINC3 for finer categories of personal income was created from responses to questions QI21A (for personal incomes of less than \$20,000) and QI21B (for personal incomes of \$20,000 or more). The variable FMINC3 for finer categories of family income was created from questions QI23A (for family incomes of less than \$20,000) and QI23B (for family incomes of \$20,000 or more). Thus, income categories in PRINC3 and FMINC3 ranged from 1 (less than \$1,000) to 30 (\$150,000 or more).

Questions QI23A or QI23B were skipped if the respondent had no other family members in the household. Therefore, when the imputed household roster variable IRFAMSKP indicated that no other family members were living in the household, FMINC3 was assigned a legitimate skip code of 99; this included situations in which data existed in questions QI23A or QI23B but IRFAMSKP subsequently indicated that no other family members lived in the household.

A third total family income variable with finer categories of income, FMLYINC3, also was created. If other family members were living in the household, FMLYINC3 was set to be equal to the response from the mutually exclusive pair of total family income questions QI23A and QI23B. If no other family members lived in the household, FMLYINC3 was set to be equal to the responses from questions QI21A or QI21B. Thus, unlike FMINC3, FMLYINC3 was not assigned legitimate skip codes if no other family members lived in the household. Respondents who reported in question QI21B that their total personal income was the maximum income category (i.e., \$150,000 or more in 2016) were not asked question QI23B about their total family income. FMLYINC3 was assigned a code of 30 (\$150,000 or more) in this situation.

12.2.3 Additional Edits to Income Variables for Inconsistencies or Logical Inferences

Relatively little additional editing was done to the income variables, aside from assigning legitimate skip codes (Section 2.4.2). This section describes the additional edits that were implemented because of inconsistencies in the data or to make logical inferences.

As noted in Section 4.2.3, interviewers entered information at the beginning of the interview about the state where the dwelling unit was located. This information was used in the income section to fill in the state-specific names for Temporary Assistance for Needy Families (TANF) programs in question QI08N. If the edited variable STATELOC for the state where the dwelling unit was located had been set to bad data because of interviewer errors, any nonblank values in the variable corresponding to question QI08N about cash assistance from a state or county welfare program (FAMPMT) were set to bad data. This edit was implemented because the incorrect state residence information could result in respondents not being asked about cash assistance from the TANF program in the states where they actually lived. In turn, if FAMPMT had been set to bad data because STATELOC had been set to bad data, and FAMSVC did not indicate welfare assistance, then nonblank values in WELMOS also were set to bad data.

Two types of edits were implemented for overall family income variables based on data about personal income:

- If the binary family income variable FINC1 indicated that the family income was less than \$20,000, and the corresponding personal income variable PINC1 had a missing value, then the respondent's personal income (PINC1) was logically inferred to be less than \$20,000. Similarly, if the binary personal income variable PINC1 was \$20,000 or more, and the family income variable FINC1 was skipped, then the total family income variable FAMINC1 indicated that the family income was \$20,000 or more.
- If the finer category for total personal income (PRINC3) was greater than the corresponding income category in FMINC3, then FMINC3 and FMLYINC3 were set to bad data. For example, if PRINC3 indicated that the respondent had an income of \$25,000 to \$29,999, and the respondent reported a total family income of only \$20,000 to \$24,999, then FMINC3 and FMLYINC3 were set to bad data.

In addition, the recoded variables PINC2, FINC2, and FAMINC2 were created in 2016 to facilitate comparison of the income data in 2016 with data from 2004 to 2014, when the maximum income category was \$100,000 or more. These recodes were created by collapsing

reported incomes of \$150,000 or more with incomes of \$100,000 to \$149,999 to create a maximum category of 29 (\$100,000 or more).

12.3 Editing for Cell Phone Variables

A question on coverage of the household by telephones other than cell phones was added to the NSDUH interview in 2015:

CELL1 Is there at least one telephone at this address that is not a cell phone?

1 YES
2 NO
DK/REF

The wording of a second question (CELL2) was modified slightly for 2015, and an interviewer note was added:

CELL2 Do you or anyone at this address have a cell phone that is currently working?

INTERVIEWER NOTE: We are asking whether you or anyone at this address has a cell phone that currently works and can be used to make or receive calls.

1 YES
2 NO
DK/REF

The order of these questions was reversed in 2016 (i.e., CELL2 appearing first) to avoid respondent confusion based on the previous order.

Minimal processing of data was done to these variables for cell phone coverage. Specifically, the unedited variables were assigned final, mnemonic variable names. Based on the question order in 2016, question CELL2 was renamed CELLWRKNG and question CELL1 was renamed CELLNOTCL.

Changing the order of these questions in 2016 was determined not to have affected the comparability of data between 2015 and 2016. Therefore, these variables retained the names CELLWRKNG and CELLNOTCL in 2016.

12.4 Imputation for the Income Variables

The imputation of income was separated into two imputation sets. The first set involved the imputation of all the binary income variables (i.e., "yes/no" questions about the following sources of income: social security, supplemental security income, welfare cash assistance, welfare noncash assistance, wages, and SNAP benefits), the number of months on welfare (the only variable that was not binary in this imputation set), and a yes/no question regarding whether the respondent's income or the respondent's household family income was \$20,000 or more (including income from the sources referenced in the previous questions). This first set was processed using the single response propensity (RP)/multiple prediction (PRD) PMN type

described in Section 3.4.3. The second imputation set for finer categories of income consisted of imputing more specific income categories for the respondent and the respondent's family in the household. This set was processed using the single RP/single PRD PMN type as described in Section 3.4.1.

The income variables for total personal income, total family income, and number of months on welfare have lower item response rates compared with most of the other variables that undergo imputation. See [Table A.28](#) in Appendix A for details on the rates of missingness for these variables.

Imputations for all income variables were conducted separately within four age groups: 12 to 17, 18 to 25, 26 to 64, and 65 or older. The segregation into age groups was done to exploit the correlation between the income variables and age and to allow parallel processing of the data (thus reducing the time it takes to implement the procedures).

All income RP and PRD models used a state-rank variable as a covariate. States were ranked by the weighted proportion of respondents whose personal incomes during the prior calendar year were greater than or equal to \$20,000. See Section 3.8.2 for a general discussion of state-rank variables.

12.4.1 Imputation for Binary Income Variables (Imputation Set 1)

The PMN imputation type used for the binary income variables was single RP/multiple PRD. The RP model is described in Section 12.4.1.1. The PRD model for the first binary income variable imputed each year, family social security (FAMSOC), is described in Section 12.4.1.2, and the provisional hot-deck step for FAMSOC is described in Section 12.4.1.3. The remaining PRD and provisional hot-deck steps described in Section 12.4.1.4 list deviations from the analogous steps for family social security. The final hot-deck step applied to all binary income variables is described in Section 12.4.1.5. Finally, a recode for the GOVTPROG variable, made from four imputation-revised binary income variables, is described in Section 12.4.1.6.

For the binary income models that predict whether a respondent had a given source of income, other sources of income were useful covariates. Therefore, provisionally imputed values were used as covariates in subsequent models within the set. The order in which missing values for the binary income variables were provisionally imputed is listed in [Table 12.2](#).

Table 12.2 Imputation Order for Binary Income Variables (Imputation Set 1)

Income Type	Edited Variable Name(s)	Imputation-Revised Variable Name(s)
Family Social Security	FAMSOC	IRFAMSOC
Family Supplemental Security Income	FAMSSI	IRFAMSSI
Family Welfare Payments	FAMPMT	IRFAMPMT
Family Other Welfare Services	FAMSVC	IRFAMSVC

Table 12.2 Imputation Order for Binary Income Variables (Imputation Set 1) (continued)

Income Type	Edited Variable Name(s)	Imputation-Revised Variable Name(s)
Family SNAP Benefits ¹	FSTAMP	IRFSTAMP
Family Months on Welfare	WELMOS	IRWELMOS
Total Personal/Family Income	PINC1, FINC1, FAMINC1	IRPINC1, IRFINC1, IRFAMIN1

¹ Beginning with the 2015 survey, respondents were asked whether any family members received benefits from the Supplemental Nutrition Assistance Program (SNAP) instead of being asked about the receipt of food stamps. However, the question indicated that SNAP was formerly known as food stamps. Consequently, this change to the question did not affect the comparability of the data with earlier years, and the variables in 2015 and 2016 continued to use the names FSTAMP and IRFSTAMP.

12.4.1.1 Response Propensity Step

The response propensity models for the binary income variables utilized the analysis weight, ANALWT. All respondents were in the domain for the binary income variables (i.e., eligible to have a valid value for these variables). For the single RP model, a domain member was considered an item respondent if and only if all of the variables listed in [Table 12.2](#) were nonmissing. See [Table D.68](#) in Appendix D for details of the covariates used in the RP models for these variables.

12.4.1.2 First Prediction Step (FAMSOC)

FAMSOC is a binary variable; after imputation, its only values are "yes" and "no" since all other missing data codes are replaced with imputed values. Using the adjusted weights that are outputs of the RP step, FAMSOC was modeled using logistic regression and, in particular, the RLOGIST procedure in SAS-callable SUDAAN. The single predicted mean used in later hot-deck steps was the predicted probability that the respondent's family-in-household received income from social security in the preceding calendar year.

12.4.1.3 First Provisional Hot-Deck Step (FAMSOC)

The provisional hot-deck applied to the FAMSOC variable is a simplified version of the final hot-deck step for the binary income variables, because its only purpose is to fill in missing values so that FAMSOC can be used as a covariate in the PRD models for binary income variables imputed later in the sequence. Section 3.4.2 describes the concept of a provisional hot-deck in more detail. The final hot-deck step is described in Section 12.4.1.5 and in tabular form in [Tables E.141](#), [E.142](#), and [E.143](#) in Appendix E. This provisional hot-deck step included the following:

- The predictive mean vector included only one element: the predicted mean from the preceding step.
- No logical constraints were used.
- The following likeness constraints were used:
 - IRFAMSKP of donor = IRFAMSKP of recipient. This is a likeness constraint in this hot-deck step, not a logical constraint. The creation of IRFAMSKP is described in Section 11.3.2 (LogC1 in [Table E.141](#)).

- Donor's predicted mean must be within 5 percent of recipient's predicted mean (delta constraint; LikC3 in [Table E.142](#)).
- Donor must match recipient with respect to whether there are adults aged 65 or older in the household. The variable used in this constraint was IRHH65, whose creation is described in Section 11.3.2 (LikC5 in [Table E.142](#)).
- In the first attempt to find a donor, all three likeness constraints were applied. In the second attempt, the delta constraint was removed.

The provisionally imputed version of FAMSOC was called INTFAMSOC. INTFAMSOC was used as a covariate in the rest of the binary income PRD models.

12.4.1.4 Analogous Prediction and Provisional Hot-Deck Steps for Remaining Binary Income Variables (FAMSSI, FAMPMT, FAMSV, FSTAMP, WELMOS, and FAMINC1)

PRD models were fit for FAMSSI, FAMPMT, FAMSV, FSTAMP, and FAMINC1¹⁶⁹ in the same manner as for FAMSOC, as described previously. Only the PRD model for WELMOS was different. The domain of the WELMOS model included only respondents who reported that their family-in-household received welfare payments and/or other welfare services during the preceding calendar year, as defined by FAMPMT and FAMSV. Least squares regression (not logistic regression) was used, where the dependent variable was a standard logit,¹⁷⁰ such that $Y = \text{logit}(p)$ and p = number of months on welfare divided by 12. The REGRESS procedure in SAS-callable SUDAAN was used to fit the model. The predicted mean from the WELMOS model was the predicted probability of receiving welfare in a given month in the previous calendar year, given that the respondent received welfare payments and/or welfare services in the previous calendar year.

The provisional imputation steps for the other variables were implemented in the same manner as for FAMSOC, as described previously, except that the likeness constraints sometimes differed slightly. The following deviations are noted:

- Only the provisional hot-deck step for FAMSOC included the constraint involving the IRHH65 variable.
- For FAMPMT and FAMSV, LikC6 in [Table E.142](#) was used: donor must match recipient with respect to whether there are children younger than 18 in the household. The variable used in this constraint was IRKID17, whose creation is described in Section 11.3.2. This likeness constraint was used in both the first and second attempts to find a donor.
- FAMINC1 did not undergo a provisional imputation step since it was the last variable in the set.

¹⁶⁹ The response variable for the model for total income was total family income (not total personal income), and the provisionally imputed version of WELMOS was not included as a covariate in the model.

¹⁷⁰ The Cox empirical logit was used when an individual was on welfare for all 12 months.

12.4.1.5 Final Hot-Deck Step

Details on the missingness patterns, constraints, and predictive mean vectors for the binary income variables' final hot-deck step are available in Appendix E. This section explains the more general ideas behind the hot-deck step applied to these variables.

Because nine imputation-revised variables are created in this step, and almost any subset of them can be missing, there are 319 missingness patterns for the binary income variables. The first 127 cover patterns where WELMOS is nonmissing; these are summarized as pattern 5 in [Table E.143](#). The remaining 192 cover patterns where WELMOS is missing; these are summarized as patterns 1 through 4 in [Table E.143](#).

- When WELMOS is nonmissing, any subset of FAMSOC, FAMSSI, FAMPMT, FAMSVC, FSTAMP, PINC1, and FINC1 may be missing. However, if *none* of them are missing, then nothing is missing and no imputation is required. This leads to $2^7 - 1 = 127$ missingness patterns.¹⁷¹
- When WELMOS is missing, the elements of the predictive mean vector related to FAMPMT, FAMSVC, and WELMOS depend on the values of FAMPMT and FAMSVC. There are six relevant combinations of FAMPMT and FAMSVC ([Table 12.3](#)). For each of the six, any subset of FAMSOC, FAMSSI, FSTAMP, PINC1, and FINC1 may be missing. This leads to $6 \times 2^5 = 192$ missingness patterns.
 - For the combinations labeled 1 in [Table 12.3](#), the recipient is known to have received welfare payments and/or welfare services and neither is missing. Here, no predictive mean vector elements are required for FAMPMT and FAMSVC,¹⁷² and there is no need to manipulate the element for WELMOS, which is already conditional on receipt of welfare as defined by FAMPMT and FAMSVC.
 - For combinations 2 and 3, the recipient is also known to have received welfare payments or welfare services, but one is missing. Here, a predictive mean vector element (i.e., a predicted mean) is required for the missing one of the two, but again there is no need to manipulate the corresponding element for WELMOS.
 - For combination 4, the recipient is known not to have received welfare payments, but welfare services is missing. Here, a predictive mean vector element is required for FAMSVC. The WELMOS predicted mean must be made conditional on the receipt of welfare services in the previous calendar year. Combination 5 is similar, but FAMPMT is the missing one.
 - For combination 6, both FAMPMT and FAMSVC are missing. Here, predictive mean vector elements are required for both. The WELMOS predicted mean must be made conditional on the receipt of welfare payments and/or welfare services. This probability is unknown, but can be crudely approximated by assuming the

¹⁷¹ Technically, if FAMPMT and FAMSVC were both missing, then WELMOS was necessarily missing. Therefore, $2^6 = 64$ out of these 127 missingness patterns are logically impossible and need not be set up.

¹⁷² In general, in either the single RP/multiple PRD or multiple RP/multiple PRD types of PMN, only the elements of the predictive mean vector corresponding to missing responses were used (see Chapter 3).

two are independent. Under that assumption, the probability of receiving either or both is one minus the probability of receiving neither, as expressed by $(1 - \text{PMT}) \times (1 - \text{SVC})$.¹⁷³

- If FAMPMT = "No" and FAMSVC = "No," then WELMOS is assigned a skip code and is therefore nonmissing. This combination is irrelevant for the purposes of this discussion, which only covers cases where WELMOS is missing.
- Combinations 1 through 3 in [Table 12.3](#) correspond to pattern 1 in [Table E.143](#), combination 4 corresponds to pattern 2, combination 5 corresponds to pattern 3, and combination 6 corresponds to pattern 4.

Table 12.3 Cross-Classification of FAMPMT and FAMSVC; Relevant Combinations for Binary Income Predictive Mean Vector when WELMOS Is Missing

		FAMSVC		
		Yes	No	Missing
FAMPMT	Yes	1	1	2
	No	1	N/A	4
	Missing	3	5	6

N/A = not applicable.

Both the likeness and logical constraints are written in such a way that they can be summarized using the five broad missingness patterns listed in [Table E.143](#). For example, likeness constraint 4 states, "If recipient is missing months on welfare, then donor must match recipient with respect to personal welfare payments (if nonmissing) and welfare services (if nonmissing)." This constraint does not apply to the first 127 missingness patterns, nor does it apply to the 64 missingness patterns in combination 6 where both FAMPMT and FAMSVC are missing. However, instead of carefully listing the exact missingness patterns to which each constraint applied, the constraints were written using conditions that allowed them to be applied only where necessary.

The few logical constraints are due to relationships between (1) FAMPMT, FAMSVC, and WELMOS and (2) PINC1, FINC1, and FAMINC1. A key likeness constraint exploited the high degree of association among welfare payments, welfare services, SNAP benefits, binary total income (at the personal and family level), and months on welfare. If a recipient required imputation for one or more of these six variables, but had information on at least one of these variables, the donors were restricted so that donors and recipients had the same values for these nonmissing variables.

As stated in the introduction to Section 12.4, the variables for total personal income, total family income, and number of months on welfare have lower item response rates compared with most of the other variables that undergo imputation. The constraints are fairly complex and extensive, especially when several of the other dichotomous income variables are missing. Also,

¹⁷³ A reasonable alternative method that requires no assumption of independence would be to model FAMPMT and FAMSVC together as a categorical variable with four levels: both, only FAMPMT, only FAMSVC, and neither. Then, the probability of receiving welfare payments and/or other welfare services is simply the sum of the first three predicted means. The assumption of independence is certainly questionable.

when several variables are missing, it is difficult to find a donor on the first try because the donor must be close to the recipient for all predicted means corresponding to the missing variables.

12.4.1.6 Recodes for Additional Analyses

A dichotomous recoded income variable GOVTPROG indicated whether the respondent participated in any government assistance programs. It was created from four imputation-revised variables: family Supplemental Security Income (IRFAMSSI), family SNAP benefits (IRFSTAMP), family welfare payments (IRFAMPMT), and family welfare services (IRFAMSV). Although a variety of recoded variables were created, only GOVTPROG is described here because it was used as a covariate in subsequent health insurance imputation models. See Chapter 13 for details on the imputation of missing values in the health insurance variables.

12.4.2 Imputation for Finer Category Income Variables (Imputation Set 2)

Three income variables resulted from editing the questions in the finer category income phase: personal total income (PRINC3), total family income if there are other family members (FMINC3), and total family income (FMLYINC3). These edited variables are described in Section 12.2.2. All three imputation-revised variables derived from their edited counterparts were created using the single RP/single PRD PMN type described in Chapter 3. The single PRD model was fit for total family income (FMLYINC3), but the item nonrespondent also received values from the donor for PRINC3 and FMINC3 if those were missing. There were no deviations from this general approach.

12.4.2.1 Response Propensity Step

The response propensity models for the finer category income variables utilized the analysis weight, ANALWT. All respondents were in the domain for this imputation set. Item respondents were those with nonmissing values for PINC2 and FINC2. See [Table D.73](#) in Appendix D for details of the covariates used in the RP models for these variables.

12.4.2.2 Prediction Step

Each of the three finer category income variables was ordinal with 30 levels. Each response category covers an interval of income, with levels as follows:

1	LESS THAN \$1,000
2	\$1,000 - \$1,999
3	\$2,000 - \$2,999
4	\$3,000 - \$3,999
5	\$4,000 - \$4,999
6	\$5,000 - \$5,999
7	\$6,000 - \$6,999
8	\$7,000 - \$7,999
9	\$8,000 - \$8,999
10	\$9,000 - \$9,999
11	\$10,000 - \$10,999
12	\$11,000 - \$11,999

13	\$12,000 - \$12,999
14	\$13,000 - \$13,999
15	\$14,000 - \$14,999
16	\$15,000 - \$15,999
17	\$16,000 - \$16,999
18	\$17,000 - \$17,999
19	\$18,000 - \$18,999
20	\$19,000 - \$19,999
21	\$20,000 - \$24,999
22	\$25,000 - \$29,999
23	\$30,000 - \$34,999
24	\$35,000 - \$39,999
25	\$40,000 - \$44,999
26	\$45,000 - \$49,999
27	\$50,000 - \$74,999
28	\$75,000 - \$99,999
29	\$100,000 - \$149,999
30	\$150,000 OR MORE

The FMLYINC3 variable was modeled using the LIFEREG procedure in SAS/STAT[®] software.¹⁷⁴ This procedure was used for regression modeling of continuous nonnegative random variables, such as survival times and income, by fitting models that are sometimes referred to as "failure time models." This particular type of model, which was assumed for the response variable representing income, can be written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where $\mathbf{y}$ is a vector of observed responses, $\mathbf{X}$ is the matrix of covariates, $\boldsymbol{\beta}$ is the parameter vector, and $\boldsymbol{\varepsilon}$ is a vector of error terms. In particular, the error terms are assumed to come from a known multivariate distribution, such as the logarithm of a three-parameter generalized gamma model, or a more common two-parameter distribution, such as gamma, Weibull, lognormal, or log-logistic. Although the underlying random variable $\mathbf{y}$ is assumed to be continuous, the LIFEREG procedure allows the variable to be reported in interval categories, consistent with the 30 NSDUH income intervals defined previously for these finer category income variables. The contribution of an individual with covariates in the matrix $\mathbf{X}$ to the overall likelihood is simply the probability mass assigned by the model to the interval (l, u) containing the actual continuous income for that individual. For this interval, l represents the lower bound and u represents the upper bound. This contribution has the form $F(u|\mathbf{X}, \boldsymbol{\beta}, \sigma^2) - F(l|\mathbf{X}, \boldsymbol{\beta}, \sigma^2)$, where F is a cumulative distribution function and σ^2 represents the variance of the individual responses. The LIFEREG procedure uses standard likelihood methods of inference and incorporates the survey weights.

LIFEREG allows several choices for the functional form of the parametric model that correspond to the error distribution, including the two-parameter log-logistic, lognormal, gamma,

¹⁷⁴ Details about the LIFEREG procedure are discussed in the *SAS/STAT[®] 13.2 User's Guide: High-Performance Procedures* (SAS Institute Inc., 2014).

and Weibull and the three-parameter generalized gamma. Compared with the other models, the gamma distribution provided a better overall fit, as indicated by likelihood techniques. Because the three-parameter generalized gamma did not significantly improve on its two-parameter special cases, when using the likelihood ratio tests as criteria for comparison, it was decided to use a two-parameter model.

The predicted mean used in the subsequent hot-deck step was the term $X\beta$, which was the predictive mean value. This value was a monotonic function of the conditional mean of the modeled income distribution at a given individual set of values of the regression covariates. Specifically, $X\beta$ was a translation of the estimated mean of log income.

12.4.2.3 Hot-Deck Step

The hot-deck step for the finer category income variables is an example of univariate matching, multivariate assignment (see [Table 3.4](#) in Chapter 3). The only predicted mean used in the hot-deck step was related to FMLYINC3, but the recipient received values for PRINC3 and FMINC3 if either or both were missing. The imputation-revised version of FMLYINC3 (IRFMLYIN3) was created from the imputation-revised versions of PRINC3 and FMINC3, called IRPRINC3 and IRFMINC3, respectively. The constraints ensure consistency with existing information such as IRPINC1 and IRFINC1 from the preceding imputation set, and PRINC3 and FMINC3 if nonmissing.

With close to 10 percent of missing cases being imputed annually, the finer category income variables have among the lowest response rates of any NSDUH variables that undergo imputation. As a result of the absence of many constraints, the single predicted mean, and the large domain, most respondents are handled in the first attempt to find a donor.

12.4.2.4 Recodes for Additional Analyses

The recoded variable INCOME classified the families of respondents into four income levels: less than \$20,000; \$20,000 to \$49,999; \$50,000 to \$74,999; and \$75,000 or more. Another recoded variable (INCOME5) was created to take advantage of an extra level of income. This variable had five levels: the first three levels were equivalent to INCOME, but the last level of INCOME was separated into two levels: \$75,000 to \$99,999 and \$100,000 or more. Both INCOME and INCOME5 were recodes of the variable IRFMLYIN3. A variety of recoded variables were created but are not discussed in this report. However, as with GOVTPROG, the variable INCOME is discussed here because it was used as a covariate in subsequent health insurance models (see Chapter 13 for details on the imputation of missing values in the health insurance variables). INCOME5, which is currently used for special requests, also is discussed because it is similar to the INCOME variable.

The recoded variables IRPINC2, IRFINC2, and IRFAMIN2 were created to be comparable with the variables of the same names created in 2004-2014, in which the finer categories of income had 29 levels with the top level being "\$100,000 OR MORE." These recodes were made by collapsing the top two levels of variables IRPRINC3, IRFMINC3, and IRFMLYIN3.

This page intentionally left blank

13. Editing and Imputation for the NSDUH Health Insurance Variables

13.1 Introduction

This chapter discusses procedures for logically editing and imputing data from the interviewer-administered sections pertaining to health insurance coverage. Respondents (or other household members serving as proxies because they were better able to provide information about the respondent's health insurance coverage) were asked whether they were currently covered by different types of health insurance.¹⁷⁵ If private health insurance coverage was reported, respondents were asked whether that included coverage for substance abuse treatment or mental health services. Data also were collected on periods when respondents never had health insurance coverage, most recent coverage (if currently without health insurance), and reasons for losing health insurance coverage or for never having had coverage (if applicable).

For the 2002-2012 National Surveys on Drug Use and Health (NSDUHs), missing values for the health insurance variables were replaced with valid values using two different predictive mean neighborhood (PMN) imputation methods: the "old" method and the "constituent variables" method. The "old" method consists of edited and imputed values for three health insurance variables in a way that is consistent with iterations of the NSDUH questionnaire prior to 2002. The "constituent variables" method edits and imputes for specific health insurance variables (e.g., Medicaid or Medicare coverage) using all relevant questions in the questionnaire from 2002 to the present and was first implemented for the health insurance section in 2002.

Starting with the 2013 NSDUH, however, the creation of the edited and imputed "old" method variables was discontinued, and the "constituent variables" method was thereafter referred to as simply the health insurance method. There were two main reasons for this decision. First, analysts have long been cautioned against performing trend analyses that combine pre-2002 survey years with 2002 and beyond. Second, including both sets of health insurance variables had the potential to confuse analysts who might use the incorrect health insurance variable set to produce estimates.

As with other variable groups (e.g., demographics, drugs), the health insurance variables were imputed in sets as follows:

- Imputation Set 1: IRMCDCHP, IRMEDICR, IRCHMPUS, and IRPRVHLT (types of health insurance)
- Imputation Set 2: IROTHHLT (any other health insurance)

¹⁷⁵ Reference is made only to "respondents" in the remainder of this chapter. However, readers are advised that the health insurance information for a respondent may have been provided by another adult household member who was serving as a proxy for the respondent because the proxy was considered to be better able to answer the health insurance questions for the respondent. See Chapter 11 for more information about proxy variables.

The five base variables (CAIDCHIP, MEDICARE, CHAMPUS, PRIVATE, ANYOTHER) that undergo imputation are simple recodes created from six "source" variables, each of which maps to a single question in NSDUH. In the 2016 survey, the source variables had the same values as the raw variables created from the responses to the questions, except that missing values were replaced by standard NSDUH missing value codes (Sections 2.2.1 and 2.4.3). With the exception of IRMCDCHP (which was a combination of MEDICAID and CHIPCOV), the source variables remained separate to form the base variables for imputation sets 1 and 2. Section 13.2.5 discusses the creation of base variables used in imputation sets 1 and 2.

In addition to the five imputation-revised variables, the health insurance imputation process also creates the overall health insurance variable, IRINSUR4. This variable combines information from the detailed imputation-revised variables into one summary measure of coverage.

13.2 Editing the Health Insurance Variables

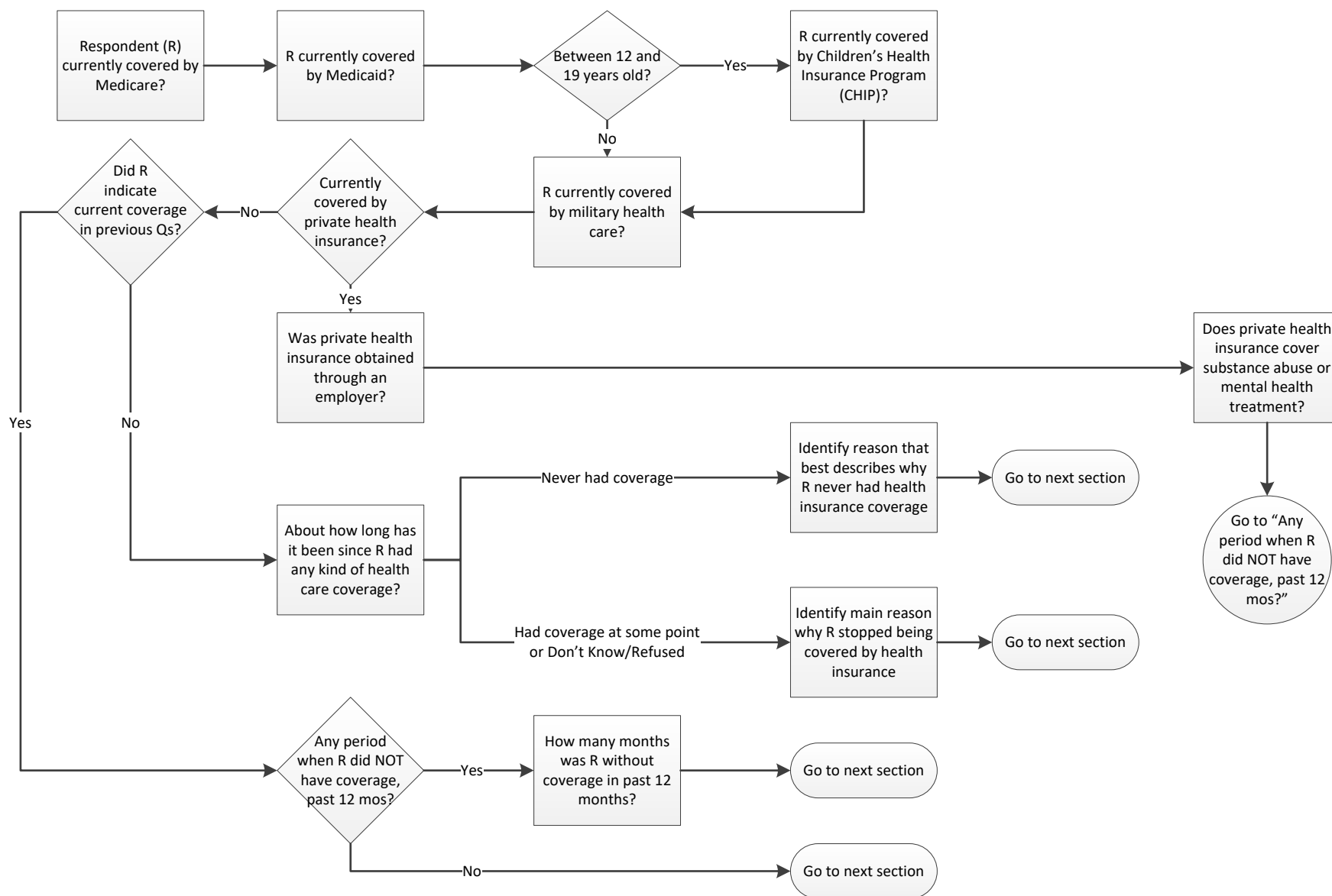
The content and mode of administration of the health insurance section have undergone relatively little change since the start of the new NSDUH baseline in 2002. In particular, the health insurance section continued to be interviewer administered in 2016, in contrast to several other sections of the questionnaire having moved from being interviewer administered to being self-administered beginning in 2015 (see Chapter 5).¹⁷⁶ Documentation of historical changes to the health insurance questions prior to the new baseline in 2002 can be found in the report on editing of interviewer-administered data that was prepared for the 2011 NSDUH methodological resource book (Kroutil, Chien, Handley, & Bradshaw, 2013).

An important aspect of editing the health insurance variables consisted of assigning legitimate skip codes based on the skip logic in this section (Section 2.4.2). [Figure 13.1](#) shows the logic for asking questions in the health insurance section. For example, if respondents answered "no" (where applicable) to questions QHI01 through QHI06 and then reported in QHI11 that they were not currently covered by any kind of health insurance (QHI11 = 2), legitimate skip codes were assigned to HLCNOTYR (anytime in the past 12 months that respondents were without health insurance, corresponding to question QHI13) and HLCNOTMO (number of months that respondents were without health insurance in the past 12 months, corresponding to question QHI14). Similarly, if respondents reported some type of current health insurance coverage in QHI01 through QHI06, edited variables corresponding to questions QHI15 through QHI18 were assigned legitimate skip codes (i.e., HLCLAST through HLNVSOR).

As was the case in prior years, question QHI18 (reasons that the respondent never had health insurance) was an "enter all that apply" question. Therefore, the edited variables corresponding to question QHI18 (HLNVCOST through HLNVSOR) were assigned a code of 1 (Response entered) if the corresponding response category was chosen from QHI18. The variables were assigned a code of 6 (Response not entered) if the corresponding response category was not chosen but at least one response had been entered in QHI18 (Section 2.4.4).

¹⁷⁶ Self-administered versions of the health insurance questions were field tested in 2012 and 2013 in preparation for the partial redesign of the 2015 NSDUH questionnaire. Based on the field test results, however, SAMHSA determined that the health insurance questions would continue to be interviewer administered for the 2015 questionnaire (Center for Behavioral Health Statistics and Quality, 2014a; 2014b).

Figure 13.1 Routing Logic for Respondents in the Health Insurance Section



*Respondents immediately go to the next section (i.e., without further health insurance questions being asked) if current health insurance coverage status is unknown.

13.2.1 Editing Because of Incorrect State Information

Question QHI02 asked whether respondents were currently covered by Medicaid, and question QHI02A asked whether respondents aged 12 to 19 were covered by the Children's Health Insurance Program (CHIP). The CAI program filled in names for state-specific Medicaid program or CHIP names to aid respondent identification. If the interviewer's report of the state where the dwelling unit was located (edited variable STATELOC) had been set to bad data because of inconsistencies in the state information for the respondent (Section 4.2.3), then the edited variable CAIDCHIP for coverage by Medicaid or CHIP (see Section 13.2.2) was usually assigned a bad data code as well. The rationale for this edit was that the CAI logic would supply an incorrect name for the state's Medicaid program or CHIP if the state information that the interviewer reported was incorrect. Consequently, the respondent would be answering QHI02 or QHI02A based on a version of the question that did not correctly correspond to where the respondent would be eligible for publicly funded health insurance coverage. For example, if a respondent was aged 12 to 19 and was living in California (FIPE4 = 5), the respondent should have been asked in QHI02A whether he or she was covered by "Healthy Families or Medi-Cal." However, if a value of 6 had been entered in FIPE4 (i.e., for Colorado), the respondent would be asked whether he or she was covered by "Child Health Plan Plus, or CHP+."

An exception to this assignment of bad data codes concerned the special situation in which respondents were routed to questions QHI15 (time since the respondent last had health insurance) and QHI17 (main reason for losing health insurance coverage). If responses to QHI15 or QHI17 indicated that the respondent did not currently have (or never had) health insurance coverage, CAIDCHIP retained a code of 2 (i.e., "no"), even if STATELOC had been set to bad data, for consistency with information from QHI15 or QHI17 that the respondent was not currently covered by any type of health insurance.

13.2.2 Determination of Any Health Insurance Coverage

Determining whether respondents were currently covered by any type of health insurance was critical for determining the subsequent questions that respondents would be asked about periods without health insurance coverage in the past 12 months, most recent health insurance coverage (if not currently covered), reasons for losing health insurance, or reasons for never having health insurance. The following content issues have been relevant since the new NSDUH baseline in 2002 for determining whether respondents had health insurance coverage:

- As noted previously, respondents who were aged 12 to 19 are asked in question QHI02A whether they were covered by CHIP. Government experts in the health insurance field advised SAMHSA that it would be virtually impossible to produce separate estimates of coverage by the Medicaid program (question QHI02) and coverage by CHIP. For this reason, the variable CAIDCHIP has been created from responses to questions QHI02 and QHI02A.
- The state-specific Medicaid or CHIP names that are used to customize questions QHI02 and QHI02A may be updated for the CAI program across years to reflect any changes to these names at the state level.

- If STATELOC had a valid value (Sections 4.2.3 and 13.2.1), CAIDCHIP was assigned a code of 1 (i.e., "yes") if an affirmative response occurred in either QHI02 or QHI02A (if applicable). CAIDCHIP was coded as 2 ("no") if QHI02 was answered as "no" and (a) QHI02A also was answered as "no" (for respondents who were aged 12 to 19) or (b) QHI02A had been legitimately skipped (for respondents aged 20 or older). Otherwise, CAIDCHIP was coded as 97 ("refused") if a code of 97 occurred in either QHI02 or QHI02A, or 94 ("don't know") if a code of 94 (and no code of 97) occurred in these items. Remaining cases that did not meet any of these criteria were coded as 98 (i.e., blank).
- Respondents who answered "no" to all questions about Medicare, Medicaid, CHIP (if applicable), military health coverage, and private health insurance were asked a follow-up question (QHI11) to determine if they were covered by *any* type of health insurance. Responses to this question were used to determine subsequent routing in the health insurance section depending on whether respondents currently had or did not have health insurance. The variable HLTINNOS was created from QHI11.
- A recoded "any health insurance" variable, ANYHLTIN, was created from responses to MEDICARE (from QHI01), CAIDCHIP (from QHI02 and QHI02A), CHAMPUS (from QHI03), PRVHLTIN (from QHI06), and HLTINNOS (from QHI11).
 - If any affirmative response was reported in any of the above variables, ANYHLTIN was coded as 1 ("yes").
 - Otherwise, if HLTINNOS (from QHI11) had been answered as "no" (and by definition, preceding questions had been answered as "no"), ANYHLTIN was coded as 2 ("no").
 - If ANYHLTIN was not already coded as 1 or 2, it was coded as 97 ("refused") or 94 ("don't know"), as follows: (a) if a code of 97 (i.e., "refused") occurred in any of the above health insurance items, ANYHLTIN was coded as 97; or (b) ANYHLTIN was coded as 94 (i.e., "don't know") if a code of 94 (but no code of 97) occurred in the above items.
 - For remaining cases (e.g., if variables had been set to bad data, or a breakoff had occurred), ANYHLTIN retained a code of 98 (OTHER MISSING).

In addition, because the names for Medicare and Medicaid sound similar, follow-up questions have been included in the health insurance section of the interview since 2003 for respondents who were (1) younger than 65 and reported being covered by Medicare (which is for people aged 65 or older or certain disabled individuals) or (2) aged 65 or older and reported being covered by Medicaid (which is for low-income individuals or disabled individuals). In these situations, respondents were asked to confirm their previous answer. If respondents indicated on follow-up that their previous answer to Medicare or Medicaid coverage was incorrect, it was determined that they were not covered by that particular type of health insurance. However, if respondents did not know or refused to confirm whether their previous answer to the Medicare or Medicaid coverage question was correct, the editing procedures continued to classify them as having that type of coverage. Stated another way, the only response on follow-up that overruled a previous report of coverage by Medicare or Medicaid was for respondents to indicate that they did not have that type of coverage.

Despite these changes in 2003 for questions about Medicare or Medicaid coverage, the names for the edited variables did not change. The edited variable for Medicare coverage retained the name MEDICARE. The edited variable for coverage by Medicaid or CHIP retained the name CAIDCHIP (see previous discussion in this section).

13.2.3 Editing of Health Insurance Follow-Up Variables Based on Current Coverage

Information from ANYHLTIN and a preliminary edited variable SKHLCCOV¹⁷⁷ were used to determine whether variables pertaining to periods without health insurance coverage in the past 12 months, most recent health insurance coverage (if not currently covered), reasons for losing health insurance, and reasons for never having health insurance should be assigned "legitimate skip" codes (Section 2.4.2).

- If respondents reported some health insurance coverage in QHI01 to QHI11, legitimate skip codes were assigned to the edited variables corresponding to questions QHI15 to QHI18. Questions QHI15 to QHI18 applied to respondents who currently were without coverage and pertained to situations in which respondents formerly or never had coverage.
- If respondents currently were without any health insurance coverage (i.e., applicable questions in QHI01 to QHI11 all were answered as "no"), legitimate skip codes were assigned to the edited variables corresponding to questions QHI13 and QHI14. Questions QHI13 and QHI14 applied to respondents who currently had coverage and pertained to periods when these respondents were without coverage in the past 12 months.
- If it was unknown from questions QHI01 to QHI11 whether or not respondents currently had health insurance coverage, the CAI program skipped respondents out of all remaining health insurance questions. The edited variables corresponding to questions QHI13 to QHI18 had missing values.

13.2.4 Miscellaneous Issues for Editing the Health Insurance Variables

Table 13.1 discusses edit specifications for the health insurance variables. For example, the data could indicate that respondents were covered currently by Medicare, Medicaid, CHIP (for respondents who were aged 12 to 19), some type of military health coverage (e.g., CHAMPUS or the VA), or private health insurance. If respondents were reported to have been *currently* covered by all of the types of insurance they were asked about, a flag was set and included on the data file. The original data were retained, but this flag was designed to alert analysts to the presence of this unlikely data pattern.

¹⁷⁷ SKHLCCOV was a precursor to the final variable ANYHLTIN but was used in some places in the editing program to determine whether respondents were legitimately skipped out of certain follow-up questions about health insurance coverage. Both variables indicate whether respondents currently had health insurance coverage. SKHLCCOV was coded as 3 if respondents reported in question QHI02A that they were covered by CHIP but they did not report coverage by Medicaid, Medicare, military health insurance (e.g., CHAMPUS), or private health insurance. These values of 3 in SKHLCCOV were recoded to 1 in ANYHLTIN.

In addition, the only types of current health insurance coverage that were asked about in 1999 were Medicare, Medicaid, some type of military health coverage, or private health insurance. Therefore, a second flag was set for comparability with a similar flag set in the 1999 data. This second flag indicated when respondents reported that they were covered currently by all four of these types of health insurance that were asked about in 1999, even if they did not report being covered by CHIP (if aged 12 to 19) or they were aged 20 or older and were skipped out of question QHI02A.

Table 13.1 Edits Pertaining to the Health Insurance Section

Issue	Edits Implemented
The respondent (R) reported being currently covered by Medicare, Medicaid, Children's Health Insurance Program (CHIP) (if aged 12 to 19), military coverage, and private health insurance.	A flag (HLCALLFG) was provided to indicate that this pattern occurred, but no further editing was done to the data.
The R reported being currently covered by Medicare, Medicaid, military coverage, and private health insurance, the only types of current coverage that were asked about in 1999.	A flag (HLCALL99) was provided to indicate that this pattern occurred, but no further editing was done to the responses. This HLCALL99 variable was comparable to the HLCALLFG variable in 1999.
The R's only indication of current health insurance coverage came from reports of coverage by Medicaid or CHIP, but the state location variable STATELOC had been set to bad data.	Nonblank values in the variables pertaining to any period in the past 12 months when the R was without health insurance (HLCNOTYR, corresponding to question QHI13) and the number of months that the R was without health insurance in the past 12 months (HLCNOTMO, corresponding to question QHI14) were replaced with bad data codes. This edit was not done if the R indicated current coverage by Medicare, the military, or private health insurance.
The R had some indication of current coverage from at least one of the five sources of insurance listed above. However, the R also was reported to have had a period in the past 12 months when he or she was without health insurance. Further, it was reported that the R had been without health insurance for 12 of those months.	No editing was done when this pattern occurred. The rationale for not doing any editing was that the R may just recently have gotten insurance or have become qualified for insurance.
The R had no indication of current coverage from any of the five sources of insurance listed above. If the R (or proxy) answered "don't know" or "refused" when asked when the R last had coverage, the R was routed to questions about what coverage the R last had, and why the R lost health insurance coverage. That is, the skip logic assumed that the R had some prior history of coverage, but that may not necessarily have been the case.	If the R was reported to have had previously some form of health insurance or medical coverage, or if some reason was given why the R lost insurance coverage, legitimate skip codes were assigned to the variables pertaining to reasons that the R never had coverage. That is, the implicit assumption made in the computer-assisted interviewing (CAI) skip logic was verified by an answer indicating some prior history of health insurance coverage. However, if nothing was reported to indicate that the R previously had health insurance, the skipped variables pertaining to reasons for never having had insurance retained codes of blank.
The R was male but reported in QHI17 that he lost health insurance coverage because he "received Medicaid or medical insurance only while pregnant."	The edited variable HLLOSRSN (corresponding to QHI17) was set to bad data.

13.2.5 Editing Types of Health Insurance and Any Other Health Insurance Variables, Imputation Sets 1 and 2

The editing process for health insurance combined the source variables (MEDICAID and CHIPCOV) to create the variable CAIDCHIP (whether someone was covered by Medicaid or one of the state children's health insurance programs). This CAIDCHIP variable and all the other source variables in [Table 13.2](#), except HLTINNOS, were used directly as base variables for imputation.

Table 13.2 Mapping of Raw Health Insurance Variables to Edited Variables

Question Number ¹	Question Text ²	Source Variable ³	Used to Create CAIDCHIP	Used to Create MEDICARE	Used to Create CHAMPUS	Used to Create PRIVATE	Used to Create ANYOTHER
QHI01, QHI01v	Is the respondent covered by Medicare?	MEDICARE (1 = yes, 2 = no)	No	Yes	No	No	Yes
QHI02, QHI02v	Is the respondent covered by Medicaid or Medical Assistance?	MEDICAID (1 = yes, 2 = no)	Yes	No	No	No	Yes
QHI02A	Is the respondent currently covered by a Children's Health Insurance Program operated by your state of residence? ⁴	CHIPCOV (1 = yes, 2 = no)	Yes	No	No	No	Yes
QHI03	Is the respondent currently covered by TRICARE, CHAMPUS, CHAMPVA, the VA, or military health care?	CHAMPUS (1 = yes, 2 = no)	No	No	Yes	No	Yes
QHI06	Is the respondent currently covered by private health insurance?	PRVHLTIN (1 = yes, 2 = no)	No	No	No	Yes	Yes

Table 13.2 Mapping of Raw Health Insurance Variables to Edited Variables (continued)

Question Number ¹	Question Text ²	Source Variable ³	Used to Create CAIDCHIP	Used to Create MEDICARE	Used to Create CHAMPUS	Used to Create PRIVATE	Used to Create ANYOTHER
QHI11	Is the respondent currently covered by any kind of health insurance, that is, any policy or program that provides or pays for medical care?	HLTINNOS (1 = yes, 2 = no, 99 = legitimate skip ⁵)	No	No	No	No	Yes

¹ The "v" questions were asked to verify the answer given in the previous question for respondents who were younger than 65 and a Medicare recipient or older than 65 and a Medicaid recipient.

² The questions provided in this table are abbreviated versions of those given in the questionnaire.

³ Missing values in these edited values were represented by standard missing value codes. CHIPCOV was replaced in the final analytic file by CAIDCHIP, a combination of MEDICAID and CHIPCOV. See Section 13.2.2 for details.

⁴ The questionnaire did not ask the question exactly in this way. It identified the specific program, depending upon the state of residence entered by the respondent and was asked only of respondents aged 12 to 19.

⁵ Respondents were assigned a legitimate skip for HLTINNOS if they answered "yes" or gave no answer to at least one of the other health insurance questions.

A respondent was routed to QHI11 (whether the respondent was covered by any kind of health insurance at the time of the survey) if they answered "no" to all the other health insurance questions. All other respondents were given a legitimate skip value to the variable HLTINNOS, as shown in Table 13.2. Therefore, it was possible that the imputation-revised versions of the four specific health insurance variables would all have had a value of "no," and the value of HLTINNOS would have been a legitimate skip, if one or more of the "no" values was imputed. In this instance, another variable was needed to reflect the fact that a respondent could have had a valid "yes" or "no" imputed value for any other health insurance, even though the respondent was never asked QHI11 and was assigned a legitimate skip code. The ANYOTHER variable was created using the preliminary edited variable, SKHLCCOV (Section 13.2.3), which indicated whether a respondent was covered by any health insurance. SKHLCCOV and ANYOTHER were defined as follows:

SKHLCCOV =

- 1 (or 3¹⁷⁸) if CAIDCHIP = 1, MEDICARE = 1, CHAMPUS = 1, PRVHLTIN = 1, or HLTINNOS = 1; else

¹⁷⁸ See Section 13.2.3 for discussion of the assignment of a code of 3 to SKHLCCOV. Respondents with SKHLCCOV = 3 were treated in the same manner as those with SKHLCCOV = 1.

- 2 if CAIDCHIP = 2, MEDICARE = 2, CHAMPUS = 2, PRVHLTIN = 2, and HLTINNOS = 2; else
- missing value code if the nonmissing values of CAIDCHIP, MEDICARE, CHAMPUS, PRVHLTIN, and HLTINNOS are all "2," and at least one of these variables had a missing response.

ANYOTHER =

- legitimate skip code (99) if CAIDCHIP = 1, MEDICARE = 1, CHAMPUS = 1, or PRVHLTIN = 1; else
- SKHLCCOV.

13.3 Imputation for the Health Insurance Variables

Beginning in the 2013 NSDUH, missing values for the health insurance variables were replaced with valid values using the health insurance method,¹⁷⁹ whereby indicator variables for more specific types of health insurance (i.e., coverage by Medicare or by Medicaid) were imputed using the PMN methodology in two imputation sets.

Imputation set 1 is described in Section 13.3.1 and includes the variables IRMCDCHP, IRMEDICR, IRCHMPUS, and IRPRVHLT. These are indicators of coverage for Medicaid/CHIP, Medicare, CHAMPUS or similar coverage for military personnel, and private health insurance, respectively. Imputation set 2 is described in Section 13.3.2 and includes the variable IROTHHLT, which indicates "any other health insurance." Together, these five imputation-revised health insurance variables were then used to create a third indicator of "any health insurance" coverage, IRINSUR4, discussed in Section 13.3.3.

The health insurance variables that undergo imputation tend to have item response rates of more than 99 percent. See Chapter 3 for details on PMN; see [Table A.29](#) in Appendix A for details on rates of missingness for the health insurance variables.

13.3.1 Types of Health Insurance Variables, Imputation Set 1

The first four imputation-revised variables were created using the single RP/multiple PRD PMN type. This approach was appropriate due to the high response rates for the variables being imputed in this set, and there were no noteworthy deviations from this general approach. The four imputation-revised variables created were yes/no indicators of whether the respondent had health insurance from Medicaid or a state children's health insurance program (IRMCDCHP), Medicare (IRMEDICR), CHAMPUS or other similar coverage for military personnel (IRCHMPUS), or private health insurance (IRPRVHLT). The health insurance

¹⁷⁹ The health insurance method varies slightly from other methods used to impute variables for NSDUH. First, it uses some uncommon constraints and covariates. For example, from the interviewer-administered demographics section, the SERVICE variable, an indicator of service in the United States armed forces, was used both as a covariate and in a likeness constraint related to CHAMPUS. Second, age groups were created to ensure reasonable domain sizes for both modeling and hot-deck steps. Finally, provisional imputation steps in imputation set 1 were not actually hot-deck steps but simple stochastic imputations based on predicted means, and in the final hot-deck step for these variables, different constraints were applied to different age groups.

indicators were imputed from base variables in the following order: CAIDCHIP, MEDICARE, CHAMPUS, and PRVHLTIN.

13.3.1.1 Response Propensity Step

The response propensity models for imputation set 1 utilized the final analysis weight, ANALWT. All respondents were in the domain for this stage. For the RP model, a domain member was considered an item respondent if all four base variables (CAIDCHIP, MEDICARE, CHAMPUS, and PRVHLTIN) were nonmissing. No respondents received skip codes for the final imputation-revised variables IRMCDCHP, IRMEDICR, IRCHMPUS, and IRPRVHLT. See [Table D.75](#) in Appendix D for details of the covariates used in the RP models for these variables.

13.3.1.2 First Prediction Step (CAIDCHIP)

CAIDCHIP was a dichotomous base variable; after imputation, its only values were "yes" and "no." Therefore, it was modeled using logistic regression, specifically the RLOGIST procedure in SAS-callable SUDAAN. The single predicted mean from this model (the probability that the respondent received Medicaid or was covered by the State Children's Health Insurance Program) is used in the final hot-deck step.

13.3.1.3 First Provisional Stochastic Imputation Step (CAIDCHIP)

The first provisional imputation step used for CAIDCHIP was not a hot-deck step, but a stochastic imputation. Each recipient was assigned a "yes" response for the intermediate imputation-revised variable INTCDCH with probability equal to the predicted mean from the model. This method was based on the idea of a centered PMN, as described in Singh, Grau, and Folsom (2004). As applied to single dichotomous variables like CAIDCHIP, the centered PMN approach reduces to a simple stochastic imputation that requires no donors and no constraints. This approach to imputation for categorical variables is further explored in the evaluation of imputation methods for the NSDUH report (Center for Behavioral Health Statistics and Quality, 2017e).

13.3.1.4 Analogous Prediction and Provisional Hot-Deck Imputation Steps (MEDICARE, CHAMPUS, and PRVHLTIN)

PRD models were fit for MEDICARE, CHAMPUS, and PRVHLTIN in the same manner as for CAIDCHIP, as described above, with a few notable deviations:

- For the MEDICARE variable, a single PRD model was fit for the 18-64 age group. This was done because (1) only a small proportion of respondents in these age groups had Medicare, particularly for the 18-25 age group, and (2) a respondent of working age could have received Medicare only if he or she was not working because of disability. This was true regardless of whether the respondent was aged 18 to 25 or 26 to 64.
- The CHAMPUS PRD models used a covariate that underwent a simplified imputation procedure. For respondents aged 18 or older, models included an indicator of whether the respondent had ever been in the military service, designated by an imputation-

revised version of the edited variable SERVICE. The variable SERVICE generally had a very low level of missingness (two missing values in the 2016 survey). Because covariates in these models were not supposed to have any missing values, the missing values in the SERVICE variable were randomly imputed as "yes" responses if the random number was greater than the mean value of SERVICE across all the other respondents, and imputed as "no" otherwise.

- PRVHLTIN did not undergo a provisional imputation step because it was the last variable in the set.

13.3.1.5 Final Hot-Deck Step

Details on the missingness patterns, constraints, and predictive mean vectors for the final hot-deck step of this first stage of imputation are available in Appendix E. This section explains the general ideas behind the hot-deck step applied to these variables.

There are 15 missingness patterns for the first stage of health insurance imputation.¹⁸⁰ No logical constraints were applied, because there were no logical relationships among the variables (or between the variables and any other NSDUH variables). The predictive mean vectors were also made simple because of the absence of logical relationships. Nonetheless, the likeness constraints were fairly sophisticated and sometimes varied across age groups, exploiting auxiliary variables from the demographics and income sections. The SERVICE variable was also used, but not an imputation-revised version of it; the constraint only applied when the item nonrespondent had a nonmissing value. The vast majority of item nonrespondents in imputation set 1 are typically handled on the first attempt to find a donor because (1) the item response rates are high, (2) the domain includes all unit respondents, and (3) the constraints are not very restrictive.

13.3.2 Any Other Health Insurance Variable, Imputation Set 2

The fifth and final imputation-revised variable, IROTHHLT, was created from the base variable ANYOTHER. This variable indicated whether respondents had any other type of health insurance, even though they reported or were imputed to have none of the four specific types of specific health insurance, as recorded by IRMCDCHP, IRMEDICR, IRCHMPUS, and IRPRVHLT. The PMN imputation type used for this variable was single RP/single PRD, and there were no noteworthy deviations from this general approach.

For this second stage, the age groups were 12 to 17, 18 to 25, and 26 or older. Three age groups were used instead of four because of the small number of respondents who would have otherwise comprised the domain for the age group of 65 or older.

13.3.2.1 Response Propensity Step

The response propensity models for imputation set 2 utilized the final analysis weight, ANALWT. The domain for the ANYOTHER variable included respondents who had either a

¹⁸⁰ Each of the four variables in the set can be missing or nonmissing. This leads to $2 \times 2 \times 2 \times 2 = 16$ possible combinations, but the combination where none are missing is one that does not require imputation.

reported or imputed "no" value to all four imputation-revised specific health insurance variables from the first stage (IRMCDCHP, IRMEDICR, IRCHMPUS, and IRPRVHLT). For a domain member to be considered an item respondent, he or she had to have complete data for the variable ANYOTHER. See [Table D.75](#) in Appendix D for details of the covariates used in the RP model for this variable.

13.3.2.2 Prediction Step

Like all the other health insurance variables that undergo imputation, ANYOTHER was a dichotomous variable; after imputation, its only values were "yes" and "no." Like the variables in the first stage, it was modeled using logistic regression, as implemented by the RLOGIST procedure in SAS-callable SUDAAN. The single predicted mean that was output by the RP step and used in the later hot-deck step was the predicted probability that the respondent received other health insurance, given that he or she was not covered by Medicaid/CHIP, Medicare, CHAMPUS, or private health insurance.

13.3.2.3 Hot-Deck Step

The hot-deck step for the ANYOTHER variable was the simplest one used in NSDUH. There were no logical constraints, the only likeness constraint was the delta constraint, and the predictive mean vector was actually a scalar. Approximately 1 percent of the records in the domain underwent imputation, and nearly every item nonrespondent was handled on the first attempt to find a donor.

13.3.3 Health Insurance Recodes for Additional Analyses

In addition to the five imputation-revised variables, the overall health insurance variable, IRINSUR4, was created during the imputation process. This variable is created by combining IRMCDCHP, IRMEDICR, IRCHMPUS, IRPRVHLT, and IROTHHLT. If a respondent had a reported or imputed "yes" value for any of these five variables, the respondent was considered to have health insurance. Otherwise, he or she did not have health insurance. Though IRINSUR4 was technically a recoded variable created from other variables, an imputation indicator (IIINSUR4) was nevertheless created. IIINSUR4 was set to "1" if the respondent had a reported "yes" value for any of the five health insurance variables or a reported "no" for all five of them; and "3" otherwise.

This page intentionally left blank

14. Editing and Imputation for the NSDUH Pair Variables

14.1 Introduction

Each household selected for the National Survey on Drug Use and Health (NSDUH) has zero, one, or two household members selected for interviewing. When two members of the same household, a pair, are selected, the data can be used to study outcome variables based on the relationship between the pair (i.e., the pair relationship). The outcome variables can be at either the person level or the pair level. The most common type of analysis is the person-level analysis, where the inferential population is defined by one of the pair members (called the focus pair member). An example of an outcome at the person level is the proportion of children who use drugs and whose parents report talking to them about drugs. The focus is on the child in a parent-child pair. By contrast, an example at the pair level is parent-child drug behavior for all possible parent-child pairs (within the child's age group). This chapter describes the techniques used to edit and impute the pair variables from the household roster (hereafter, "pair variables") for the 2016 NSDUH.¹⁸¹ The variables described in this chapter can be divided into three stages:

- Stage One: Pair Relationships,
- Stage Two: Multiplicity Counts, and
- Stage Three: Household-Level Person Counts.

The creation of the edited and imputation-revised pair variables was conducted in three stages because the variables from earlier stages were needed for the creation of variables in later stages. Stage one consisted of the creation and imputation of the variables that identify the pair relationships. The multiplicity and household-level person counts were created and imputed in stages two and three, respectively. Missing values in all three stages were imputed using the predictive mean neighborhood (PMN) imputation procedure, which uses predicted means from models to find donors in a nearest neighbor hot deck. Chapter 3 of this report provides background information about imputation in general (including hot-deck imputation) and details about the PMN methodology in particular.

Though this chapter presents first the editing procedures applied to the variables in each stage and then the imputation procedures for each stage, it is important to note that the actual order of processing was by stage; that is, both editing and imputation were completed for the variables in each stage before moving on to the next stage.

¹⁸¹ Edits to the variables for the parenting experiences section, which also involves selection of a pair, are discussed in Section 9.6 in Chapter 9. The parenting experiences section was administered if (1) an adult and an adolescent pair were selected for an interview, and (2) the adult was the parent of the adolescent who also was selected for an interview.

14.1.1 Pair Relationship Variables

The pair relationship variables were derived from the household composition (roster) variables, described in Chapter 11 of this report. Pair relationship variables include the edited and imputation-revised pair relationship variables PAIRREL¹⁸² and IRPRREL, as well as the imputation indicator IIPRREL, which summarizes how the data in IRPRREL were obtained. PAIRMEM was created to indicate that a respondent was part of a respondent pair. RELMATCH was created to detail how well the pair members matched on their relationship to one another. The variable PRNTIND identified whether the respondent was a parent in a parent-child relationship. Finally, three additional variables were created to aid in pair analyses by giving details about the other pair member: AGEOTHER, SEXOTHER, and PAIRID. AGEOTHER contained the age of the other respondent in the pair; SEXOTHER contained the gender of the other respondent in the pair; and PAIRID contained the questionnaire ID (QUESTID) of the other pair member.

14.1.2 Multiplicity Count Variables

"Multiplicity" is the term used to describe the complication that arises in the analysis of pair data at the person level for a given pair domain; several pairs in the household could be associated with the same person. Whereas, when analyses are at the pair level, the pair domain is completely and uniquely defined by the pair relationship. For example, to tabulate the number of sibling-sibling pairs where both siblings have used marijuana, it is necessary to know only whether a pair of respondents contains two siblings. By contrast, for person-level analyses, the pair domain depends upon which pair member is the focus. In this case, to analyze the influence that older siblings have on younger siblings in terms of drug use, it is necessary to know which pair member is the older sibling. The multiplicity count is a count of the number of pairs in the household that can be associated with the person of focus because—to continue the example above—a child may have more than one older sibling. If the multiplicity is not considered, then individuals of focus who are members of more pairs receive a weight that is too large (i.e., they are counted too many times in the target population), and individuals of focus who are members of fewer pairs receive a weight that is too small. Adjustment for multiplicity was not done a priori because that would lead to the addition of 12 weights to the file, one for each of the 12 possible pair relationships that require this adjustment (e.g., excluding the spouse-spouse and spouse-spouse with children relationships). Note that the multiplicity problem does not arise if there is only one inclusion possibility (e.g., a single-parent household, if the child is the focus) or if the analysis is a pair-level analysis (e.g., parent-child pair drug behavior).

To illustrate how multiplicities appear in the definitions of parameters and estimates, consider estimating the total number of children who used drugs in the past year, where a parent reported talking to them about drugs. Let $y_{hip}(d)$ be defined as the drug-related behavior outcome for pair p containing the individual i belonging to domain d in household h . Now, for the population of all individuals who belong to the domain d , the total parameter is defined as

¹⁸² The levels of PAIRREL are provided in [Table 14.3](#).

$$\tau_y(d) = \sum_{h=1}^H \sum_{i=1}^{N_h(d)} \sum_{p=1}^{M_{hi}(d)} \frac{y_{hip}(d)}{M_{hi}(d)}$$

(i.e., total of averages over pairs (p) associated with the individual i , over all i in domain d and in the household h) (Chromy & Singh, 2001). Here $M_{hi}(d)$ denotes the multiplicity (i.e., the number of pairs associated with the individual i in domain d), and $N_h(d)$ can be thought of as the multiplicity count for the household h (i.e., the number of people in the household that are in domain d).

Multiplicity count variables were only created for specific relationships of interest. These variables are listed in Table 14.1. With the exception of the spouse-spouse multiplicity variables listed at the bottom of this table, these variables all underwent imputation.

Table 14.1 Edited and Imputation-Revised Multiplicity Count Variables

Variable Description	Edited Variable Name	Imputation-Revised Variable Name
Number of parents of a child aged 12-14 who is a member of a parent-child pair	MCPCC14	IRMPCC14
Number of children aged 12-14 belonging to a parent who is a member of a parent-child pair	MCPCP14	IRMPCP14
Number of parents of a child aged 15-17 who is a member of a parent-child pair	MCPCC57	IRMPCC57
Number of children aged 15-17 belonging to a parent who is a member of a parent-child pair	MCPCP57	IRMPCP57
Number of parents of a child aged 12-17 who is a member of a parent-child pair	MCPCC17	IRMPCC17
Number of children aged 12-17 belonging to a parent who is a member of a parent-child pair	MCPCP17	IRMPCP17
Number of parents of a child aged 12-20 who is a member of a parent-child pair	MCPCC20	IRMPCC20
Number of children aged 12-20 belonging to a parent who is a member of a parent-child pair	MCPCP20	IRMPCP20
Number of siblings aged 12-14 for a respondent aged 15-17 who is a member of a sibling-sibling pair	MCS1417	IRMS1417
Number of siblings aged 15-17 for a respondent aged 12-14 who is a member of a sibling-sibling pair	MCS1714	IRMS1714
Number of siblings aged 12-17 for a respondent aged 18-25 who is a member of a sibling-sibling pair	MCS1725	IRMS1725
Number of siblings aged 18-25 for a respondent aged 12-17 who is a member of a sibling-sibling pair	MCS2517	IRMS2517
Pair relationship is spouse-spouse	MCSPSP	N/A
Pair relationship is spouse-spouse with children younger than 18	MCSPSPWC	N/A

N/A = not applicable.

14.1.3 Household-Level Person Count Variables

The household-level person count variable is the number of individuals of focus in the household for a given pair domain, regardless of which pair (or whether a pair) was actually selected. For example, if two parents were in the household with three children aged 12 to 14, the household person count for the *parent* focus of the parent-child (12-14) domain would be 2, and the household person count for the *child* focus would be 3. If the parents indicated that they had a spousal relationship, the household person count for the spouse-spouse with children domain would be 2. The rest of the household counts would be zero (even the sibling focus) because they do not fall in the age groups of the domains that are edited and imputed; that is, counts were not created for siblings that were both aged 12 to 14. Also, note that household person counts for all domains were calculated for every respondent, even when only one respondent was selected in the household or when a selected pair did not fall in a particular domain. The household-level person count variables are listed in [Table 14.2](#), and all underwent imputation.

Table 14.2 Edited and Imputation-Revised Household-Level Person Count Variables

Variable Description	Edited Variable Name	Imputation-Revised Variable Name
Number of children aged 12-14 in the household with at least one parent living with them	HCPCC14	IRHPCC14
Number of parents in the household with at least one child aged 12-14 living with them	HCPCP14	IRHPCP14
Number of children aged 12-17 in the household with at least one parent living with them	HCPCC17	IRHPCC17
Number of parents in the household with at least one child aged 12-17 living with them	HCPCP17	IRHPCP17
Number of children aged 12-20 in the household with at least one parent living with them	HCPCC20	IRHPCC20
Number of parents in the household with at least one child aged 12-20 living with them	HCPCP20	IRHPCP20
Number of household members aged 15-17 with a sibling aged 12-14 living with them	HCS1417	IRHS1417
Number of household members aged 18-25 with a sibling aged 12-17 living with them	HCS1725	IRHS1725
Number of spouse-spouse pairs	HCSPSP	IRHCSPSP
Number of spouse-spouse pairs with children younger than 18	HCSPSPWC	IRHCSPWC

14.2 Stage One Editing: Pair Relationships

14.2.1 Editing the Household Roster of Each Pair Member

Prior to identifying the relationships between selected pair members, a key step is editing the questionnaire household rosters for each pair member. This involves identifying situations where the relationship listed in the roster for a particular roster member was not possible given

the roster member's age and relationship to the respondent. In the majority of cases where the relationships could not be determined, this resulted in setting the relationship code to bad data and sometimes setting the roster member's age to bad data as well. In general, no effort was made to try to match the values of roster-derived household composition variables between pair members, because interviews of the different members of the same household could have taken place at different times. However, information from other pair members was sometimes used to change a relationship code from one value to another, instead of setting the relationship code to bad data.

14.2.2 Creating the Pair Relationship Variable (PAIRREL)

The process of identifying the pair relationships involved two steps: (1) match the household rosters of the pair members, and (2) determine the pair relationships using the relationship codes and ages of the matched rosters, if they could be determined. The first step is described in Section 14.2.2.1 and Appendix H, and the second step is described in Section 14.2.2.2 and Appendix H. Any relationships that could not be determined by this process were imputed as described in Section 14.5.2.

No restrictions were placed on the types of pairs that could be selected for inclusion in the NSDUH sample. However, the variable PAIRREL includes a limited number of relationships. This is because the identification of the particular relationships between a given pair was limited by the relationship codes that were available: parent, child, grandparent, grandchild, sibling, spouse, unmarried partner, roommate, parent-in-law, child-in-law, boarder, other relative, and other nonrelative. (This precluded the possibility of identifying an uncle-nephew relationship, for example.)

Furthermore, because the creation of the multiplicity factors required complicated programming logic, multiplicities could not be created for all possible pair relationships. Only certain pair relationships considered to be of higher analytic interest were included in the creation of multiplicities:

- parent-child, child aged 12 to 14;
- parent-child, child aged 12 to 17;
- parent-child, child aged 15 to 17;
- parent-child, child aged 12 to 20;
- sibling-sibling, younger sibling aged 12 to 14, older sibling aged 15 to 17;
- sibling-sibling, younger sibling aged 12 to 17, older sibling aged 18 to 25;
- spouse-spouse (includes partner-partner), with children younger than 18;¹⁸³ and
- spouse-spouse (includes partner-partner), with or without children.

¹⁸³ The spouse-spouse pair relationship included respondents who were legally married, as well as respondents who lived together as though married (partners). Although the questionnaire distinguished between "spouses" and "partners," the pair relationship variable being described here did not distinguish between the two. In rare instances, a spouse-spouse pair included one pair member who identified the second pair member as a spouse, whereas the second pair member identified the first as a partner.

The levels in PAIRREL, summarized in [Table 14.3](#), do not correspond exactly with those given above, but the relevant pair relationships can be derived from the value of PAIRREL. For example, a value of PAIRREL = 3 indicates that, among the pair relationships given above, the pair relationship was a parent-child pair with a child aged 18 to 20.

Table 14.3 Levels of the Variable PAIRREL

Value of PAIRREL	Interpretation	Domain of Interest
1	Respondent is part of a parent-child (12-14) pair	Yes
2	Respondent is part of a parent-child (15-17) pair	Yes
3	Respondent is part of a parent-child (18-20) pair	Yes, indirectly
4	Respondent is part of a parent-child (21+) pair	No
5	Respondent is part of a sibling (12-14)-sibling (15-17) pair	Yes
6	Respondent is part of a sibling (12-17)-sibling (18-25) pair	Yes
7	Respondent is part of another sibling-sibling pair	No
8	Respondent is part of a spouse-spouse ¹ pair, with children in the household younger than 18	Yes
9	Respondent is part of a spouse-spouse pair, with no children in the household younger than 18	Yes
10	Respondent is part of a spouse-spouse pair, but it is unclear whether children younger than 18 in the household belong to the pair	Yes
11	Respondent is part of a grandparent-grandchild pair	No
12	Respondent is part of another clearly identifiable pair	No
13	Respondent is part of a pair that is not clearly identifiable, but it is clear from the relationship codes that it is not within codes 1 through 11	No
14	Respondent is part of a pair that is not clearly identifiable, and it could be any pair relationship	Maybe
15	Respondent is part of a pair that is not clearly identifiable, but it could be pair codes 1 or 12	Maybe
16	Respondent is part of a pair that is not clearly identifiable, but it could be pair codes 2 or 12	Maybe
17	Respondent is part of a pair that is not clearly identifiable, but it could be pair codes 3 or 12	Maybe
18	Respondent is part of a pair that is not clearly identifiable, but it could be pair codes 4 or 12	No
19	Respondent is part of a pair that is not clearly identifiable, but it could be pair codes 5 or 12	Maybe
20	Respondent is part of a pair that is not clearly identifiable, but it could be pair codes 6 or 12	Maybe
21	Respondent is part of a pair that is not clearly identifiable, but it could be pair codes 7 or 12	No
22	Respondent is part of a pair that is not clearly identifiable, but it could be pair codes 8 or 12	Maybe

Table 14.3 Levels of the Variable PAIRREL (continued)

Value of PAIRREL	Interpretation	Domain of Interest
23	Respondent is part of a pair that is not clearly identifiable, but it could be pair codes 9 or 12	Maybe
24	Respondent is part of a pair that is not clearly identifiable, but it could be pair codes 8, 9, or 12	Maybe
25	Respondent is part of a pair that is not clearly identifiable, but it could be pair codes 11 or 12	No
99	Respondent is not a member of a pair	No

¹ The pair relationship labeled "spouse-spouse" includes partner-partner pair relationships.

14.2.2.1 Matching the Household Rosters

For the purpose of discussing how to match the household rosters in this report, the pair members are identified as pair member "A" and pair member "B." For the household roster of pair member A, it was necessary to determine which household member listed in A's roster corresponded to the other selected pair member. The same had to be done for pair member B. This was accomplished using the age and gender of the pair members, in addition to the variable MBRSEL, which was used to identify the roster member corresponding to the other selected pair member.

In a perfect setting, the questionnaire age and gender of pair member B (AGE and IRSEX, respectively) would have corresponded exactly to the age and gender entered for one of the members of pair member A's household roster (RAGE and RSEX). Moreover, the value of MBRSEL for this matched roster member would have been 1, and the value of MBRSEL for all other roster members would have been zero or missing. Furthermore, in this perfect setting, matches with exactly one MBRSEL = 1 correctly identifying the other pair member also would have been found with pair member B's roster. This did not always occur, however, so some effort was required to determine the roster member most likely to correspond to the other selected pair member.

The quality of the match between pair members varied depending upon the quality of the roster entries and the time between interviews. A number of if-then-else conditions, called priority conditions (because of the hierarchical nature of the conditions), gave a pair match that was considered valid in the vast majority of cases. These conditions are provided in Appendix H. In general, the conditions matched IRSEX and AGE for the one pair member against the gender and age of the roster members in the other pair member's roster, using MBRSEL to help identify the appropriate roster member. For a match to be considered valid, it was necessary that at least one of the two pair members have a match.

In the cases where a single roster member had to be selected among duplicates and where the duplicates had the same relationship code, it was necessary to limit the relationship codes to child or sibling. In some cases, because of the poor quality of the rosters of the pair members, it was not possible to locate the household member listed in A's roster that corresponded to pair member B, and vice versa. The determination of the pair relationships for these cases was left to

imputation. Even when a pair of roster members was successfully identified, it was not always possible to successfully determine the pair relationship, as described in the next section.

14.2.2.2 Determining the Pair Relationship Using the Relationship Codes of the Matched Rosters

Once the pair was identified, two observations per household resulted, each with a relationship code corresponding to the other selected pair member. The relationship codes for these two observations had to be matched to determine the pair relationship. For example, suppose a 15-year-old and a 38-year-old were selected to be interviewed. If the 38-year-old was subsequently identified as the parent on the 15-year-old's roster, and the 15-year-old was identified as the child of the 38-year-old on the 38-year-old's roster, then the pair relationship would be identified as PAIRREL = 2, according to the levels of PAIRREL provided in [Table 14.3](#). Thus, these two individuals would belong to the following pair relationships of interest: parent-child with child aged 15 to 17, parent-child with child aged 12 to 17, and parent-child with child aged 12 to 20. As noted earlier, the pair relationship of interest was derived from the values of PAIRREL. In particular, the parent-child with child aged 12 to 17 and parent-child with child aged 12 to 20 domains were derived from the parent-child pair relationships created using 12- to 14-year-olds, 15- to 17-year-olds, and 18- to 20-year-olds, the levels referenced in PAIRREL. Moreover, the overall spouse-spouse domain was derived from the two spouse-spouse pair relationships with and without children.¹⁸⁴

As with the procedure used to match the household rosters, a series of conditions was used to identify the relationship between pair members. These priority conditions used ages and relationship codes to identify the pair relationships. In a perfect setting, the relationship codes would be nonmissing and in agreement between the pair members, as in the example given in the previous paragraph. In some instances, however, either the relationship codes were missing or they did not agree across the pair members. The priority conditions offer a method for interpreting the relationship codes in such cases.

Below are the strategies used to identify a pair relationship in an imperfect setting:

1. If a relationship code was missing on one side of the pair but not on the other, then the pair relationship was assumed to be identified by the nonmissing relationship code. The exception to this rule occurred if the identified relationship was parent-child with a child younger than 18, the "parent" was less than 10 years older than the child, and the "parent" answered the parenting experiences question (FIPE3) by saying that the other respondent was not his or her child. In this case, the nonmissing relationship code was considered spurious, and the relationship was left missing.
2. If it was not possible to definitively determine the relationship between the pair members using the relationship codes, but the relationship codes on both sides indicated that the unknown pair relationship was not a relationship of interest, then the pair relationship was identified as such and no imputation was required. For example, if pair member "A" identified pair member "B" as a "boarder," but pair member "B" identified pair member "A" as "other relative," then the relationship was

¹⁸⁴ The spouse-spouse pair relationship includes partner-partner pair relationships.

not a relationship of interest and the variable PAIRREL would have been assigned the value "13" (Table 14.3).

3. If it was not possible to definitively determine the relationship between the pair members using the relationship codes, but a parent-child relationship was possible given the relationship code in one of the pair member's rosters, then the FIPE3 variables were used to assist in the determination of a pair relationship. For example, consider a pair member who was a stepparent and refers to his or her stepchild as "child," but the child refers to the stepparent as "other nonrelative." Membership in a parent-child relationship where the child was younger than 18 was indicated if the stepparent answered FIPE3 affirmatively, thereby proceeding to the parenting experiences section. On the other hand, if the stepparent answered FIPE3 negatively, then the stepparent was not considered the parent. A third scenario arose if the FIPE3 answer was not given. In this case, a parent-child relationship was assumed if the stepparent was legally married and the child identified the spouse of the other pair member as "parent."

The quality of the match for PAIRREL levels 1 through 25 is indicated by the variable RELMATCH, the levels of which are summarized in Table 14.4. In general, imputation was required for values of RELMATCH of 0 or 4, or if PAIRREL = 10. PAIRREL = 10 was a special case because it was clear that a relationship "of interest" always would have been involved. For this value of PAIRREL, the value of RELMATCH was equal to 1 or 2. However, imputation was still required because it was not clear whether children were in the household. The number of cases that were matched or not matched, as indicated by the RELMATCH variable (or PAIRREL = 10), for the 2016 survey is provided in Table 14.5. The amount of imputation required was dependent upon the quality of the rosters. The attributes of the roster are described in Chapter 11.

Table 14.4 Values of PAIRREL That Correspond to the Levels of the Variable RELMATCH

Value of RELMATCH	Values of PAIRREL	Interpretation
0	14	FAILURE: The relationship was not identifiable and could have been a relationship of interest.
1	1-9, 11-13	SUCCESS: The relationship was clearly identifiable using information from both pair members or was unmistakably not a relationship of interest.
1	10	FAILURE: A spouse-spouse ¹ relationship was definitively established using information from both pair members, but it was unclear whether the pair had children in the household.
1.5	8	SUCCESS: A spouse-spouse relationship was definitively established using information from both pair members, and children younger than 18 were in both rosters. Relationship codes on one side of the pair indicated children belonged to the pair, and on the other side of the pair, the relationship codes corresponding to the children were missing.

Table 14.4 Values of PAIRREL That Correspond to the Levels of the Variable RELMATCH (continued)

Value of RELMATCH	Values of PAIRREL	Interpretation
2	1-9, 11-13	SUCCESS: The relationship was clearly identifiable using information from one pair member, while the relationship code from the other pair member was missing.
2	10	FAILURE: A spouse-spouse relationship was definitively established using information from one pair member, while the relationship code from the other pair member was missing. It was unclear whether the pair had children in the household.
3	1-8, 12, 13	SUCCESS: Relationship information was conflicting between the pair members, but conclusions were drawn anyway for some parent-child pairs, some sibling-sibling pairs, and some spouse-spouse pairs using either information outside the household roster or logical reasoning. ²
4	15-25	FAILURE: Relationship information was not identifiable. Information was in conflict between the pair members, where one pair member indicated relationship of interest and the other did not. However, ages supported a relationship of interest (may be used to limit imputation).

¹ The pair relationship labeled "spouse-spouse" includes partner-partner pair relationships.

² In the case of potential parent-child pairs, further evidence that a parent-child relationship was involved or not involved was obtained by looking at the FIPE3 variable to see whether a stepparent had a spouse who corresponded to a child's parent or to see the ages of the respondents. For spouse-spouse relationships, two situations occurred. In the case where the respondents were not legally married, the children of one pair member were considered the children of the pair in the household, even though they were not identified as such by the other pair member. In the case where only one pair member referred to the other as a "married" or "unmarried partner," and if both had the same children, they were considered "spouse-spouse-with-children." The other pair member was usually referred to as a "roommate" or "other nonrelative."

Table 14.5 Frequencies of the Levels of the Variable RELMATCH: 2016

RELMATCH	Frequency (Percent)	Requires Imputation
0	43 (0.24)	Yes
1 (PAIRREL $\neq$ 10)	17,373 (97.34)	No
1 (PAIRREL = 10)	47 (0.26)	Yes
1.5	1 (0.01)	No
2 (PAIRREL $\neq$ 10)	87 (0.49)	No
2 (PAIRREL = 10)	1 (0.01)	Yes
3	115 (0.64)	No
4	180 (1.01)	Yes

14.3 Stage Two Editing: Multiplicity Counts

As stated earlier, multiplicities were required to account for analyses that were performed at the person level, even though the pair weights were calculated at the pair level. Because the multiplicities were relevant only at the person level, the definition of multiplicity required the

identification of the focus member of the pair. Using the pair relationships determined in Section 14.2, the following domains were considered:

1. parent-child (child 12 to 14), parent focus;
2. parent-child (child 12 to 14), child focus;
3. parent-child (child 15 to 17), parent focus;
4. parent-child (child 15 to 17), child focus;
5. parent-child (child 12 to 17), parent focus;
6. parent-child (child 12 to 17), child focus;
7. parent-child (child 12 to 20), parent focus;
8. parent-child (child 12 to 20), child focus;
9. sibling (12 to 14)-sibling (15 to 17), sibling (15 to 17) focus;
10. sibling (12 to 14)-sibling (15 to 17), sibling (12 to 14) focus;
11. sibling (12 to 17)-sibling (18 to 25), sibling (18 to 25) focus;
12. sibling (12 to 17)-sibling (18 to 25), sibling (12 to 17) focus;
13. spouse-spouse (includes partner-partner) with children younger than 18; and
14. spouse-spouse (includes partner-partner).

Determining the multiplicity entailed finding the number of roster pairs in the domain of interest that contained the focus member in the pair. In broad terms, the process of determining the multiplicity count involved two steps: (1) calculate the multiplicity count for each pair member, and (2) use the screener, quality of roster, and other means to determine the appropriate count if each pair member's counts did not match. The first step is described in Section 14.3.1, and the second step is described in Section 14.3.2 and Appendix I. Multiplicities that could not be determined through these steps were left to imputation as described in Section 14.5.3.

Because the pair weights reflect selection done at the time of screening, the multiplicity count should reflect the household makeup at that time. However, this was not possible in all cases, because for some households the screener roster was not as complete as the questionnaire roster, and recorded relationships in the screener roster were relative to the head of the household rather than to each pair member. No account was made for cases where a change in the household makeup occurred between the screening date and the dates of both interviews. In other words, due to the passage of time only, the observed change in household makeup could have occurred because of an intervening birthday or because a roster member left or entered the household after screening. Because an adjustment to the multiplicity counts would have been extremely complicated to implement for the small number of cases to which it applied, no such adjustment was made. Nevertheless, when possible, the screener was used to resolve cases where there were disagreements between pair members on the value of the multiplicity count.

14.3.1 Determining the Multiplicity Count for Each Pair Member

The multiplicity counts for each pair member consisted of a direct count and an indirect count. The direct count was obtained by looking at the focus pair member. It was a count of the roster members who could have been selected, where the same pair domain would have resulted. The indirect count was obtained by looking at the pair member who was not the focus. It was a count of the pair member himself or herself, plus other roster members who, by virtue of their relationship code, would have had the same pair relationship had they been selected.

A summary of the ways of determining the direct count and indirect count for each pair domain are provided in Table 14.6. For these domains, neither the direct nor the indirect count could be zero, because the pair member who was not the focus had to be part of the count. For spouse-spouse counts, no work was necessary to determine multiplicity counts. If a respondent was in a spouse-spouse pair, the multiplicity count was necessarily 1 in almost all cases, because only one spouse-spouse pair could have been selected that included that pair member. If the true multiplicity count exceeded 1, then the multiplicity count was set to 1.¹⁸⁵ Note that other spouse-spouse pairs in the household (e.g., one spouse's parents) would have been of interest in the household counts discussed in subsequent sections.

Table 14.6 Multiplicity Counts for Each Pair Member

Pair Relationship	Focus Member	Direct Count	Indirect Count
Parent-Child	Child	From child: number of parents	From parent: self + spouse/partner
Parent-Child	Parent	From parent: number of children in appropriate age range	From child: self + number of siblings in the appropriate age range
Sibling-Sibling	Older sibling	From older sibling: number of siblings in younger age range	From younger sibling: self + number of siblings in younger age range
Sibling-Sibling	Younger sibling	From younger sibling: number of siblings in older age range	From older sibling: self + number of siblings in older age range

14.3.2 Determining the Final Multiplicity Count

Once the counts were determined for each pair member, it was necessary to resolve any differences between these counts across pair members. In most cases, the direct and indirect counts agreed, with no bad relationship codes for either pair member, resulting in a straightforward determination of the final multiplicity count. This was usually possible if one pair member had bad relationship codes or had a count of zero, which meant that the final multiplicity count came from the pair member with good data.¹⁸⁶ For some cases, both pair members had bad relationship codes, which meant that the final multiplicity was left to

¹⁸⁵ In rare cases, it was possible for a respondent to have two or more spouses who lived in the same household. Determining the appropriate multiplicity count in these cases required knowledge of which spouse was the focus, which would be arbitrary. Because having multiple spouses was an extremely rare occurrence, and because of the complexity of determining the appropriate multiplicity count, these situations were not explicitly addressed during data processing.

¹⁸⁶ There were some exceptions to this rule. If the bad relationship codes were only within the relevant age ranges, then the count from the good side was used only if the age ranges in the good side matched the screener.

imputation. Among the remaining cases, some could be reconciled and some could not. In the cases where reconciliation was possible, many of the disagreements between the pair members were resolved by going to the screener. The method used to reconcile differing counts depended upon the domain. For the parent-child domains, for example, in addition to the screener, the parent/legal guardian variable, FIPE3, was used to help reconcile differences. Detailed rules for reconciling differences between pair members are provided in Appendix I.

If reconciliation between the counts from the two pair members in the household and the screener was not possible, upper and lower bounds within which the imputed value had to reside were determined from the counts for each pair member and the counts for the screener. The amount of imputation required for the multiplicity counts for the 2016 survey is shown in [Table 14.7](#). From this table, it is apparent that the greatest degree of uncertainty came with the determination of the number of parents in the child-focus parent-child domains. This occurred because, even though the parent-child pair relationship had been established, it often was unclear whether there was a second "parent" in the household.

Other domains had very little uncertainty. The counts of the number of children in the parent-focus parent-child domain, for example, were almost always definitively determined.

Table 14.7 Amount of Imputation Required for Multiplicities in Various Pair Domains: 2016

Pair Domain	Multiplicity	Missing Cases
Parent-Child (12-14), Child Focus	Number of parents	220
Parent-Child (12-14), Parent Focus	Number of children	84
Parent-Child (15-17), Child Focus	Number of parents	176
Parent-Child (15-17), Parent Focus	Number of children	82
Parent-Child (12-17), Child Focus	Number of parents	310
Parent-Child (12-17), Parent Focus	Number of children	80
Parent-Child (12-20), Child Focus	Number of parents	328
Parent-Child (12-20), Parent Focus	Number of children	78
Sibling (12-14)-Sibling (15-17), Older Sibling Focus	Number of younger siblings	92
Sibling (12-14)-Sibling (15-17), Younger Sibling Focus	Number of older siblings	90
Sibling (12-17)-Sibling (18-25), Older Sibling Focus	Number of younger siblings	86
Sibling (12-17)-Sibling (18-25), Younger Sibling Focus	Number of older siblings	90

14.4 Stage Three Editing: Household-Level Person Counts

In order to improve the quality of the estimates from the pair data through poststratification of the appropriate weights,¹⁸⁷ it was necessary to identify the household-level person counts for each domain. This entailed finding the number of individuals in the household who belonged to a particular domain, given that one member of a domain was selected as the

¹⁸⁷ See the questionnaire dwelling unit-level and person pair-level sampling weight calibration report in the 2016 NSDUH methodological resource book (Center for Behavioral Health Statistics and Quality, 2018b) for details.

focus. These counts were more difficult to derive than the multiplicity counts because all households were considered. Within each household, counts for any of the domains of interest were derived, regardless of whether the pair belonged to that domain or even whether a pair was selected at all. The counts were derived for 10 of the 14 pair domains described in Section 14.3. For the parent-child counts where the child was between 15 and 17, calculating the household counts was unnecessary.¹⁸⁸ For the other two remaining sibling-sibling domains, the reason is historical: They were added after the procedures were first developed, and there was insufficient time to develop the household counts for those domains. The domains where these counts were created are listed below:

1. parent-child (child 12 to 14), parent focus;
2. parent-child (child 12 to 14), child focus;
3. parent-child (child 12 to 17), parent focus;
4. parent-child (child 12 to 17), child focus;
5. parent-child (child 12 to 20), parent focus;
6. parent-child (child 12 to 20), child focus;
7. sibling (12 to 14)-sibling (15 to 17), sibling (15 to 17) focus;
8. sibling (12 to 17)-sibling (18 to 25), sibling (18 to 25) focus;
9. spouse-spouse (includes partner-partner) with children younger than 18; and
10. spouse-spouse (includes partner-partner).

Determining the household-level person counts was a two-step process: (1) calculate the household count for each respondent, whether a member of a pair or a single respondent; and (2) use the screener, quality of roster, and other means to determine the appropriate final count either by attempting to reconcile differing counts between pair members or by attempting to determine the appropriate count when information from only one roster was available. For households where only one respondent was selected, the matching step (step 2) was unnecessary. The first step is described in Section 14.4.1, and the second step is described in Section 14.4.2 and Appendix J. Household counts that could not be determined by this process were left up to imputation as described in Section 14.5.4.

Because the pair weights reflected selection at the time of screening, the household-level person counts should have reflected the household makeup at that time. As with the multiplicity counts, however, this was not entirely possible, so no account was made for cases where a change in the household makeup occurred between the screening time and the time of each interview. An explanation for why this was not possible for the multiplicity counts is described in Section 14.3. Moreover, as stated in that section, to implement such an adjustment would have been extremely complicated for the household-level person counts. Nevertheless, in cases where

¹⁸⁸ Because household counts were defined for everybody, it was possible to derive these counts using the counts for the parent-child domains where the child was between 12 and 14 and where the child was between 12 and 17. However, the multiplicity counts for the parent-child (15 to 17) domain had to be calculated and could not have been derived in as straightforward a way, because the multiplicity counts were only defined if the pair relationship corresponded to the pair domain of interest.

there were disagreements between pair members on the value of the household-level person count, the screener was used to resolve those disagreements.

14.4.1 Determining the Household-Level Person Count for Each Respondent

14.4.1.1 Parent-Child Domains

When obtaining household-level person counts for parent-child domains, the six parent-child domains previously listed were considered. In any household, the household-level person counts for parent-child domains were nonzero if at least one parent was present in the household with children within the relevant age range. In this instance, the child-focus count would have been the number of children in the household within that age range that belonged to the parent in the household, and the parent-focus counts would have been the number of parents. If more than one "family unit" (mother and/or father with children) lived within the household, the child-focus counts should have counted children from more than one set of parents, and the parent-focus counts should have counted two or more parents, at least one for each set of children.

One situation where this could occur was where three generations lived within the same household, with children in both the youngest and the second generations within the relevant age range. Using the youngest generation as the reference point, some of the parent's siblings (the grandparents' other children) were within the relevant age range. In this instance, the parent-child domains of the number of children would have included both the children of the parents and the children of the grandparents who were in that age range. The count of the number of parents included both the parents and the grandparents (and exceeded 2). Identifying more than one family unit in a household with children within the relevant age range under other scenarios (e.g., two sisters both with children within the relevant age range, both living within the same household) could not be determined from the data and had to be disregarded. Regardless of how many family units were in the household, counts had to be determined in different ways depending upon whether a parent-child pair "of interest" was selected or not.

Descriptions of how to obtain the household-level person counts are provided below for the parent-child domains outlined above. Parent-child pairs of interest with parent-focus and child-focus domains considered together are described first. In this instance, the pair actually belonged to a pair relationship where analysis using one or more of the domains listed was possible. This is followed by descriptions for other pairs and single respondents with parent-focus and child-focus domains considered separately.

14.4.1.1.1 Obtaining Counts for Parent-Child Domains (Parent-Focus and Child-Focus): Parent-Child Pairs, Child Younger than 21

If the pair was identified as parent-child and the three-generation situation described above was not apparent, the household-level child-focus person count was given by the parent-focus multiplicity count. Similarly, the household-level parent-focus person count was given by the child-focus multiplicity count. If a three-generation situation was identified and the grandparent also had children within the relevant age range, the number of children and the number of parents were adjusted appropriately. The final household count in this instance was

greater than the imputation-revised multiplicity count, which did not include all of the children in the household within the relevant age range.

14.4.1.1.2 Obtaining Counts for Child-Focus Parent-Child Domains: Other Pairs and Single Respondents

For other pairs¹⁸⁹ and single respondents, the following conditions were required to determine the household count for the number of children of parents in the household:

1. If the age of the respondent was within the relevant age range and that respondent had at least one parent, then the child-focus counts were determined in the same way as the parent-focus multiplicity counts: The count was of the "self" plus the respondent's siblings within the relevant age range. If the respondent's parents were not identified as living with him or her in the household, then the count was set to zero.
2. If the respondent had children within the relevant age range, then the count was of the respondent's children within that range. If the respondent also had older children who had children of their own within the relevant age range, then the count was of the respondent's children and grandchildren within the relevant age range.
3. If the age of the respondent was outside the relevant age range, but the respondent had parents living with them in the household and had siblings within the relevant age range, then the count was of the number of the respondent's siblings.
4. If the respondent had grandchildren within the relevant age range and the respondent also had children older than 25 or children-in-law living with them, then the count was the number of the respondent's grandchildren. The assumption was that the respondent's children or children-in-law were the parents of the respondent's grandchildren. The likelihood of this not being the case was small. In the case where a pair was selected, this was resolved by looking at the count of the other pair member.

14.4.1.1.3 Obtaining Counts for Parent-Focus Parent-Child Domains: Other Pairs and Single Respondents

For other pairs and single respondents, the following conditions were required to determine the household count for the number of parents of children in the household:

1. If the age of the respondent was within the relevant age range, then the count was of the number of the respondent's parents (which could be zero).
2. If the age of the respondent was outside the relevant age range but the respondent had siblings within the relevant age range, then the count was of the number of the respondent's parents (again, this could be zero).

¹⁸⁹ "Other pairs" included pairs that were not within a domain of interest because the age of at least one of the pair members was outside the relevant age range. For parent-child pairs, this applied to a pair with a child who was 21 or older. For sibling-sibling pairs, this applied to siblings where both were within the same age range (both were 12 to 14, 15 to 17, or 18 to 25) or at least one of the siblings was older than 25. "Other pairs" also are referenced in Sections 14.4.1.1.3 and 14.4.1.2.2.

3. If the respondent had children within the relevant age range, then the parent-focus counts were determined in the same way as the child-focus multiplicity counts: The count was of the self plus the spouse or unmarried partner. If the respondent also had older children (older than 25 and living with him or her) who had children of their own (identified as grandchildren) within the relevant age range, then the count was at least two. If the respondent had a spouse or unmarried partner in the household, then the count was incremented by one, and if a child-in-law was in the household, then the count also was incremented by one. (Note that, under these scenarios, the number of parents could range between two and four.)
4. If the respondent had grandchildren within the relevant age range but no children in that range, and the respondent had a child older than 25 or a child-in-law living with them, then the count was 2 if both the child older than 25 and the child-in-law were living in the household, and the count was 1 otherwise.

14.4.1.2 Sibling-Sibling Domains

When obtaining household-level person counts for sibling-sibling domains, only the two sibling-sibling domains previously listed were considered. As with the parent-child counts, the household-level person counts for sibling-sibling domains were nonzero if at least one sibling-sibling pair was present in the household within the relevant age ranges, in which the count was the number of appropriately aged siblings. If sets of siblings from more than one "family unit" (sets of siblings from different parents) resided within the same household, the sibling-sibling counts should have counted possible pairs from within each set. However, sets of siblings that did not involve the respondent's family unit could not have been identified from the data. Regardless of how many sets of siblings were in the household, counts had to be determined in different ways depending upon whether a sibling-sibling pair "of interest" was selected or not.

Descriptions of how to obtain the household-level person counts are provided below for the sibling-sibling domains outlined above. Sibling-sibling pairs of interest are described first. In this instance, the pair actually belonged to a pair relationship where analysis using one or more of the domains listed was possible. This is followed by descriptions for other pairs and single respondents. In each case, the descriptions apply regardless of which sibling-sibling domain was considered.

14.4.1.2.1 Obtaining Counts for Sibling-Sibling Domains: Sibling-Sibling Pairs of Interest

If the pair was identified as sibling-sibling within a relevant domain, the multiplicity count was the number of younger siblings because the older sibling was the focus. The household-level sibling-sibling person counts were determined in a similar manner to the multiplicity count, except that the count of interest was for the number of older siblings. If the pair member was the older sibling, then the household count was the self plus the number of siblings in the older age range. The count for the younger sibling pair member was the number of siblings within the same older age range. Unlike the case with the parent-child household-level counts, inconsistencies in the sibling-sibling counts when the pair selected was sibling-sibling still needed to be resolved. However, the rules for resolving inconsistencies followed directly from those used for the multiplicity counts when counting the number of younger siblings

(Appendix I). Note that a pair that was within one sibling-sibling pair domain had to be outside the other sibling-sibling pair domain.

14.4.1.2.2 Obtaining Counts for Sibling-Sibling Domains: Other Pairs and Single Respondents

For other pairs and single respondents, the following conditions were required to determine the household count for the number of siblings within the older age ranges of the domains of interest in the household:

1. If the age of the respondent was within the age range of the older sibling and that respondent had at least one sibling in the younger age range, then the count was the self plus the respondent's siblings within the older age range. If the respondent did not have any siblings within the younger age range, then the count was set to zero.
2. If the age of the respondent was within the age range of the younger sibling and that respondent had at least one sibling in the older age range, then the count was the number of the respondent's siblings in the older age range.
3. If the age of the respondent was outside the age range of the older or younger sibling but had at least one sibling in each of the older and younger age ranges, then the count was the number of siblings in the older age range.
4. If the age of the respondent was outside the age range of the older or younger sibling but the respondent had children within both the older and the younger age ranges, then the count was set to the number of respondent's children in the older age range.
5. If the age of the respondent was outside the age range of the older or younger sibling but the respondent had grandchildren within both the older and younger age ranges, then the count was the number of grandchildren in the older age range. If the respondent's grandchildren were cousins rather than siblings, then there was no way of deciphering this from the data. This had to be resolved by looking at the information from the other pair member, if another pair member was selected.

14.4.1.3 Spouse-Spouse Domains

What is referred to as a "spouse-spouse domain" was actually derived from spouse-spouse and partner-partner pair relationships. The following conditions were required for the number of spouse-spouse (including partner-partner) pairs to be incremented by 1. Some of these conditions were applied to the same household:

1. The respondent was part of a spouse-spouse (or partner-partner) pair.
2. The respondent was not part of a spouse-spouse pair but had a spouse (or unmarried partner).
3. The respondent had two parents living in the house.
4. The respondent had two parents-in-law living in the house.

5. The respondent had two grandparents living in the house.
6. The respondent had a child and a child-in-law living in the house.

The following conditions were required for the number of spouse-spouse pairs with children younger than 18 to be incremented by one. (These also include partner-partner pairs with children younger than 18.) Some of these conditions were applied to the same household:

1. The respondent was part of a spouse-spouse (or partner-partner) pair with children younger than 18.
2. The respondent was not part of a spouse-spouse pair¹⁹⁰ but had a spouse (or unmarried partner) and children younger than 18.
3. The respondent had two parents living in the house and was either younger than 18 or had siblings younger than 18.
4. The respondent had a child and a child-in-law living in the house and had grandchildren younger than 18.

14.4.2 Determining the Final Household-Level Person Count

For a particular type of household-level person count, there are three types of households from a sample selection perspective. For the first type, a pair was selected where the pair relationship corresponded directly to the pair domain being counted and both pair members responded. In this case, the household-level person count was usually easy to obtain using the multiplicity counts, although an adjustment was sometimes required if more than one family unit was in the household. For example, if a parent-child pair was selected where the child was 12 years old, the household-level person counts for the parent-focus parent-child (12-14) domain could usually be obtained from the multiplicity count that was calculated in stage two. In the second type of household, a pair also was selected and both pair members responded, but in this type the pair relationship did not correspond directly to the pair domain being counted. In this case, determining the final count was sometimes more difficult, particularly if one or more of the counts was a count of zero. A count of zero from a roster with good data did not necessarily mean that the final count should be zero. For example, suppose a household consisted of a man, his wife, brother, and two sons, and suppose one of the sons and his uncle (the man's brother) were selected. If the uncle's roster had a count of zero for all domains of interest—because he and his brother are not of the ages of interest and all of the other household members were "other relatives" from his perspective—then no nonzero parent-child count could be obtained. The final count would have to be determined from imputation. In the third type of household, only one respondent was selected. In this case, it was not possible to match counts from different pair members, but determining the final count could still be difficult if the count was zero for a household where the value was not truly zero.

For situations where a pair was selected and both pair members had good roster data, if the counts agreed between the pair members and were not zero, then a straightforward

¹⁹⁰ All spouse-spouse pairs were excluded here because spouse-spouse pairs with children were already accounted for, and spouse-spouse pairs without children had already been defined (possibly by imputation) not to have children younger than 18.

determination of the final household-level count was possible. This occurred in a majority of cases. If one pair member had a bad roster with no information in it and the other had a good roster, this was treated in the same way as if a single respondent was selected with a good roster. In either of these cases, the final count could be determined, provided a considerable number of conditions were satisfied. The conditions used to accept a good roster's count, when either the other pair member's roster was bad or no pair was selected, are provided in Appendix J. If these conditions were not met, the final household-level person count was left to imputation. Imputation also was required if two pair members were selected, both with bad rosters.

Among the remaining cases, some could be reconciled and some could not. In the cases where reconciliation was possible, some of the disagreements were caused by the pair members' rosters having different age and gender compositions. In these cases, many of the disagreements between the pair members were resolved by going to the screener. However, the screener did not provide much help if the age and gender composition of the pair members' rosters were identical, yet the counts still disagreed, as was the case with the uncle-nephew pair described above. In that example, one count was zero and the other was nonzero. Under conditions set out in Appendix J, it was possible to determine that the disagreement in this case was due to the uncle not being able to identify the parent-child domains, and the nonzero count was used. More detailed rules for reconciling differences between pair members are described in Appendix J.

If the attempt to reconcile differences in the household-level person counts between pair members was unsuccessful, upper and lower bounds within which the imputed value must reside were determined from the counts for each pair member and the counts for the screener.

14.5 Imputation for the Pair Variables

Imputation was required for variables in all three stages, and the imputation models for the pair variables required the inclusion of covariates at the pair level. The creation of these covariates is described in Section 14.5.1. The imputation process for stages one, two, and three are described in Sections 14.5.2 through 14.5.4, respectively. The final covariates for the fitted models are listed in Appendix D.

14.5.1 Creation of Covariates for Imputing Pair-Level Variables

Imputation was performed at the household level rather than at the respondent level. Thus, it was necessary to model covariates defined at the household level. Segment-level covariates were used for this purpose because they were automatically defined at the household level, using external information that was constant regardless of when the interviews were conducted. In addition to these segment-level covariates, information from the questionnaire would also have been useful as modeling variables. The logical choices for questionnaire-derived variables include the household composition variables IRHHSIZE (household size), IRKID17 (number in household younger than 18), IRHH65 (number in household aged 65 or older), and IRFAMSKP (presence of other family members in household indicator).

However, because interviews between pair members could have been conducted at different times, these variables were not necessarily consistent across pair members. Therefore, new count variables were needed that were consistent across the pair members (i.e., used

screeners information to reconcile disagreements between them) within a household. These variables were created in a two-step process: (1) create the count variables for each pair member, and (2) attempt to reconcile disagreeing values between pair members. The following sections describe these two steps in the creation of household size, household composition age count variables, and household composition age count variables "for males only," each of which were consistent across pair members. Note that household composition age count variables were not created for females because, for a given age range, the number of females could be obtained by subtracting the number of males from the total number within that age range. These variables also had to be created for respondents who were not part of a pair, for the purposes of creating and imputing the household-consistent person counts of various domains.

14.5.1.1 Household Size

The new variable created to represent a household size that was consistent across selected pair members was called HHSIZE and constructed as follows. First, the edited household size, TOTPEOP, was compared between pair members. If the values for TOTPEOP agreed across pair members and were both nonmissing and greater than 1,¹⁹¹ then HHSIZE was set to that value. If the values of TOTPEOP disagreed across pair members because the count for one pair member was missing and the count for the other was not missing and was greater than 1, then a natural choice for HHSIZE was the valid, nonmissing value contributed by one of the pair members. If the values of TOTPEOP either (1) disagreed across pair members without a clear indication of which one was valid or (2) were both missing or equal to 1, then the tools used to determine the final value of HHSIZE included the reported and edited household size variables, QD54 and TOTPEOP, as well as other measures of household size and "quality of roster" measures. These "other measures" included the screener household size and two sums of total valid ages within a pair member's roster.

The first sum was a simple total count of the number of roster members with valid ages, obtained by summing the counts within age groups. The second sum counted the ages of the pair members as reported during questionnaire administration. The two sums differed if a roster count (the first sum) was less than the number of pair members in a given age category. For example, if a household roster had one 12- to 17-year-old, but two 12- to 17-year-olds were selected, then the value of HHSIZE would be increased by 1. An additional situation occurred where the household size counts could not be easily determined by looking at both pair members. If the counts for both pair members were missing, then the screener household size was used to define HHSIZE.

In still other cases, disagreement between pair members with regard to the true household size could not be easily resolved. The screener household size did not support either household size in these cases, and the age counts mentioned above also did not resolve the disagreement. A decision had to be made as to which pair member's household size should be believed. This decision depended upon the "quality of the roster," where the household size was determined by the pair member with a "better" roster quality. One obvious way to measure roster quality was by noting the number of cases where the ages, relationship codes, or genders were missing in the

¹⁹¹ In households where a pair of respondents was selected, the household size had to be greater than or equal to two.

roster. If a roster was missing one or more of these three variables for some of the roster members, the roster was considered to be of "poorer quality" than a roster with these variables nonmissing for all roster members.

If only one household member was selected as a respondent, referred to as a "nonpair household," the rules for creating HHSIZE were the same as those that were used if two household members were selected in a pair but only one of the pair members had a nonmissing, acceptable value for a reported household size. Note that if only one household member was selected as a respondent, it was permissible to have a reported household size of one, whereas in a selected pair, a reported household size of one was considered "bad data."

In summary, the variables used to determine HHSIZE included (for each pair member) the reported and edited household sizes, the number of cases with valid ages in the roster, the number of cases with valid ages with the count in some age categories replaced by the minimum possible in that age category, and a roster quality count of the number of roster members with missing information. The screener household size, which was the same for each pair member, also was used. Using all of these tools, HHSIZE did not have any missing values in the 2016 survey, nor did it have any in surveys from previous years. General points about the creation of the household size variable are provided in Appendix K.

14.5.1.2 Household Composition Age Count Variables

It would seem logical to assert that the ages of other household members would be good covariates for the domain to which a pair might belong. Such variables also would be important for imputing multiplicity and household-level domain counts. The household-consistent age counts were limited to the following age ranges: younger than 12, 12 to 14, 15 to 17, 12 to 17, 12 to 20, 18 to 25, 26 to 34, 35 to 49, and 50 or older. These variables were called AGE011, AGE1214, AGE1517, AGE1217, AGE1220, AGE1825, AGE2634, AGE3549, and AGE50P, respectively.

The first step in this process was to count the nonmissing ages for roster members in the household for each pair member. In some cases, it was necessary to adjust the count because the ages could not be matched exactly. For example, suppose a 38-year-old and a 17-year-old were interviewed, and the 17-year-old was interviewed first. Suppose also that the 17-year-old turned 18 (i.e., had his or her 18th birthday) before the 38-year-old was interviewed. Hence, the 17-year-old would have had an age of 18 in the 38-year-old's roster. However, because the younger pair member was 17 at the time of his or her interview, the ages of interest for this pair domain were defined to be 17 and 38. Hence, it was necessary to account for this by creating a new roster age variable that matched the age provided in the other pair member's questionnaire. The age counts using this new roster age variable were equivalent to subtracting 1 from the previously obtained 18-25 count and adding 1 to the previously obtained 12-17 count in the 38-year-old's roster. These adjustments were made for all cases where a match was made between one pair member's roster and another pair member's interview age and gender and the ages did not match exactly.

If no roster ages were missing, the sum of these counts was equal to the edited household size TOTPEOP. Note that the reported household size was not considered here, because the counts were obtained from an edited roster. As with household size, a series of priority conditions was

used to obtain the most likely count within each age group. If the appropriate count was ambiguous because of disagreement between the pair members, the quality of the roster and the age of the respondent (in that order) were used to determine the appropriate count. The roster quality was determined by the number of bad or missing roster entries and the quality of the match between the pair member's roster and the other pair member's questionnaire age and gender.

If only one household member was selected as a respondent, the rules were the same as when two household members were selected in a pair but only one of the pair members had nonmissing data for the roster ages. One important exception to these rules was that when determining minimum possible counts for various age groups, it was not necessary to incorporate information from another pair member to increment the minimum for that pair member. General points about the creation of the age variables are provided in Appendix K.

14.5.1.3 Household Composition Age Counts of Males

For some pair variables, particularly spouse-spouse pairs, knowledge of the gender of the roster member was important for imputation purposes. In a similar manner to that used in the creation of the household composition age count variables, variables counting the number of males within the given age ranges were created. Disagreements between pair members were resolved in a similar manner to what was done with the household composition age count variables. The names of the male age counts were MALE011, MALE1214, MALE1517, MALE1217, MALE1220, MALE1825, MALE2634, MALE3549, and MALE50P.

14.5.2 Stage One Imputation: Pair Relationships

Missing pair relationships were filled in using the single response propensity (RP)/single prediction (PRD) type of the PMN method. In this stage, the imputation set involved only the edited variable PAIRREL, with its corresponding imputation-revised variable IRPRREL. Because the pair relationship varies according to the ages of the respondents, modeling and imputation were done independently within each of 11 age group pairs. [Table 14.8](#) presents these 11 age group pairs, as well as the pair relationships prevalent within each age group pair.

Table 14.8 Age Group Pairs with Associated Possible Pair Relationships

Age Group Pair Number	Age Group Pair	Pair Relationships Appearing in Age Group Pair (in Order of Prevalence) ¹	
		≥ 10% Prevalence ²	< 10% Prevalence
0	12-14/12-14	Sibling-sibling	Other relationship
1	12-14/15-17	Sibling-sibling	Other relationship
2	12-14/18-25	Sibling-sibling	Other relationship; parent-child; spouse-spouse**
3	15-17/15-17	Sibling-sibling	Other relationship; spouse-spouse*
4	15-17/18-25	Sibling-sibling	Other relationship; spouse-spouse*; parent-child*
5	18-20/18-25	Other relationship; sibling-sibling; spouse-spouse	Parent-child**

Table 14.8 Age Group Pairs with Associated Possible Pair Relationships (continued)

Age Group Pair Number	Age Group Pair	Pair Relationships Appearing in Age Group Pair (in Order of Prevalence) ¹	
		≥ 10% Prevalence ²	< 10% Prevalence
6	21-25/21-25	Spouse-spouse; other relationship; sibling-sibling	Parent-child**
7	12-14/26+	Parent-child	Other relationship; grandparent-grandchild; sibling-sibling*
8	15-17/26+	Parent-child	Other relationship; grandparent-grandchild; sibling-sibling; spouse-spouse**
9	18-20/26+	Parent-child	sibling-sibling; spouse-spouse; grandparent-grandchild; other relationship
10	21+/26+	Spouse-spouse; parent-child; other relationship	Sibling-sibling; grandparent-grandchild

* Pair relationships occur in less than 1 percent of the overall total number of pair relationships.

**The pair relationship is so rare that it does not appear in the age group pair in every survey year.

¹ The pair relationship labeled "spouse-spouse" includes partner-partner pair relationships. The spouse-spouse domain as listed here actually consists of two domains (spouse-spouse-with-children and spouse-spouse-without-children) that have been collapsed for the purposes of making the table easier to read. "Other relationship" refers to a relationship other than sibling-sibling, parent-child, grandparent-grandchild, or spouse-spouse.

² The pair relationships each form at least 10 percent of the overall total number of pair relationships within the given age group pair, and the total is at least 85 percent of the overall total.

14.5.2.1 Response Propensity Step

For a respondent pair to be considered complete, the pair relationship must be definitively established. In terms of the variable PAIRREL, this meant that the pair had to have a value of PAIRREL within the range of 1 to 9 or equal to 11 or 12. A value of PAIRREL equal to 13 also was considered complete, even though the pair relationship was not definitively established, because it was known that the pair relationship was not a relationship of interest. Response propensity adjustments then were calculated for each age group pair in order to make the respondent pair weights representative of the entire sample of pairs. Because the modeling of the final pair weight adjustments was not completed at the time of the pair imputations, the pair-level sample design weights were adjusted to account for nonresponse at the household level using a simple ratio adjustment. These adjustments were calculated using an item response propensity model, which is a special case of the generalized exponential model. See Appendix A of the questionnaire dwelling unit-level and person pair-level sampling weight calibration report in the 2016 NSDUH methodological resource book (Center for Behavioral Health Statistics and Quality, 2018b) for technical details about this procedure.

14.5.2.2 Prediction Step

After the weights were adjusted using the item response propensity model within each age group pair, logistic regression models were fitted using the adjusted weights. Preferred covariates in these models included the age count variables described in Section 14.5.1.2. However, variables with missing values for some observations cannot be used as covariates in

the models. To allow for better models whenever possible, the data were partitioned into two groups: those with nonmissing values for all age count variables, and those where one or more values of the age count variables were missing. When all of the age count variables were nonmissing, the predicted mean model was fitted using these variables. Otherwise, the model was fitted using the overall household count. This resulted in two predicted mean models for each of the 11 age group pairs.¹⁹² These groups were combined again before the hot-deck imputation step. Further details about the variables used in these models can be found in Appendix D.

All modeling was done using the SUDAAN procedure MULTLOG; however, the number of levels in the response variable varied based on the age group pair.¹⁹³ For age group pairs 0 through 4 and 7 through 9, dichotomous logistic regression models were built. Because there were three outcomes with age group pairs 5, 6, and 10, polytomous logistic models were fitted for these age group pairs. All the models incorporated the weights and were calibrated to account for item nonresponse (where a pair responded to the survey but the pair relationship was unknown), using the item response propensity models, as described in Section 14.5.2.1.

Ideally, each type of pair relationship within an age group pair would constitute a response category in a multinomial response model. However, the number of cases corresponding to some pair relationships within each age group pair was very small. Hence, it was not feasible to fit multinomial models that cover all the possible pair relationships for a given age group pair. Therefore, in the modeling step, some of the response categories were combined with separate assignments of imputed values within each of the 11 age group pairs. Priority was given to placing the pair relationships "of interest" into separate categories. In some cases, pair relationships that were not of interest were combined with other categories, even if there were sufficient numbers to have a separate category in the multinomial model. Table 14.9 presents the response categories that were used for modeling. The delineation between categories that were combined for modeling was determined in the hot-deck step.

As an example, consider age group pair 5. In this age group pair, there are typically four types of pair relationships that have a sufficient number of respondent pairs to fit an adequate model, including both spouse-spouse domains, sibling-sibling pairs, and all others. Models with fewer response levels are generally easier to fit because there are more observations in each response level. Because only two of those four were pair relationships of interest, the two spouse-spouse domains were used as levels in the response variable. The third level was obtained by combining the sibling-sibling and other relationship pairs. There are typically a small number of parent-child pairs, which also were combined with the other relationship pairs.

¹⁹² Partitioning the observations into two groups is similar to including an indicator variable in the PRD model. Though two models were fitted in practice, in theory this is similar to fitting only one PRD model. Thus, this is still considered the single RP/single PRD type of PMN.

¹⁹³ Though the SUDAAN procedure RLOGIST could have been used for all age group pairs except 5, 6, and 9, MULTLOG was used for all models for coding simplicity.

Table 14.9 Modeled Pair Relationships within Age Group Pairs

Age Group Pair Number	Age Group Pair	Number of Levels in Response	Levels of Modeled Response
0	12-14/12-14	2	Sibling-sibling; all others
1	12-14/15-17	2	Sibling-sibling; all others
2	12-14/18-25	2	Sibling-sibling; all others
3	15-17/15-17	2	Sibling-sibling; all others
4	15-17/18-25	2	Sibling-sibling; all others
5	18-20/18-25	3	Both spouse-spouse pair relationships; ¹ all others
6	21-25/21-25	3	Both spouse-spouse pair relationships; ¹ all others
7	12-14/26+	2	Parent-child; all others
8	15-17/26+	2	Parent-child; all others
9	18-20/26+	2	Parent-child; all others
10	21+/26+	3	Both spouse-spouse pair relationships; ¹ all others

¹ The two spouse-spouse pair relationships are spouse-spouse and spouse-spouse-with-children-younger-than-18. The pair relationships labeled "spouse-spouse" include partner-partner pair relationships.

14.5.2.3 Hot-Deck Step

Likeness constraints used in imputation of the pair relationship were generally based on the number of household members in various age groups and on the marital status and genders of the respondents. Logical constraints were limited to the information that was already known about the pair, as denoted by the level of the variable PAIRREL. If, for example, PAIRREL = 14, then no information was available about the identity of the pair relationship and no logical constraint was needed. On the other hand, if PAIRREL = 15, this meant that the pair relationship was either a parent-child pair where the child was aged 12 to 14 or it was some relationship other than spouse-spouse, parent-child, grandparent-grandchild, or sibling-sibling.

14.5.3 Stage Two Imputation: Multiplicity Counts

In many cases where the pair relationships were not defined, multiplicity counts also were not defined. In addition, there were a handful of cases where multiplicity counts were not determined, even when the pair relationship was known. In all of these cases, imputation was required to determine the multiplicity count, and the single RP/single PRD type of PMN was used for imputation.

The multiplicity count variables were divided into six imputation sets: four sets for the multiplicities associated with the four sibling-sibling pair domains; one set for the parent-child, child focus domains; and one set for the parent-child, parent focus domains. The variables in each imputation set are provided in [Table 14.10](#). Because the parent-child counts are hierarchical (i.e., the count for 12-17 must be less than or equal to the count for 12-20), only the counts for the 12-20 age group were modeled. Using the predicted means from these models, a single donor pair for each focus was selected from which the multiplicity counts were determined for the 12-14, 12-17, 15-17, and 12-20 parent-child pairs. No imputation was required for the spouse-

spouse multiplicity counts, because a selected respondent in a spouse-spouse pair was assumed to have had only one spouse.

Table 14.10 Modeled Multiplicities within Domains

Imputation Set	Domain	Model Type	Base Variables	Imputation-Revised Variables
1	Parent-Child, Child Focus	Logistic	MCPCC14, MCPCC57, MCPCC17, MCPCC20	IRMPCC14, IRMPCC57, IRMPCC17, IRMPCC20
2	Parent-Child, Parent Focus	Poisson	MCPCP14, MCPCP57, MCPCP17, MCPCP20	IRMPCP14, IRMPCP57, IRMPCP17, IRMPCP20
3	Sibling-Sibling (12-14/15-17) Older Focus	Poisson	MCS1417	IRMS1417
4	Sibling-Sibling (12-14/15-17) Younger Focus	Poisson	MCS1714	IRMS1714
5	Sibling-Sibling (12-17/18-25) Older Focus	Poisson	MCS1725	IRMS1725
6	Sibling-Sibling (12-17/18-25) Younger Focus	Poisson	MCS2517	IRMS2517

14.5.3.1 Imputation for Parent-Child Multiplicity, Child Focus (Imputation Set 1)

The first imputation set included the four parent-child, child focus multiplicity variables. Modeling was done only for the parent-child (12-20) variable, and multivariate assignment of all four variables was done in the hot-deck step.

14.5.3.1.1 Response Propensity Step

For a respondent pair to be considered complete with regard to the parent-child multiplicities, the multiplicity had to be nonmissing for the domains with children aged 12 to 20. A nonmissing multiplicity for this domain would automatically guarantee nonmissing multiplicities for the subset parent-child domains. Response propensity adjustments were then calculated in order to make the respondent pair weights representative of the entire sample of pairs.

14.5.3.1.2 Prediction Step

For the child-focus parent-child domains, the count being modeled was the number of parents of children aged 12 to 20 who were part of a parent-child pair. In most cases, only two responses were possible: one parent or two parents. There were rare instances where the child could have three parents living in the household, perhaps due to some combination of biological, step, foster, or adoptive parents. For the purposes of modeling, these cases were collapsed with the two-parent households. Similar to the procedure described in Section 14.5.2.2, the data were divided into two groups based on whether the age count variables were nonmissing, and separate models were fitted for each group. The fitted models were binomial logistic regression models using the SUDAAN procedure RLOGIST.

14.5.3.1.3 Hot-Deck Step

Though modeling was only done for the 12-20 age group, the multiplicity counts for the 12-14, 12-17, 15-17, and 12-20 parent-child pairs were assigned from a single donor pair. Likeness constraints were generally based on the pair relationship, the household size, and the number of household members in various age groups. Logical constraints were based on the bounds created during the editing process. These constraints are described in detail in [Tables E.155](#) and [E.156](#) in Appendix E.

14.5.3.2 Imputation for All Other Multiplicities (Imputation Sets 2 through 6)

The imputation process for the other responses (parent-focus parent-child and sibling-sibling multiplicity counts) was similar to that for the child-focus parent-child counts. The main difference is in the PRD step where Poisson regression models were used to model the counts. The counts of the number of children or siblings were underdispersed for a Poisson distribution so that the data had to be scaled using the observed variance.

As with the child-focus parent-child multiplicity counts, the parent-focus parent-child multiplicity counts were only modeled for the 12-20 age group, and counts from all four age groups were assigned from a single donor pair. For the sibling-sibling multiplicity counts, there was only one variable in each imputation set, and each count was modeled separately.

14.5.4 Stage Three Imputation: Household-Level Person Counts

Because of the difficulty in definitively determining household-level counts in many cases, imputation was not uncommon. Household-level person counts were divided into five imputation sets based on the domains listed in Section 14.4: one for each of the two sibling-sibling counts, one for the spouse-spouse counts, one for the spouse-spouse with children counts, and one for the parent-child counts. The first four imputation sets were handled using the single RP/single PRD type of PMN, whereas the imputation of the parent-child household counts was done using the multiple RP/multiple PRD type of PMN. For these counts, separate models were fit for the child-focus and parent-focus counts, and the predicted means from both models were brought together in one hot-deck step.

As with the multiplicities, the parent-child domains were hierarchical, so the imputations could not have been conducted independently if consistency was to be maintained. Hence, models were fitted only to the parent-child domains for the 12-20 age group, and the household-level person counts were assigned for the 12-14, 12-17, and 12-20 parent-child pair domains from a single donor. The household-level person counts for the 15-17 parent-child domains were not determined as they can be easily derived. The spouse-spouse household-level person counts were also hierarchical in that knowledge of whether a spouse-spouse pair was in the household was required before one could say that the pair had children. Therefore, imputations of the spouse-spouse counts were processed first, followed by the imputations of whether the spouse-spouse pairs in the household had children.

Household-level person counts were defined for all respondents, regardless of which pair they belonged to, or even whether they were within a pair at all. For modeling purposes,

respondents were partitioned into two groups based on whether they belonged to a pair, and the entire imputation process was conducted separately for each group.

14.5.4.1 Imputation for Sibling-Sibling (12-14/15-17), Older Focus Household Counts (Imputation Set 1)

The first imputation set included one base variable, HCS1417, and the corresponding imputation-revised variable IRHS1417. The imputation process for this variable is detailed in Sections 14.5.4.1.1 through 14.5.4.1.3.

14.5.4.1.1 Response Propensity Step

For a pair or single respondent to be considered complete, the household-level person counts had to be nonmissing for all the variables being imputed. In addition to being separated into pair and nonpair households, respondents also were split by age for both the modeling of the response propensity and the predicted means. For the pairs, households where both pair members were younger than 18 were placed in one group, and the remaining pairs were in the other age group. For the single respondents, one age group consisted of respondents who were younger than 18, and the other consisted of respondents who were 18 or older. For pairs, response propensity adjustments were calculated to make the household weights representative of the entire sample of pairs. For single respondents, household weights also were used, and the adjustments were calculated to make the respondent household weights representative of the entire sample of households that were not part of a pair.

14.5.4.1.2 Prediction Step

After the response propensity adjustment was complete, the data were split into two groups based on whether the household-level age count variables were nonmissing, as previously described in Section 14.5.2.2. The outcome variable for the household-level sibling-sibling (12-14/15-17) count models was the number of household members aged 15 to 17 with a sibling aged 12 to 14 living with them. These counts could have a value of zero, which distinguished them from the multiplicities from a modeling point of view. Poisson regression was used to fit the models for the household-level person counts corresponding to the sibling-sibling domains. The data were underdispersed for a Poisson distribution so that the data had to be scaled using the observed variance. Modeling was done using the SUDAAN procedure LOGLINK.

14.5.4.1.3 Hot-Deck Step

After the modeling steps were complete, the two age groups were combined for one hot-deck imputation step. Imputation was conducted separately for pair and nonpair households. Likeness constraints used in imputation of household counts were generally based on the number of families in the household, the household size, and the number of household members in various age groups. Logical constraints were based on the bounds created during the editing process. These constraints are described in detail in [Tables E.164](#) and [E.165](#) in Appendix E.

In those instances where an imputed value could not be found after loosening all the likeness constraints, the imputed value was determined by doing a random imputation within bounds derived from the household composition.

14.5.4.2 Imputation for Sibling-Sibling (12-17/18-25) and Both Spouse-Spouse Household Counts (Imputation Sets 2 through 4)

The second, third, and fourth imputation sets were processed similarly to the first. One large difference was that when modeling the spouse-spouse-with-children counts, the data were not separated according to age groups. This applied to both the response propensity adjustments and the calculation of predicted means. Because of the hierarchical relationship between the spouse-spouse and the spouse-spouse-with-children counts, the response propensity adjustment for the spouse-spouse-with-children domain adjusted the weights to be representative of all spouse-spouse pairs rather than the entire sample. Missing counts for the spouse-spouse-with-children domain were not replaced via imputation until it was known definitively, after the hot-deck step of the PMN imputation, whether a household had spouse-spouse pairs.

Polytomous logistic regression was used to model the count of spouse-spouse pairs, with the possible responses being zero, one, and two or more spouse-spouse pairs in the household. Whether or not the spouse-spouse pairs had children younger than 18 was modeled with binomial logistic regression. These models were fitted using the SUDAAN MULTILOG procedure.

In some cases, two family units were in a household. If these resulted in unusual household-level person counts, they were excluded from the modeling step and were considered nonrespondents for the purposes of weight adjustment. No predicted mean was calculated in these cases, and instead of matching donors and recipients using predicted means, the imputed value was determined using random imputation within the preset bounds. One case where this may have occurred was with the spouse-spouse-with-children counts. Having two spouse-spouse pairs with children younger than 18 was an extremely rare category. Therefore, the two response categories that resulted for the spouse-spouse-with-children models were zero or one or more. Households with two family units did not need to be excluded from the spouse-spouse models, because having two spouse-spouse pairs in a household, though not common, was not rare.

14.5.4.3 Imputation for Parent-Child Household Counts (Imputation Set 5)

In contrast to the first four imputation sets, the parent-child household counts were imputed using the multiple RP/multiple PRD type of PMN. This imputation set included the three child-focus counts, HCPCC14, HCPCC17, and HCPCC20, as well as the three parent-focus counts, HCPCP14, HCPCP17, and HCPCP20. The corresponding imputation-revised variables are IRHPCC14, IRHPCC17, IRHPCC20, IRHPCP14, IRHPCP17, and IRHPCP20, respectively. The child-focus and parent-focus counts were modeled separately and joined together in one final hot-deck step. Just as was done for the first four imputation sets, the respondents were grouped into pair and nonpair households, and the entire imputation process was completed separately for each group. Unlike most imputation sets processed using the multiple RP/multiple PRD type of PMN, no provisional hot-deck steps were implemented. Because the child-focus counts were not used as covariates for the parent-focus counts and the parent-focus counts were not used as covariates for the child-focus counts, no provisional hot-deck steps were necessary, and the RP and PRD steps for the child-focus counts were run in parallel with the RP and PRD steps for the parent-focus counts.

14.5.4.3.1 First Response Propensity Step: Child Focus

All respondents were in the domain for all household counts. For a pair or single respondent to be considered complete, the three child-focus counts had to be nonmissing. Similar to the processing of the sibling-sibling counts and the spouse-spouse counts, the data were divided into two age groups based on whether the respondent or pair of respondents was 18 or older.

14.5.4.3.2 First Prediction Step: Child Focus

For the child-focus counts, the modeled response was the number of children aged 12 to 20 in the household with at least one parent living with them. This was modeled using Poisson regression, where the data were scaled using the observed variance to account for underdispersion. The LOGLINK procedure in SUDAAN was used to fit the model. The predicted mean used in the final hot-deck step was the predicted number of children aged 12 to 20 in the household with at least one parent living with them.

14.5.4.3.3 Second Response Propensity Step: Parent Focus

This RP step was identical to the RP step for the child-focus counts, except item respondents were those whose three parent-focus counts were nonmissing.

14.5.4.3.4 Second Prediction Step: Parent Focus

For the parent-focus counts, the modeled response was a three-level variable, based on the number of parents in the household with children aged 12 to 20: zero, one, or two or more. Polytomous logistic regression was used to fit the model as implemented by the MULTLOG procedure in SUDAAN. The predicted means used in the final hot-deck step were the predicted probabilities associated with each of the three levels of the response variable.

14.5.4.3.5 Hot-Deck Step

The predicted means from the models fit for the child-focus and parent-focus counts were brought together in one hot-deck imputation step. Because there are six variables in the imputation set, and any combination of them could be missing, there were 63 possible missingness patterns. These patterns are enumerated in [Table E.170](#) in Appendix E. As with the other four imputation sets, logical constraints were based on the bounds created during the editing process, and likeness constraints were generally based on the number of families in the household, the household size, and the number of household members in various age groups. These constraints are described in detail in [Tables E.164](#) and [E.165](#).

In cases where there were two family units in the household, resulting in unusual counts, the counts were not included in the parent-focus models, and no predicted means were calculated. Even though two-family households were included in the model for the child-focus parent-child counts, the resulting predicted means were not used. This was because the parent-focus and child-focus parent-child counts were in the same imputation set, and the predicted means could not be used in the imputation of the parent-focus parent-child counts when two families were in the household. In these cases, imputation was random between the bounds.

This page intentionally left blank

References

- Allison, P. D. (2008). Missing data. In R. E. Millsap & A. Maydeu-Olivares (Eds.), *The SAGE handbook of quantitative methods in psychology*. Thousand Oaks, CA: Sage Publications, Inc.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (DSM-IV) (4th ed.). Washington, DC: Author.
- Ault, K., Aldworth, J., Barnett-Walker, K., Carpenter, L., Copello, E., Frechtel, P., Liu, B., & Martin, P. (2009). Imputation report. In *2007 National Survey on Drug Use and Health: Methodological resource book* (Section 11, prepared for the Substance Abuse and Mental Health Services Administration, Contract No. 283-2004-00022, Deliverable No. 39, RTI/0209009.377.007). Research Triangle Park, NC: RTI International. Retrieved from <https://www.samhsa.gov/data/>
- Ault, K., Barnett-Walker, K., Carpenter, L., Copello, E., Cummiskey, C., Frechtel, P., Laufenberg, J., Liu, B., Martin, P., & Moore, A. (2010). Imputation report. In *2008 National Survey on Drug Use and Health: Methodological resource book* (Section 11, prepared for the Substance Abuse and Mental Health Services Administration, Contract No. 283-2004-00022, Deliverable No. 39, RTI/0209009.477.007). Research Triangle Park, NC: RTI International. Retrieved from <https://www.samhsa.gov/data/>
- Ault, K., Barnett-Walker, K., Carpenter, L., Cummiskey, C., Frechtel, P., Laufenberg, J., Martin, P., Moore, A., & Scott, V. (2011). Imputation report. In *2009 National Survey on Drug Use and Health: Methodological resource book* (Section 11, prepared for the Substance Abuse and Mental Health Services Administration, Contract No. 283-2004-00022, Deliverable No. 39, RTI/0209009.577.007). Research Triangle Park, NC: RTI International. Retrieved from <https://www.samhsa.gov/data/>
- Center for Behavioral Health Statistics and Quality. (2014a). *National Survey on Drug Use and Health: 2012 Questionnaire field test final report*. Rockville, MD: Substance Abuse and Mental Health Services Administration.
- Center for Behavioral Health Statistics and Quality. (2014b). *National Survey on Drug Use and Health: 2013 Dress rehearsal final report*. Rockville, MD: Substance Abuse and Mental Health Services Administration.
- Center for Behavioral Health Statistics and Quality. (2015). *2013 National Survey on Drug Use and Health: Methodological resource book (Section 10, Editing and imputation report)*. Retrieved from <https://www.samhsa.gov/data/>
- Center for Behavioral Health Statistics and Quality. (2016a). *2014 National Survey on Drug Use and Health: Methodological resource book (Section 10, Editing and imputation report)*. Retrieved from <https://www.samhsa.gov/data/>

Center for Behavioral Health Statistics and Quality. (2016b). *Key substance use and mental health indicators in the United States: Results from the 2015 National Survey on Drug Use and Health* (HHS Publication No. SMA 16-4984, NSDUH Series H-51). Retrieved from <https://www.samhsa.gov/data/>

Center for Behavioral Health Statistics and Quality. (2017a). *2015 National Survey on Drug Use and Health: Methodological resource book (Section 10, Editing and imputation report)*. Retrieved from <https://www.samhsa.gov/data/>

Center for Behavioral Health Statistics and Quality. (2017b). *2016 National Survey on Drug Use and Health final analytic file codebook*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (2017c). *2016 National Survey on Drug Use and Health: Methodological resource book (Section 2, Sample design report)*. Retrieved from <https://www.samhsa.gov/data/>

Center for Behavioral Health Statistics and Quality. (2017d). *2016 National Survey on Drug Use and Health: Methodological summary and definitions*. Retrieved from <http://www.samhsa.gov/data/>

Center for Behavioral Health Statistics and Quality. (2017e). *Evaluation of imputation methods for the National Survey on Drug Use and Health*. Retrieved from <https://www.samhsa.gov/data/>

Center for Behavioral Health Statistics and Quality. (2017f). *Results from the 2016 National Survey on Drug Use and Health: Detailed tables*. Retrieved from <https://www.samhsa.gov/data/>

Center for Behavioral Health Statistics and Quality. (2018a). *2016 National Survey on Drug Use and Health: Methodological resource book (Section 11, Person-level sampling weight calibration)*. Retrieved from <https://www.samhsa.gov/data/>

Center for Behavioral Health Statistics and Quality. (2018b). *2016 National Survey on Drug Use and Health: Methodological resource book (Section 12, Questionnaire dwelling unit-level and person pair-level sampling weight calibration)*. Retrieved from <https://www.samhsa.gov/data/>

Chromy, J. R. (1979). Sequential sample selection methods. In *Proceedings of the 1979 American Statistical Association, Survey Research Methods Section, Washington, DC* (pp. 401-406). Washington, DC: American Statistical Association. [Available as a PDF at <http://www.amstat.org/ASA/Membership/Sections-and-Interest-Groups.aspx>]

Chromy, J. R., & Singh, A. C. (2001). Estimation for person-pair drug-related characteristics in the presence of pair multiplicities and extreme sampling weights. In *Proceedings of the 2001 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Atlanta, GA* [CD-ROM]. Alexandria, VA: American Statistical Association. [Available as a PDF at <http://www.amstat.org/ASA/Membership/Sections-and-Interest-Groups.aspx>]

Cox, B. G. (1980). The weighted sequential hot deck imputation procedure. In *Proceedings of the 1980 American Statistical Association, Survey Research Methods Section, Houston, TX* (pp. 721-726). Washington, DC: American Statistical Association. [Available as a PDF at <http://www.amstat.org/ASA/Membership/Sections-and-Interest-Groups.aspx>]

Cox, B. G., & Cohen, S. B. (1985). *Methodological issues for health care surveys*. New York: Marcel Dekker, Inc.

Cox, D. R., & Snell, E. J. (1989). *The analysis of binary data* (2nd ed.). (Monographs on Statistics and Applied Probability Series). Boca Raton, FL: CRC Press.

Iannacchione, V. (1982). Weighted sequential hot deck imputation macros. In *Proceedings of the Seventh Annual SAS Users Group International Conference* (pp. 759-763). Cary, NC: SAS Corporation.

Kroutil, L. A., Chien, C., Handley, W., & Bradshaw, M. R. (2013). Procedures for editing interviewer-administered data in the 2011 NSDUH computer-assisted interview. In *2011 National Survey on Drug Use and Health: Methodological resource book* (Section 10, prepared for the Substance Abuse and Mental Health Services Administration, Contract No. HHSS283200800004C, Deliverable No. 39, RTI/0211838.207.003). Research Triangle Park, NC: RTI International. Retrieved from <https://www.samhsa.gov/data/>

Kroutil, L. A., Handley, W., & Bradshaw, M. R. (2013). General principles and procedures for editing drug use data in the 2011 NSDUH computer-assisted interview. In *2011 National Survey on Drug Use and Health: Methodological resource book* (Section 10, prepared for the Substance Abuse and Mental Health Services Administration, Contract No. HHSS283200800004C, Deliverable No. 39, RTI/0211838.207.003). Research Triangle Park, NC: RTI International. Retrieved from <https://www.samhsa.gov/data/>

Kroutil, L., & Myers, L. (2002). Development of editing rules for CAI substance use data. In J. Gfroerer, J. Eyerman, & J. Chromy (Eds.), *Redesigning an ongoing national household survey: Methodological issues* (HHS Publication No. SMA 03-3768, pp. 85-109). Rockville, MD: Substance Abuse and Mental Health Services Administration.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: John Wiley & Sons.

Office of Applied Studies. (2009). *Results from the 2008 National Survey on Drug Use and Health: National findings* (HHS Publication No. SMA 09-4434, NSDUH Series H-36). Retrieved from <https://www.samhsa.gov/data/>

Office of Applied Studies. (2010a). *Results from the 2009 National Survey on Drug Use and Health: Volume I. Summary of national findings* (HHS Publication No. SMA 10-4586Findings, NSDUH Series H-38A). Retrieved from <https://www.samhsa.gov/data/>

Office of Applied Studies. (2010b). *Results from the 2009 National Survey on Drug Use and Health: Volume II. Technical appendices and selected prevalence tables* (HHS Publication No. SMA 10-4586Appendices, NSDUH Series H-38B). Retrieved from <https://www.samhsa.gov/data/>

Office of Management and Budget. (1997). Revisions to the standards for the classification of federal data on race and ethnicity. *Federal Register*, 62(210), 58781-58790.

Office of Management and Budget. (2009). *Update of statistical area definitions and guidance on their uses* (OMB Bulletin No. 10-02). Washington, DC: The White House. Retrieved from <https://www.whitehouse.gov/omb/bulletins/#2009-2016>

RTI International. (2013). *SUDAAN® language manual, Release 11.0.1*. Research Triangle Park, NC: RTI International.

Rubin, D. B. (1986). Statistical matching using file concatenation with adjusted weights and multiple imputations. *Journal of Business and Economic Statistics*, 4(1), 87-94.

SAS Institute Inc. (2014). *SAS/STAT® 13.2 user's guide: High-performance procedures*. Cary, NC: SAS Institute Inc.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data* (No. 72, Monographs on Statistics and Applied Probability). Boca Raton, FL: Chapman and Hall/CRC.

Shiffman, S., Hickcox, M., Gnys, M., Paty, J. A., & Kassel, J. D. (1995, March). *The Nicotine Dependence Syndrome Scale: Development of a new measure*. Poster presented at the annual meeting of the Society for Research on Nicotine and Tobacco, San Diego, CA.

Shiffman, S., Waters, A. J., & Hickcox, M. (2003). The Nicotine Dependence Syndrome Scale: A multi-dimensional measure of nicotine dependence. Unpublished manuscript.

Singh, A., Grau, E., & Folsom, R. (2004). Imputation and unbiased estimation: Use of centered predictive mean neighborhoods method. In *Proceedings of the 2004 Joint Statistical Meetings, American Statistical Association, Section on Survey Research Methods, Toronto, Ontario, Canada* (pp. 4351-4358). Alexandria, VA: American Statistical Association. [Available as a PDF at <http://www.amstat.org/ASA/Membership/Sections-and-Interest-Groups.aspx>]

Appendix A: Percentage of Cases Imputed or Logically Assigned, 2016 Survey

This page intentionally left blank

Table A.1 Percentage of Cases Imputed for General Cigarette Use Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Cigarette Recency	Lifetime Cigarette Users	33,150	310	1	311	0.38%	0.00%	0.38%
Cigarette 30-Day Frequency	Past Month Cigarette Users	12,693	32	83	115	0.19%	0.71%	0.89%
Cigarette Age at First Use	Lifetime Cigarette Users	33,150	426	18	444	0.96%	0.05%	1.01%
Cigarette Day of First Use	Lifetime Cigarette Users	33,150	33,150	0	33,150	100.00%	0.00%	100.00%
Cigarette Month of First Use	Lifetime Cigarette Users	33,150	31,793	15	31,808	98.35%	0.01%	98.36%
Cigarette Year of First Use	Lifetime Cigarette Users	33,150	31,704	3	31,707	98.25%	0.00%	98.25%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.2 Percentage of Cases Imputed for Daily Cigarette Use Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Cigarette Daily Use	Lifetime Cigarette Users	33,150	16	30	46	0.04%	0.13%	0.17%
Cigarette Age at First Daily Use	Daily Cigarette Users	17,018	123	2	125	0.78%	0.00%	0.78%
Cigarette Day of First Daily Use	Daily Cigarette Users	17,018	17,018	0	17,018	100.00%	0.00%	100.00%
Cigarette Month of First Daily Use	Daily Cigarette Users	17,018	16,573	2	16,575	98.97%	0.01%	98.98%
Cigarette Year of First Daily Use	Daily Cigarette Users	17,018	16,545	1	16,546	98.91%	0.00%	98.91%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.3 Percentage of Cases Imputed for Cigarette Dependence Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Smoke Cigarettes Regularly throughout the Day	Past Month Cigarette Users	12,693	0	0	0	0.00%	0.00%	0.00%
Smoke Same Number of Cigarettes from Day to Day	Past Month Cigarette Users	12,693	4	0	4	0.02%	0.00%	0.02%
Number Cigarettes Smoked per Day Often Changes	Past Month Cigarette Users	12,693	10	0	10	0.07%	0.00%	0.07%
Number Cigarettes Smoked per Day Influenced by Other Things	Past Month Cigarette Users	12,693	2	0	2	0.01%	0.00%	0.01%
Smoke Same Number of Cigarettes on Weekends As on Weekdays	Past Month Cigarette Users	12,693	12	0	12	0.05%	0.00%	0.05%
Smoke Lots of Cigarettes in an Hour, Then No Cigarettes Until Hours Later	Past Month Cigarette Users	12,693	5	0	5	0.04%	0.00%	0.04%
Amount of Smoking Has Increased Since Started Smoking	Past Month Cigarette Users	12,693	8	0	8	0.04%	0.00%	0.04%
Must Smoke Much More Now Before Start To Feel Anything	Past Month Cigarette Users	12,693	26	0	26	0.29%	0.00%	0.29%
Need To Smoke a Lot More To Be Satisfied	Past Month Cigarette Users	12,693	4	0	4	0.05%	0.00%	0.05%
Crave Cigarettes when Haven't Smoked for a Few Hours	Past Month Cigarette Users	12,693	0	0	0	0.00%	0.00%	0.00%
Need To Smoke To Feel Less Irritable	Past Month Cigarette Users	12,693	11	0	11	0.11%	0.00%	0.11%
Feel in Control of Smoking	Past Month Cigarette Users	12,693	3	0	3	0.01%	0.00%	0.01%
Cravings for Cigarettes like Force Can't Control	Past Month Cigarette Users	12,693	3	0	3	0.02%	0.00%	0.02%

Table A.3 Percentage of Cases Imputed for Cigarette Dependence Variables (continued)

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Worry about Running Out of Cigarettes	Past Month Cigarette Users	12,693	0	0	0	0.00%	0.00%	0.00%
Smoking Not Affected by Other Things	Past Month Cigarette Users	12,693	12	0	12	0.11%	0.00%	0.11%
Tend To Avoid Places That Don't Allow Smoking	Past Month Cigarette Users	12,693	13	0	13	0.11%	0.00%	0.11%
No Travel by Airplane Because No Smoking Allowed	Past Month Cigarette Users	12,693	12	0	12	0.14%	0.00%	0.14%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.4 Percentage of Cases Imputed for Cigar Use Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Cigar Lifetime Use	All Respondents	67,942	13	0	13	0.02%	0.00%	0.02%
Cigar Recency	Lifetime Cigar Users	19,559	239	1	240	0.61%	0.00%	0.61%
Cigar 30-Day Frequency	Past Month Cigar Users	3,464	13	26	39	0.26%	0.83%	1.09%
Cigar Age at First Use	Lifetime Cigar Users	19,559	387	0	387	1.96%	0.00%	1.96%
Cigar Day of First Use	Lifetime Cigar Users	19,559	19,559	0	19,559	100.00%	0.00%	100.00%
Cigar Month of First Use	Lifetime Cigar Users	19,559	18,039	3	18,042	96.17%	0.00%	96.18%
Cigar Year of First Use	Lifetime Cigar Users	19,559	17,954	0	17,954	96.00%	0.00%	96.00%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.5 Percentage of Cases Imputed for Smokeless Tobacco Use Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Smokeless Tobacco Lifetime Use	All Respondents	67,942	53	0	53	0.05%	0.00%	0.05%
Smokeless Tobacco Recency	Lifetime Smokeless Tobacco Users	11,033	186	3	138	1.06%	0.03%	1.09%
Smokeless Tobacco 30-Day Frequency	Past Month Smokeless Tobacco Users	2,592	11	13	24	0.39%	1.06%	1.45%
Smokeless Tobacco Age at First Use	Lifetime Smokeless Tobacco Users	11,033	216	0	216	1.79%	0.00%	1.79%
Smokeless Tobacco Day of First Use	Lifetime Smokeless Tobacco Users	11,033	11,033	0	11,033	100.00%	0.00%	100.00%
Smokeless Tobacco Month of First Use	Lifetime Smokeless Tobacco Users	11,033	10,266	3	10,269	96.34%	0.01%	96.36%
Smokeless Tobacco Year of First Use	Lifetime Smokeless Tobacco Users	11,033	10,214	0	10,214	96.08%	0.00%	96.08%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.6 Percentage of Cases Imputed for Pipe Use Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Pipe Lifetime Use	All Respondents	67,942	15	0	15	0.02%	0.00%	0.02%
Pipe Past Month Use	Lifetime Pipe Users	6,065	1	0	1	0.02%	0.00%	0.02%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.7 Percentage of Cases Imputed for Alcohol Use Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Alcohol Lifetime Use	All Respondents	67,942	19	0	19	0.03%	0.00%	0.03%
Alcohol Recency	Lifetime Alcohol Users	48,526	783	12	795	1.17%	0.01%	1.18%
Alcohol 12-Month Frequency	Past Year Alcohol Users	40,875	783	1,182	1,965	1.56%	2.03%	3.59%
Alcohol 30-Day Frequency	Past Month Alcohol Users	30,917	639	153	792	1.72%	0.53%	2.25%
Alcohol 4+/5+ Drinks Past Month	Past Month Alcohol Users	30,917	948	70	1,018	2.72%	0.24%	2.97%
Alcohol Age at First Use	Lifetime Alcohol Users	48,526	703	0	703	1.42%	0.00%	1.42%
Alcohol Day of First Use	Lifetime Alcohol Users	48,526	48,526	0	48,526	100.00%	0.00%	100.00%
Alcohol Month of First Use	Lifetime Alcohol Users	48,526	44,995	12	45,007	97.05%	0.01%	97.06%
Alcohol Year of First Use	Lifetime Alcohol Users	48,526	44,774	0	44,774	96.88%	0.00%	96.88%

Note: Beginning in 2015, binge alcohol use was defined as consuming five or more drinks for males and four or more drinks for females on the same occasion. Prior to 2015, binge alcohol use was defined as consuming five or more drinks on the same occasion for both males and females.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.8 Percentage of Cases Imputed for Marijuana Use Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Marijuana Lifetime Use	All Respondents	67,942	44	0	44	0.07%	0.00%	0.07%
Marijuana Recency	Lifetime Marijuana Users	28,550	391	1	392	0.86%	0.00%	0.86%
Marijuana 12-Month Frequency	Past Year Marijuana Users	12,355	379	794	1,173	2.48%	4.65%	7.13%
Marijuana 30-Day Frequency	Past Month Marijuana Users	7,738	189	55	244	1.95%	0.52%	2.46%
Marijuana Age at First Use	Lifetime Marijuana Users	28,550	234	0	234	0.69%	0.00%	0.69%
Marijuana Day of First Use	Lifetime Marijuana Users	28,550	28,550	0	28,550	100.00%	0.00%	100.00%
Marijuana Month of First Use	Lifetime Marijuana Users	28,550	26,624	7	26,631	96.92%	0.01%	96.93%
Marijuana Year of First Use	Lifetime Marijuana Users	28,550	26,491	0	26,491	96.71%	0.00%	96.71%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.9 Percentage of Cases Imputed for Cocaine Use Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Cocaine Lifetime Use	All Respondents	67,942	22	0	22	0.03%	0.00%	0.03%
Cocaine Recency	Lifetime Cocaine Users	7,870	103	31	134	0.79%	0.32%	1.11%
Cocaine 12-Month Frequency	Past Year Cocaine Users	1,575	73	120	193	4.67%	7.83%	12.50%
Cocaine 30-Day Frequency	Past Month Cocaine Users	518	27	18	45	3.98%	8.79%	12.78%
Cocaine Age at First Use	Lifetime Cocaine Users	7,870	63	103	166	0.85%	1.34%	2.20%
Cocaine Day of First Use	Lifetime Cocaine Users	7,870	7,870	0	7,870	100.00%	0.00%	100.00%
Cocaine Month of First Use	Lifetime Cocaine Users	7,870	7,287	1	7,288	96.33%	0.01%	96.34%
Cocaine Year of First Use	Lifetime Cocaine Users	7,870	7,251	0	7,251	96.14%	0.00%	96.14%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.10 Percentage of Cases Imputed for Crack Use Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Crack Lifetime Use	All Respondents	67,942	26	0	26	0.04%	0.00%	0.04%
Crack Recency	Lifetime Crack Users	1,738	11	0	11	0.47%	0.00%	0.47%
Crack 12-Month Frequency	Past Year Crack Users	190	8	12	20	3.25%	4.05%	7.29%
Crack 30-Day Frequency	Past Month Crack Users	76	3	0	3	1.39%	0.00%	1.39%
Crack Age at First Use	Lifetime Crack Users	1,738	15	0	15	0.70%	0.00%	0.70%
Crack Day of First Use	Lifetime Crack Users	1,738	1,738	0	1,738	100.00%	0.00%	100.00%
Crack Month of First Use	Lifetime Crack Users	1,738	1,691	0	1,691	98.69%	0.00%	98.69%
Crack Year of First Use	Lifetime Crack Users	1,738	1,689	0	1,689	98.64%	0.00%	98.64%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.11 Percentage of Cases Imputed for Heroin Use Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Heroin Lifetime Use	All Respondents	67,942	40	0	40	0.04%	0.00%	0.04%
Heroin Recency	Lifetime Heroin Users	1,144	28	0	28	2.10%	0.00%	2.10%
Heroin 12-Month Frequency	Past Year Heroin Users	284	19	38	57	7.42%	9.64%	17.06%
Heroin 30-Day Frequency	Past Month Heroin Users	132	7	3	10	4.95%	2.49%	7.44%
Heroin Age at First Use	Lifetime Heroin Users	1,144	11	0	11	0.84%	0.00%	0.84%
Heroin Day of First Use	Lifetime Heroin Users	1,144	1,144	0	1,144	100.00%	0.00%	100.00%
Heroin Month of First Use	Lifetime Heroin Users	1,144	1,048	0	1,048	95.15%	0.00%	95.15%
Heroin Year of First Use	Lifetime Heroin Users	1,144	1,044	0	1,044	94.96%	0.00%	94.96%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.12 Percentage of Cases Imputed for Hallucinogen Use, Dependence, and Abuse Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Hallucinogen Lifetime Use	All Respondents	67,942	290	0	290	0.33%	0.00%	0.33%
Hallucinogen Recency	Lifetime Hallucinogen Users	9,774	259	47	306	1.77%	0.34%	2.12%
Hallucinogen 12-Month Frequency	Past Year Hallucinogen Users	1,909	140	104	244	5.83%	3.98%	9.81%
Hallucinogen 30-Day Frequency	Past Month Hallucinogen Users	533	55	6	61	10.81%	0.55%	11.36%
Hallucinogen Age at First Use	Lifetime Hallucinogen Users	9,774	146	121	267	1.31%	1.24%	2.55%
Hallucinogen Day of First Use	Lifetime Hallucinogen Users	9,774	9,774	0	9,774	100.00%	0.00%	100.00%
Hallucinogen Month of First Use	Lifetime Hallucinogen Users	9,774	8,984	40	9,024	96.09%	0.19%	96.28%
Hallucinogen Year of First Use	Lifetime Hallucinogen Users	9,774	8,937	16	8,953	95.87%	0.08%	95.95%
Hallucinogen Lifetime Use Other than LSD, PCP, Ecstasy, Ketamine, DMT/AMT/Foxy, or <i>Salvia divinorum</i>	All Respondents	67,942	215	0	215	0.25%	0.00%	0.25%
Hallucinogen Dependence	Past Year Hallucinogen Users	1,909	87	0	87	2.93%	0.00%	2.93%
Hallucinogen Abuse	Past Year Hallucinogen Users	1,909	82	0	82	2.81%	0.00%	2.81%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.13 Percentage of Cases Imputed for LSD Use Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
LSD Lifetime Use	All Respondents	67,942	71	0	71	0.07%	0.00%	0.07%
LSD Recency	Lifetime LSD Users	5,539	104	1	105	1.08%	0.01%	1.09%
LSD Age at First Use	Lifetime LSD Users	5,539	108	3	111	1.84%	0.01%	1.85%
LSD Day of First Use	Lifetime LSD Users	5,539	5,539	0	5,539	100.00%	0.00%	100.00%
LSD Month of First Use	Lifetime LSD Users	5,539	5,056	0	5,056	96.06%	0.00%	96.06%
LSD Year of First Use	Lifetime LSD Users	5,539	5,013	2	5,015	95.71%	0.03%	95.74%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.14 Percentage of Cases Imputed for PCP Use Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
PCP Lifetime Use	All Respondents	67,942	52	0	52	0.06%	0.00%	0.06%
PCP Recency	Lifetime PCP Users	1,005	30	0	30	1.49%	0.00%	1.49%
PCP Age at First Use	Lifetime PCP Users	1,005	39	0	39	2.72%	0.00%	2.72%
PCP Day of First Use	Lifetime PCP Users	1,005	1,005	0	1,005	100.00%	0.00%	100.00%
PCP Month of First Use	Lifetime PCP Users	1,005	984	0	984	99.25%	0.00%	99.25%
PCP Year of First Use	Lifetime PCP Users	1,005	982	0	982	99.18%	0.00%	99.18%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.15 Percentage of Cases Imputed for Ecstasy Use Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Ecstasy Lifetime Use	All Respondents	67,942	78	0	78	0.10%	0.00%	0.10%
Ecstasy Recency	Lifetime Ecstasy Users	5,276	98	1	99	1.44%	0.01%	1.45%
Ecstasy Age at First Use	Lifetime Ecstasy Users	5,276	68	1	69	1.44%	0.01%	1.45%
Ecstasy Day of First Use	Lifetime Ecstasy Users	5,276	5,276	0	5,276	100.00%	0.00%	100.00%
Ecstasy Month of First Use	Lifetime Ecstasy Users	5,276	4,793	0	4,793	93.94%	0.00%	93.94%
Ecstasy Year of First Use	Lifetime Ecstasy Users	5,276	4,758	0	4,758	93.61%	0.00%	93.61%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.16 Percentage of Cases Imputed for Ketamine Use Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Ketamine Lifetime Use	All Respondents	67,942	103	0	103	0.11%	0.00%	0.11%
Ketamine Recency	Lifetime Ketamine Users	934	15	1	16	1.17%	0.05%	1.21%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.17 Percentage of Cases Imputed for DMT/AMT/Foxy Use Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
DMT/AMT/Foxy Lifetime Use	All Respondents	67,942	118	0	118	0.17%	0.00%	0.17%
DMT/AMT/Foxy Recency	Lifetime DMT/AMT/Foxy Users	730	18	3	21	2.98%	0.24%	3.22%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.18 Percentage of Cases Imputed for *Salvia divinorum* Use Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
<i>Salvia divinorum</i> Lifetime Use	All Respondents	67,942	134	0	134	0.20%	0.00%	0.20%
<i>Salvia divinorum</i> Recency	Lifetime <i>Salvia divinorum</i> Users	1,620	16	6	22	1.00%	0.13%	1.12%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.19 Percentage of Cases Imputed for Inhalant Use, Dependence, and Abuse Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Inhalant Lifetime Use	All Respondents	67,942	175	0	175	0.17%	0.00%	0.17%
Inhalant Recency	Lifetime Inhalant Users	6,785	264	3	267	2.16%	0.02%	2.18%
Inhalant 12-Month Frequency	Past Year Inhalant Users	743	98	63	161	9.30%	6.66%	15.96%
Inhalant 30-Day Frequency	Past Month Inhalant Users	222	28	8	36	6.90%	4.08%	10.98%
Inhalant Age at First Use	Lifetime Inhalant Users	6,785	504	0	504	4.86%	0.00%	4.86%
Inhalant Day of First Use	Lifetime Inhalant Users	6,785	6,785	0	6,785	100.00%	0.00%	100.00%
Inhalant Month of First Use	Lifetime Inhalant Users	6,785	6,409	3	6,412	97.26%	0.01%	97.27%
Inhalant Year of First Use	Lifetime Inhalant Users	6,785	6,349	0	6,349	96.76%	0.00%	96.76%
Inhalant Dependence	Past Year Inhalant Users	743	68	0	68	6.34%	0.00%	6.34%
Inhalant Abuse	Past Year Inhalant Users	743	65	0	65	5.76%	0.00%	5.76%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.20 Percentage of Cases Imputed for Methamphetamine Use, Dependence, and Abuse Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Methamphetamine Lifetime Use	All Respondents	67,942	87	0	87	0.09%	0.00%	0.09%
Methamphetamine Recency	Lifetime Methamphetamine Users	2,961	44	0	44	0.89%	0.00%	0.89%
Methamphetamine 12-Month Frequency	Past Year Methamphetamine Users	414	27	35	62	5.43%	7.06%	12.48%
Methamphetamine 30-Day Frequency	Past Month Methamphetamine Users	168	7	3	10	3.37%	0.77%	4.14%
Methamphetamine Age at First Use	Lifetime Methamphetamine Users	2,961	34	0	34	1.13%	0.00%	1.13%
Methamphetamine Day of First Use	Lifetime Methamphetamine Users	2,961	2,961	0	2,961	100.00%	0.00%	100.00%
Methamphetamine Month of First Use	Lifetime Methamphetamine Users	2,961	2,837	0	2,837	98.06%	0.00%	98.06%
Methamphetamine Year of First Use	Lifetime Methamphetamine Users	2,961	2,829	0	2,829	97.97%	0.00%	97.97%
Methamphetamine Dependence	Past Year Methamphetamine Users	426	12	0	12	2.17%	0.00%	2.17%
Methamphetamine Abuse	Past Year Methamphetamine Users	426	12	0	12	2.17%	0.00%	2.17%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.21 Percentage of Cases Imputed for Pain Reliever Use, Misuse, Dependence, and Abuse Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Pain Reliever Any Use Recency	All Respondents	67,942	723	0	723	0.68%	0.00%	0.68%
Pain Reliever Misuse Recency	Lifetime Pain Reliever Any Users	36,773	290	0	290	0.53%	0.00%	0.53%
Pain Reliever 30-Day Frequency	Past Month Pain Reliever Misusers	969	15	12	27	1.21%	1.61%	2.81%
Pain Reliever Past Year Initiation	Past Year Pain Reliever Misusers	3,361	42	0	42	1.04%	0.00%	1.04%
Pain Reliever Age at First Misuse	Past Year Initiates of Pain Reliever Misuse	747	120	0	120	18.75%	0.00%	18.75%
Pain Reliever Day of First Misuse	Past Year Initiates of Pain Reliever Misuse	747	747	0	747	100.00%	0.00%	100.00%
Pain Reliever Month of First Misuse	Past Year Initiates of Pain Reliever Misuse	747	165	0	165	26.91%	0.00%	26.91%
Pain Reliever Year of First Misuse	Past Year Initiates of Pain Reliever Misuse	747	124	0	124	20.41%	0.00%	20.41%
Pain Reliever Dependence	Past Year Pain Reliever Misusers	3,361	53	0	53	1.55%	0.00%	1.55%
Pain Reliever Abuse	Past Year Pain Reliever Misusers	3,361	41	0	41	1.24%	0.00%	1.24%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.22 Percentage of Cases Imputed for OxyContin® Use and Misuse Variables

		Unweighted Frequencies				Weighted Percentages		
	Domain	Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
OxyContin Past Year Any Use	All Respondents	67,942	265	236	501	0.33%	0.21%	0.54%
OxyContin Past Year Misuse	Past Year OxyContin Any Users	2,301	11	1	12	0.32%	0.01%	0.33%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.23 Percentage of Cases Imputed for Tranquilizer Use, Misuse, Dependence, and Abuse Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Tranquilizer Any Use Recency	All Respondents	67,942	521	0	521	0.64%	0.00%	0.64%
Tranquilizer Misuse Recency	Lifetime Tranquilizer Any Users	12,691	76	0	76	0.60%	0.00%	0.60%
Tranquilizer 30-Day Frequency	Past Month Tranquilizer Misusers	584	9	4	13	1.47%	0.21%	1.69%
Tranquilizer Past Year Initiation	Past Year Tranquilizer Misusers	1,872	20	0	20	0.89%	0.00%	0.89%
Tranquilizer Age at First Misuse	Past Year Initiates of Tranquilizer Misuse	543	21	0	21	4.44%	0.00%	4.44%
Tranquilizer Day of First Misuse	Past Year Initiates of Tranquilizer Misuse	543	543	0	543	100.00%	0.00%	100.00%
Tranquilizer Month of First Misuse	Past Year Initiates of Tranquilizer Misuse	543	47	0	47	9.61%	0.00%	9.61%
Tranquilizer Year of First Misuse	Past Year Initiates of Tranquilizer Misuse	543	25	0	25	5.42%	0.00%	5.42%
Tranquilizer Dependence	Past Year Tranquilizer Misusers	1,872	26	0	26	0.95%	0.00%	0.95%
Tranquilizer Abuse	Past Year Tranquilizer Misusers	1,872	20	0	20	0.76%	0.00%	0.76%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.24 Percentage of Cases Imputed for Stimulant Use, Misuse, Dependence, and Abuse Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Stimulant Any Use Recency	All Respondents	67,942	317	0	317	0.35%	0.00%	0.35%
Stimulant Misuse Recency	Lifetime Stimulant Any Users	9,158	62	0	62	0.48%	0.00%	0.48%
Stimulant 30-Day Frequency	Past Month Stimulant Misusers	618	6	1	7	0.64%	0.07%	0.71%
Stimulant Past Year Initiation	Past Year Stimulant Misusers	2,060	20	0	20	0.98%	0.00%	0.98%
Stimulant Age at First Misuse	Past Year Initiates of Stimulant Misuse	523	16	0	16	4.60%	0.00%	4.60%
Stimulant Day of First Misuse	Past Year Initiates of Stimulant Misuse	523	523	0	523	100.00%	0.00%	100.00%
Stimulant Month of First Misuse	Past Year Initiates of Stimulant Misuse	523	33	0	33	8.08%	0.00%	8.08%
Stimulant Year of First Misuse	Past Year Initiates of Stimulant Misuse	523	17	0	17	4.83%	0.00%	4.83%
Stimulant Dependence	Past Year Stimulant Misusers	2,060	18	0	18	0.62%	0.00%	0.62%
Stimulant Abuse	Past Year Stimulant Misusers	2,060	17	0	17	0.63%	0.00%	0.63%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.25 Percentage of Cases Imputed for Sedative Use and Misuse Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Sedative Any Use Recency	All Respondents	67,942	379	0	379	0.46%	0.00%	0.46%
Sedative Misuse Recency	Lifetime Sedative Any Users	7,364	56	0	56	0.56%	0.00%	0.56%
Sedative 30-Day Frequency	Past Month Sedative Misusers	107	1	2	3	0.51%	1.18%	1.69%
Sedative Past Year Initiation	Past Year Sedative Misusers	389	4	0	4	0.51%	0.00%	0.51%
Sedative Age at First Misuse	Past Year Initiates of Sedative Misuse	97	7	0	7	7.88%	0.00%	7.88%
Sedative Day of First Misuse	Past Year Initiates of Sedative Misuse	97	97	0	97	100.00%	0.00%	100.00%
Sedative Month of First Misuse	Past Year Initiates of Sedative Misuse	97	11	0	11	9.49%	0.00%	9.49%
Sedative Year of First Misuse	Past Year Initiates of Sedative Misuse	97	7	0	7	7.88%	0.00%	7.88%
Sedative Dependence	Past Year Sedative Misusers	389	4	0	4	0.54%	0.00%	0.54%
Sedative Abuse	Past Year Sedative Misusers	389	4	0	4	0.67%	0.00%	0.67%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.26 Percentage of Cases Imputed for Demographic Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
General								
Education Level	All Respondents	67,942	64	0	64	0.07%	0.00%	0.07%
Marital Status	Respondents at Least 15 Years of Age	59,519	10	0	10	0.04%	0.00%	0.04%
Employment Status	Respondents at Least 15 Years of Age	59,519	725	0	725	1.12%	0.00%	1.12%
Employment Status: 18+	Respondents at Least 18 Years of Age	50,861	601	0	601	1.11%	0.00%	1.11%
Immigrant Status								
Born in the United States	All Respondents	67,942	289	251	540	0.37%	0.33%	0.70%
Immigrant Age of Entry in the United States	Respondents Not Born in the United States	8,364	252	0	252	2.72%	0.00%	2.72%
Hispanic or Latino Origin								
Hispanic or Latino Origin	All Respondents	67,942	222	4	226	0.14%	0.01%	0.15%
Hispanic or Latino Origin Group	Hispanic or Latino Respondents	12,150	83	94	177	0.46%	0.85%	1.30%
Single/Multiple Hispanic or Latino Origin Group	Hispanic or Latino Respondents	12,150	88	99	187	0.48%	0.86%	1.34%
Race								
Native Hawaiian	All Respondents	67,942	2,597	1	2,598	3.73%	0.00%	3.73%
Guamanian/Chamorro	All Respondents	67,942	2,597	1	2,598	3.73%	0.00%	3.73%
Samoan	All Respondents	67,942	2,597	0	2,597	3.73%	0.00%	3.73%
Other Pacific Islander	All Respondents	67,942	2,597	2	2,599	3.73%	0.00%	3.73%
Asian	All Respondents	67,942	2,606	343	2,949	3.75%	0.53%	4.28%
Black or African American	All Respondents	67,942	2,640	27	2,667	3.79%	0.03%	3.82%
American Indian/Alaska Native	All Respondents	67,942	2,733	30	2,763	3.80%	0.02%	3.82%
White	All Respondents	67,942	2,748	139	2,887	3.83%	0.23%	4.06%
Detailed Race: 15 Levels	All Respondents	67,942	2,809	514	3,323	3.92%	0.78%	4.70%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.27 Percentage of Cases Imputed for Roster Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Household Roster								
Number of People in Household	All Respondents	67,942	50	0	50	0.09%	0.00%	0.09%
Number of Children in Household Aged < 18	All Respondents	67,942	222	11	233	0.34%	0.02%	0.37%
Number of People in Household Aged $\geq$ 65	All Respondents	67,942	282	65	347	0.35%	0.10%	0.44%
Family Roster								
Presence of Family Members in Household	All Respondents	67,942	62	0	62	0.11%	0.00%	0.11%
Number of Respondent's Family Members in Household including Foster Relationships	All Respondents	67,942	72	8	80	0.12%	0.02%	0.14%
Number of Respondent's Family Members in Household Aged < 18 including Foster Relationships	All Respondents	67,942	157	0	157	0.26%	0.00%	0.26%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.28 Percentage of Cases Imputed for Income Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Amount of Income								
Respondent's Total Income > or < \$20,000	All Respondents	67,942	945	26	971	2.08%	0.02%	2.11%
Respondent's Total Income (Finer Categories)	All Respondents	67,942	2,040	0	2,040	4.59%	0.00%	4.59%
Total Family Income > or < \$20,000	All Respondents	67,942	2,705	0	2,705	4.30%	0.00%	4.30%
Total Family Income (Finer Categories)	All Respondents	67,942	6,360	0	6,360	10.39%	0.00%	10.39%
Source of Income								
Family Received Social Security or Railroad Retirement Payments	All Respondents	67,942	568	0	568	0.73%	0.00%	0.73%
Family Received Public Assistance	All Respondents	67,942	415	0	415	0.46%	0.00%	0.46%
Family Received Supplemental Security Income	All Respondents	67,942	571	0	571	0.76%	0.00%	0.76%
Respondent/Other Family Member Received Food Stamps	All Respondents	67,942	308	0	308	0.42%	0.00%	0.42%
Family Received Welfare/Job Placement/Child Care	All Respondents	67,942	316	0	316	0.36%	0.00%	0.36%
Number of Months on Welfare	Family Receives Public Assistance or Welfare/Job Placement/Child Care	3,292	141	0	141	4.52%	0.00%	4.52%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.29 Percentage of Cases Imputed for Health Insurance Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Respondent Has Health Insurance								
Overall Health Insurance	All Respondents	67,942	567	0	567	0.56%	0.00%	0.56%
Type of Insurance								
Private	All Respondents	67,942	464	0	464	0.56%	0.00%	0.56%
Medicare	All Respondents	67,942	297	0	297	0.34%	0.00%	0.34%
Military Health Care (CHAMPUS, TRICARE, CHAMPVA, VA)	All Respondents	67,942	166	0	166	0.20%	0.00%	0.20%
Medicaid/CHIP	All Respondents	67,942	524	0	524	0.63%	0.00%	0.63%
Other	Respondent Does Not Have Private Health Insurance, Medicare, Medicaid/CHIP, or CHAMPUS	8,333	190	0	190	1.58%	0.00%	1.58%

CHIP = Children's Health Insurance Program; CHAMPUS = Civilian Health and Medical Program of the Uniformed Services; CHAMPVA = Civilian Health and Medical Program of the Department of Veteran's Affairs; VA = Department of Veteran's Affairs.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

Table A.30 Percentage of Cases Imputed for Pair Variables

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Pair Relationships								
Family Pair Relationship Indicator	All Pair Members	35,694	538	4	542	1.62%	0.04%	1.66%
Multiplicities								
Multiplicity Count: Child-Parent, Parent Focus, Child is 12-14	Parent-Child Pair, Child is 12-14	5,498	2	0	2	0.04%	0.00%	0.04%
Multiplicity Count: Child-Parent, Parent Focus, Child is 12-17	Parent-Child Pair, Child is 12-17	10,096	6	0	6	0.08%	0.00%	0.08%
Multiplicity Count: Child-Parent, Parent Focus, Child is 12-20	Parent-Child Pair, Child is 12-20	11,692	20	0	20	0.24%	0.00%	0.24%
Multiplicity Count: Child-Parent, Parent Focus, Child is 15-17	Parent-Child Pair, Child is 15-17	4,598	4	0	4	0.12%	0.00%	0.12%
Multiplicity Count: Child-Parent, Child Focus, Child is 12-14	Parent-Child Pair, Child is 12-14	5,498	136	0	136	2.29%	0.00%	2.29%
Multiplicity Count: Child-Parent, Child Focus, Child is 12-17	Parent-Child Pair, Child is 12-17	10,096	234	0	234	2.15%	0.00%	2.15%
Multiplicity Count: Child-Parent, Child Focus, Child is 12-20	Parent-Child Pair, Child is 12-20	11,692	268	0	268	2.16%	0.00%	2.16%
Multiplicity Count: Child-Parent, Child Focus, Child is 15-17	Parent-Child Pair, Child is 15-17	4,598	98	0	98	2.01%	0.00%	2.01%
Multiplicity Count: Sibling-Sibling (12-14/15-17), 12-14 Focus	Sibling-Sibling Pair: Younger is 12-14, Older is 15-17	2,842	36	0	36	1.52%	0.00%	1.52%
Multiplicity Count: Sibling-Sibling (12-14/15-17), 15-17 Focus	Sibling-Sibling Pair: Younger is 12-14, Older is 15-17	2,842	38	0	38	1.24%	0.00%	1.24%

Table A.30 Percentage of Cases Imputed for Pair Variables (continued)

	Domain	Unweighted Frequencies				Weighted Percentages		
		Respondents in Domain	Imputed	Logically Assigned	Imputed or Logically Assigned	Imputed	Logically Assigned	Imputed or Logically Assigned
Multiplicity Count: Sibling-Sibling (12-17/18-25), 12-17 Focus	Sibling-Sibling Pair: Younger is 12-17, Older is 18-25	3,252	54	0	54	1.43%	0.00%	1.43%
Multiplicity Count: Sibling-Sibling (12-17/18-25), 18-25 Focus	Sibling-Sibling Pair: Younger is 12-17, Older is 18-25	3,252	50	0	50	1.35%	0.00%	1.35%
Household Person-Level Count								
Household Count: Number of Spouse-Spouse Pairs in Household	All Respondents	67,942	151	0	151	0.30%	0.00%	0.30%
Household Count: Number of Spouse-Spouse Pairs with Children	All Respondents	67,942	78	0	78	0.11%	0.00%	0.11%
Household Count: Child-Parent, Parent Focus, Child is 12-14	All Respondents	67,942	305	0	305	0.31%	0.00%	0.31%
Household Count: Child-Parent, Parent Focus, Child is 12-17	All Respondents	67,942	470	0	470	0.52%	0.00%	0.52%
Household Count: Child-Parent, Parent Focus, Child is 12-20	All Respondents	67,942	546	0	546	0.60%	0.00%	0.60%
Household Count: Child-Parent, Child Focus, Child is 12-14	All Respondents	67,942	90	0	90	0.10%	0.00%	0.10%
Household Count: Child-Parent, Child Focus, Child is 12-17	All Respondents	67,942	156	0	156	0.20%	0.00%	0.20%
Household Count: Child-Parent, Child Focus, Child is 12-20	All Respondents	67,942	227	0	227	0.28%	0.00%	0.28%
Household Count: Sibling-Sibling (12-14/15-17), 15-17 Focus	All Respondents	67,942	64	0	64	0.08%	0.00%	0.08%
Household Count: Sibling-Sibling (12-17/18-25), 18-25 Focus	All Respondents	67,942	134	0	134	0.21%	0.00%	0.21%

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016.

**Appendix B: Detailed Documentation of Edits for Variables
from the Self-Administered Demographics, Substance Use,
and Additional Self-Administered Sections in the
2016 NSDUH**

This page intentionally left blank

Appendix B: Detailed Documentation of Edits for Variables from the Self-Administered Demographics, Substance Use, and Additional Self-Administered Sections in the 2016 NSDUH

B.1 Introduction

This appendix provides detailed documentation of edits that were implemented (where applicable) for variables from the self-administered demographics, substance use, and additional self-administered sections (i.e., sections that also had been self-administered prior to 2016). The remaining sections of the appendix are organized as follows:

- Section B.2 presents tables describing edits for selected self-administered demographics variables corresponding to editing procedures that were discussed in Chapter 5.
- Section B.3 presents tables describing edits for substance use variables corresponding to editing procedures that were discussed in Chapter 6.
- Section B.4 presents tables describing edits for variables in additional self-administered sections corresponding to editing procedures that were discussed in Chapter 7.

The issues that are shown in the tables in these sections reflect changes to the questionnaire for 2016 (Section 1.4.1). Issues from 2014 or prior years are no longer discussed if they have been resolved because of the questionnaire changes in 2016. For questions or sections of the interview that did not change between 2014 and 2016, the issues that are shown in these tables reflect cumulative response patterns that have been observed in the data since a given section or set of questions was added to the interview. However, not all of these patterns may be encountered in a given survey year. In addition, no editing was done for some types of responses that are presented in the tables. Nevertheless, this information is included to document the types of noteworthy responses for which the data were not edited.

B.2 Detailed Documentation of Edits for Selected Self-Administered Demographics Variables

This section presents tables describing edits for the current education and the employment and workplace variables from the self-administered demographics.

- [Table B.1](#) presents edits pertaining to variables in the current education section.
- [Table B.2](#) presents edits pertaining to variables in the employment and workplace section.

Table B.1 Edits Pertaining to the Current Education Section

Issue	Edits Implemented
The respondent (R) did not know or refused to report whether he or she was going to school, reported being on a holiday or break from school (QD17A = 1), but reported that he or she did not intend to return to school once the break was over (QD17B = 2).	The R was defined as not going to school (EDUSCHLGO = 2). The variables pertaining to the current grade through the number of days that the R skipped school in the past 30 days (EDUSCHGRD, EDUFULPAR, EDUSCKMON, and EDUSKPMON) were assigned legitimate skip codes.
For Rs aged 12 to 17 who completed the 11th grade or lower, the current grade (QD18) was potentially inconsistent with the highest grade that the R reported completing (QD11) despite the presence of a consistency check between QD11 and QD18. Also, the consistency check generally applied to the completion of levels of education beyond the 11th grade and therefore was less applicable to respondents aged 12 to 17.	An algorithm was developed for Rs aged 12 to 17 who had completed the 11th grade or lower and who were currently in the 12th grade or lower. The algorithm compared the self-reported current and highest grades with the R's current age (see text). An edited variable for the highest grade completed (EDUHGHEDT) also was created. Edits were generally implemented as follows: • When both the current grade and the highest completed grade were potentially consistent with the R's age, the edits picked the response from QD18 or QD11 that would yield the most consistent data. The second variable in the pair then was edited for consistency with the response that was picked as being most consistent. • When the current grade was more consistent with the R's current age than was the reported highest level of education from the interviewer-administered demographics, EDUHGHEDT was logically assigned a code to indicate that the R had completed the lower grade that was adjacent to his or her current grade. • When the highest completed level of education was more consistent with the R's current age than was the reported current grade, the edited current grade (EDUSCHGRD) was logically assigned a code to indicate that the R was in the next highest grade relative to the one he or she had completed, or else EDUSCHGRD was coded as bad data. When neither the current grade (QD18) nor the highest completed grade (QD11) was consistent with the R's age, either EDUHGHEDT or EDUSCHGRD (or both) were coded as bad data. • If the current grade was lower than the expected grade, then the answer that was most consistent with the R's current age was chosen and the other variable in the pair was coded as bad data. • If the current grade was exactly two grades higher than the last grade but the highest grade was lower than the expected highest grade, then EDUHGHEDT was coded as bad data. • If the current grade was more than two grades higher than the last grade but the current grade was lower than the expected current grade, then EDUSCHGRD was coded as bad data. • Otherwise, if the current grade was more than two grades higher than the last grade, then both EDUHGHEDT and EDUSCHGRD were coded as bad data.

Table B.1 Edits Pertaining to the Current Education Section (continued)

Issue	Edits Implemented
The R was aged 12 to 17 and reported being in college.	No editing was done in the following situations: • The report of currently being in college was consistent with the R's age and highest level of education. • The R was aged 16 or 17, received a high school diploma or a general educational development (GED) certificate, and reported currently being in the first year of college. EDUHGHEDT or EDUSCHGRD (or both) were coded as bad data in the following situations: • If the highest level of education was consistent with the R's current age but the current year in college was not, then EDUSCHGRD was coded as bad data. • If the R was aged 16 or 17 and reported completing the 12th grade without receiving a diploma, then EDUSCHGRD was coded as bad data. • If both the highest level of education and the report of currently being in college were inconsistent with the R's current age, then both EDUHGHEDT and EDUSCHGRD were coded as bad data.
The R was aged 12 to 17, was in the 1st through 12th grades, and reported receiving a doctorate or professional degree.	EDUHGHEDT was coded as bad data.
The R was aged 18 or older, was currently in school, completed the 11th grade or lower, was currently in the 12th grade or lower, and the current grade in QD18 was zero, one, or two grade levels higher than the highest grade in QD11.	No editing was done to EDUHGHEDT or EDUSCHGRD in these situations.
The R was aged 18 or older, was currently in school, completed the 11th grade or lower, was currently in the 12th grade or lower, and the current grade in QD18 was lower than the highest grade or more than two grade levels higher than the highest grade from QD11.	EDUSCHGRD was coded as bad data. No editing was done to EDUHGHEDT.
The R was aged 18 or older, reported completing the 12th grade without receiving a diploma (QD11 = 13), and was currently in the 12th grade.	No editing was done to EDUHGHEDT or EDUSCHGRD. Adults who had not gotten a high school diploma could report that they were in the 12th grade if they were getting their GED certificate of high school completion.

Table B.1 Edits Pertaining to the Current Education Section (continued)

Issue	Edits Implemented
The R was aged 18 or older and reported being in college. However, the R also reported completing the 11th grade or lower or completing the 12th grade without receiving a high school diploma.	The following edits were implemented if the R was (a) aged 18 or 19 and reported being in the <i>first or second</i> year of college, (b) aged 20 to 22 and reported being in the <i>first through fourth</i> year of college, or (c) aged 23 or older and in the fifth or higher year of college: • If the highest level of education was answered as 1 or 2, then EDUHGHEDT was set to a value of 32 (Regular high school diploma LOGICALLY ASSIGNED) because answers of 1 or 2 in QD11 could be typographical errors for answers of 12. • If the highest level of education was answered as 4, then EDUHGHEDT was set to a value of 34 (GED certificate of high school completion LOGICALLY ASSIGNED) because QD11 = 4 could be a typographical error for an answer of 14. • If the highest grade was the 10th grade or lower, then EDUHGHEDT was coded as 85 because the current year in college was more consistent with the R's age. EDUSCHGRD was coded as bad data in the following situations: • The R had completed the 10th or 11th grades or had completed the 12th grade without receiving a high school diploma. • The R reported receiving a bachelor's degree and was in the 12th grade or lower.
The R was aged 23 or older, the highest level of education was a high school diploma or lower, and the current grade was lower than the highest level of education.	EDUSCHGRD was coded as bad data.
The R was aged 23 or older, and the current grade was beyond the 12th grade. The values for the highest level of education and the current grade differed by 2 or more, and the value of the highest level of education plus a factor of 10 was lower than the current year in school.	EDUSCHGRD was coded as bad data.
The highest level of education was a high school diploma, and the R reported currently being in the 12th grade (regardless of age).	EDUSCHGRD was coded as bad data.
The R received a GED and reported being in the 1st through 12 grades (regardless of age).	The following age-specific edits were implemented: • EDUHGHEDT was coded as bad data for Rs aged 12 to 17. • EDUSCHGRD was coded as bad data for Rs aged 18 or older.
The R received some college credit but no degree, and the R was currently in the 8th grade or lower (regardless of age).	EDUSCHGRD was coded as bad data.

Table B.1 Edits Pertaining to the Current Education Section (continued)

Issue	Edits Implemented
The R received an associate's degree, and the R was currently in the 12th grade or lower (regardless of age).	The following age-specific edits were implemented: • EDUHGHEDT was coded as bad data for Rs aged 12 to 17. • EDUSCHGRD was coded as bad data for Rs aged 18 or older.
The R received a master's degree and was in the 12th grade or lower.	The following age-specific edits were implemented: • EDUHGHEDT was coded as bad data for Rs aged 12 to 17. • EDUSCHGRD was coded as bad data for Rs aged 18 or older.
The R did not report being a full-time student in QD19.	EDUSCKMON and EDUSKPMON were assigned legitimate skip codes if the R reported being a part-time student. EDUSCKMON and EDUSKPMON retained values of 98 (blank) if QD19 was answered as "don't know" or "refused."
Question QD21 was skipped because the R reported in QD20 that school was not in session in the past 30 days (QD20 = 90).	EDUSKPMON was assigned a legitimate skip code, and EDUSCKMON retained a code of 90.
The number of days in the past 30 that the R missed school because the R was sick and the number of days that the R skipped school added up to more than 30.	No editing was done in this situation.
The R reported being currently enrolled in school but reported skipping school all 30 days in the past 30 days.	A code of 11 was assigned to the school variable EDUSCHLGO. This code was intended to indicate to analysts that there was some uncertainty about whether the R was currently going to school.

Table B.2 Edits Pertaining to the Employment and Workplace Section

Issue	Edits Implemented
The R did not know or refused to report in question QD26 whether he or she worked in the past year. However, the R also reported in question QD33 (edited variable WRKDPSTYR) that he or she did not have a job in the past 12 months.	The R was logically inferred not to have worked in the past week (WRKDPSTWK = 4) and not to have had a job in the past week (WRKHADJOB = 4), where 4 = No LOGICALLY ASSIGNED. Subsequent employment and workplace variables that could be assigned legitimate skip codes were edited as though QD26 and QD27 had been answered as "no."
The R reported working in the past week. However, the R subsequently reported missing work for all 30 of the past 30 days because he or she was sick or did not want to be at work (or both). Because the past week was included in the 30 days prior to the interview, it would be inconsistent for an R to report working in the past week but missing work for every day in the past month.	The following edits were implemented in this situation: • If the R reported that he or she missed work for all 30 days in the past month because he or she was sick, the edited variable (WRKSICKMO) was assigned a bad data code. • If the R reported missing work for all 30 days in the past month because he or she did not want to be there, the edited variable (WRKSKIPMO) was assigned a bad data code.
The respondent (R) reported working in the past week in question QD26. However, the R subsequently reported being without a job at some point in the past 12 months and reported being without a job during all 52 weeks in the past 12 months. Because all 52 weeks of the 12-month period prior to the interview would include the week prior to the interview, it would be inconsistent for an R to report working in the past week but not working for all 52 weeks in the past year.	The edited variable pertaining to the number of weeks without a job in the past 12 months (WRKNJBWKS) was assigned a bad data code.
The R answered question QD26 (worked in the past week) as "no" but answered question QD27 (having a job in the past week) as "don't know." Edit logic prior to 2003 left the variable pertaining to the number of hours worked in the past week WRKDHRSWK as blank (i.e., a legitimate skip code was not assigned to WRKDHRSWK). The prior logic for assigning legitimate skip codes to WRKDHRSWK was part of the logic for assigning legitimate skips when both QD26 and QD27 were answered as "no." However, only the response to question QD26 truly applies to WRKDHRSWK.	Since 2003, a legitimate skip code has been assigned to WRKDHRSWK when QD26 is answered as "no" (QD26 = 2), independent of how QD27 was answered.
The reported year when the R last worked for pay was fewer than 5 years from the R's birth year (including situations where the year the R reported last working for pay was earlier than the year the R was born).	The edited variables WRKLASTMO and WRKLASTYR were assigned bad data codes.

Table B.2 Edits Pertaining to the Employment and Workplace Section (continued)

Issue	Edits Implemented
The R was not asked whether he or she was self-employed in the past 12 months because the R had previously reported not working in the past week because he or she was self-employed and did not have any business (QD30 = 5).	The edited variable pertaining to self-employment in the past 12 months (WRKSELFEM) was assigned a code to indicate that "yes" could be logically inferred. This was done instead of assigning a legitimate skip code.

B.3 Detailed Documentation of Edits for Substance Use Variables

This section presents tables that detail the edits that were applied to variables in the substance use sections.

- [Tables B.3 to B.5](#) present edits pertaining to variables for the frequency of use of different substances in the past 12 months or past 30 days, corresponding to the discussion of procedures in Section 6.2.4.
- [Tables B.6 to B.8](#) present edits pertaining to variables for the initiation of use of a given substance, corresponding to the discussion of editing procedures in Sections 6.2.5 and 6.3.6.
- [Table B.9](#) presents miscellaneous edits pertaining to the tobacco variables.
- Finally, in this section, [Table B.10](#) presents edits pertaining to how respondents obtained the last prescription psychotherapeutic drugs (i.e., pain relievers, tranquilizers, stimulants, and sedatives) the last time they misused them in the past 12 months.

Table B.3 Edits Pertaining to 12-Month and 30-Day Frequency Variables for Tobacco through Methamphetamine Other than for Parent/Child Relationships or Binge Alcohol Use

Issue	Edits Implemented
Number of days that the respondent (R) used in the past 12 months is greater than 365 days (e.g., if the R reported using 31 days per month on average, then $31 \times 12 = 372$ days).	Nonmissing values greater than 365 were trimmed back to a maximum of 365 days. A flag was set to indicate that this trimming was done.
Number of days that the R reported using in the past 30 days implies less use in the past 12 months than what the R reported in the 12-month frequency (e.g., if the R reported using on only 1 day in the past 30 days and did not use on 29 days, the maximum number of days that the R could have used in the past 12 months would be 336 days); this applied to alcohol through methamphetamine (but not to the child hallucinogens).	The 12-month frequency was trimmed to be consistent with the number of days that the R reported using in the past 30 days. The relevant flag variable indicated that this trimming was done.
Rs who answered the 30-day frequency questions as "don't know" (DK) or "refused" (REF) had the opportunity to give their best estimate of the number of days they used in the past 30 days. These best estimate variables were categorical as opposed to being continuous (e.g., 1 = 1 or 2 days, 2 = 3 to 5 days). For alcohol through methamphetamine, the maximum possible value of the "estimated" 30-day frequency could imply less use in the past 12 months than what the R reported in the 12-month frequency. For example, if the R reported using a substance on "10 to 19 days" in the past 30 days, the R would not have used the substance on at least 11 days in that period. It would therefore be inconsistent for the R to have reported using the substance on more than 354 days in the past 12 months.	The 12-month frequency was trimmed to be consistent with the minimum number of days that the R estimated using in the past 30 days. For example, if the R estimated using a substance on "10 to 19" days in the past 30 days, it was assumed that the R could have used on 345 days in the past 12 months, which would be consistent with use on 10 days in the past 30 days (and nonuse on 20 days). This value would not be considerably different from that derived by assuming use on the maximum number of estimated days in the past 30 days (e.g., 354 possible days of use in the past 12 months, if the R reported using on 10 to 19 days in the past 30 days).
For Rs who estimated their 30-day frequency, the final 12-month frequency was consistent with some, but not all, values in the range for the estimated 30-day frequency.	No editing was done to the 12-month frequency. For example, if an R reported using a substance on "6 to 9 days" in the past 30 days and reported using on 8 days in the past 12 months, the 12-month frequency was within the range of 6 to 8 days. Thus, a 12-month frequency of 8 days would be consistent with the estimated 30-day frequency, as long as the R used on 6 to 8 days in the past 30 days (and did not use on 9 days in that period). In this situation, minimum and maximum possible values were created for use by the statistical imputation team in assigning a final 30-day frequency. This procedure narrowed the allowable range of the 30-day frequency for consistency with the 12-month frequency, with the final 30-day frequency being picked from within that narrowed range. In the above example, the allowable range for assigning the final 30-day frequency was 6 to 8 days rather than 6 to 9 days.

Table B.3 Edits Pertaining to 12-Month and 30-Day Frequency Variables for Tobacco through Methamphetamine Other than for Parent/Child Relationships or Binge Alcohol Use (continued)

Issue	Edits Implemented
R initiated use at some point in the past 12 months. Relative to when the R was interviewed, the R therefore could not have used over the entire 12-month period.	The value of the 12-month frequency was reduced (i.e., prorated) according to the maximum possible allowable number of days that the R could have used in the past 12 months, relative to when the R was interviewed. If the R answered the 12-month frequency in terms of a total number of days used in that period and the answer was greater than the maximum number of days that the R could have used, the 12-month frequency was reduced to this upper limit. The maximum possible allowable number of days was determined from the interview date and the month- and year-of-first use or the date of the R's last birthday, whichever yielded the smallest number of days that the R could have used. If the R reported first using in the survey year but did not report the month when he or she first used, it was assumed that the R potentially started use in January of the survey year.
The R initiated use in the past 12 months but definitely initiated use more than 30 days prior to the interview date. The R also reported use on only 1 day in the past 12 months. However, the recency had a value of 1 (Used in the past 30 days) or 8 (Used at some point in the past 12 months LOGICALLY ASSIGNED).	The 12-month frequency of 1 day was set to bad data. If the R last used in the past 30 days and the R initiated use more than 30 days ago but within the past 12 months, the R logically would have used on at least 1 day in the past 30 days and on at least 1 other day in the period that was more than 30 days ago but within the past 12 months. Similarly, a recency of 8 indicated that there was uncertainty about whether the R last used in the past 30 days or more than 30 days ago but within the past 12 months. Consequently, use on only 1 day in the past 12 months and initiation more than 30 days ago within the past 12 months would have forced the recency to be more than 30 days ago. In addition, when the recency had indicated use in the past 30 days and the above edit had changed a 12-month frequency from 1 day to bad data, whatever values that existed in the 30-day variables were retained. In particular, if the 30-day frequency had a missing value, that missing value was retained. (Other edits described below would have logically inferred that the R used on 1 day in the past 30 days if the R also had indicated use on exactly 1 day in the past 12 months.)

Table B.3 Edits Pertaining to 12-Month and 30-Day Frequency Variables for Tobacco through Methamphetamine Other than for Parent/Child Relationships or Binge Alcohol Use (continued)

Issue	Edits Implemented
Information was needed for establishing constraints on donors for the imputation of the 12-month frequency-of-use variables. For example, if a 12-month frequency variable had been changed from 1 day to bad data, the imputation team needed to constrain the imputation of the 12-month frequency to a value greater than 1 day.	Indicator flags were created to aid the statistical imputation team in establishing specific constraints for the imputation of the 12-month frequency-of-use variables. The indicator flags had the following values: 1 = REQUIRE DONOR WITH 12-MONTH FREQUENCY OF USE GREATER THAN 1 DAY. 2 = REQUIRE DONOR WITH 12-MONTH FREQUENCY OF USE GREATER THAN RECIPIENT'S 30-DAY FREQUENCY OF USE. (Did not apply to the psychotherapeutics, which do not have 30-day frequency variables.) 3 = REQUIRE DONOR WITH 12-MONTH FREQUENCY OF USE GREATER THAN RECIPIENT'S 30-DAY FREQUENCY OF BINGE ALCOHOL USE. (Applied only to alcohol).
R initiated use at some point in the past 30 days. Relative to when the R was interviewed, however, the 30-day frequency (or the minimum value of the "estimated" 30-day frequency) exceeded the maximum possible number of days that the R could have used the substance in that period.	If the R gave an exact number of days for the 30-day frequency, the value was reduced to agree with the maximum possible number of days that the R could have used the substance. If the R estimated the 30-day frequency as a range of a possible number of days, the range for the estimated 30-day frequency was reduced to agree with the maximum possible number of days that the R could have used. For example, if the R estimated using a substance on 20 to 29 days in the past 30 days but could have used the substance at most on 15 days in that period, the estimated 30-day frequency was revised to the range indicating use on 10 to 19 days.
The 30-day frequency was greater than 12-month frequency; this applied to alcohol through methamphetamine (but not to the child hallucinogens).	The value for the 30-day frequency was retained, and the 12-month frequency was set to a missing value (i.e., bad data). (Any of the corresponding 12-month frequency source variables that the respondent answered were set to bad data as well.) This edit conserved data reported by the R that suggest more frequent use in the past 30 days. Although both the 12-month and 30-day frequencies could have been set to missing (i.e., bad data) and subsequently imputed, a drawback of this approach is that a donor could be chosen whose 30-day and 12-month frequency data were consistent but whose data were considerably different from what the recipient R originally reported, especially if the recipient R's original answer to the 30-day frequency was at or close to 30.

Table B.3 Edits Pertaining to 12-Month and 30-Day Frequency Variables for Tobacco through Methamphetamine Other than for Parent/Child Relationships or Binge Alcohol Use (continued)

Issue	Edits Implemented
For cigarettes, smokeless tobacco, and cigars, Rs who answered the 30-day frequency question as DK or REF were asked to give their best estimate of the number of days they used these tobacco products in the past 30 days. These best estimate variables were categorical as opposed to being continuous (e.g., 1 = 1 or 2 days, 2 = 3 to 5 days). Because there were no 12-month frequency questions for tobacco products, inconsistencies would not exist between these "estimated" 30-day frequency variables and other data for these tobacco products.	No editing was done to the corresponding 30-day frequency variables (i.e., leaving the original answer of DK or REF). A final value was assigned by the statistical imputation team.
For alcohol through methamphetamine, if Rs gave their best estimate of the number of days they used in the past 30 days, it was possible for the bottom end of the range for an "estimated" 30-day frequency to exceed the value for the 12-month frequency (i.e., after any opportunities that the R had for inconsistency resolution).	The value indicating the range for the "estimated" 30-day frequency was retained, and the 12-month frequency was set to a missing value (i.e., bad data). For example, if the R reported using a substance on 5 days in the past 12 months but estimated use of that substance on 6 to 9 days in the past 30 days, there would be no way for these two answers to be consistent. Setting the 12-month frequency to bad data in this situation is consistent with the edit when Rs report an exact number of days for the final 30-day frequency that is greater than the final 12-month frequency.
The bottom end of the range for the final "estimated" number of days that the R used a substance in the past 30 days is <i>exactly</i> equal to the final 12-month frequency; this applied to alcohol through inhalants (but not to LSD, PCP, or Ecstasy).	The final, edited 30-day frequency was set to the value for the 12-month frequency. The code was retained for the "estimated" 30-day frequency to indicate to analysts where the value came from for the 30-day frequency. For example, if the R reported using a substance on exactly 6 days in the past 12 months and estimated using that substance on 6 to 9 days in the past 30 days, the 30-day frequency was set to exactly 6 days. The two answers in this example are not necessarily inconsistent, as long as the R used on exactly 6 days in the past month. Thus, the response of 6 days in the past 12 months could be thought of as "trapping" the possible number of days that the R could have used in the past 30 days.
The 30-day frequency has a final response of DK or REF (i.e., the R did not give a best estimate for the 30-day frequency), but the 12-month frequency indicates use on exactly 1 day.	The final, edited 30-day frequency was logically inferred to be 1 day. If the R last used in the past 30 days and used on only 1 day in the past 12 months, that 1 day of use had to have occurred in the 30 days prior to the interview.

Table B.3 Edits Pertaining to 12-Month and 30-Day Frequency Variables for Tobacco through Methamphetamine Other than for Parent/Child Relationships or Binge Alcohol Use (continued)

Issue	Edits Implemented
The value for the final 12-month frequency is somewhere within the range for the final "estimated" number of days that the R used a substance in the past 30 days; this applied to alcohol through inhalants (but not to LSD, PCP, or Ecstasy).	No further editing was done because the data were not necessarily inconsistent. For example, if the R reported using a substance on 8 days in the past 12 months and estimated use of that substance on 6 to 9 days in the past 30 days, use of the substance on 6, 7, or 8 days in the past 30 days would still be consistent with use on 8 days in the past 12 months. In this situation, minimum and maximum possible values were created for use by the statistical imputation team in assigning a final 30-day frequency. This procedure narrowed the allowable range of the 30-day frequency for consistency with the 12-month frequency, with the final 30-day frequency being picked from within that narrowed range. In the above example where the R reported using on 8 days in the past 12 months and 6 to 9 days in the past 30 days, the allowable range for the 30-day frequency would be 6 to 8 days, not 6 to 9, in order for the 30-day frequency to be consistent with the 12-month frequency.
Overall 12-month frequency does not agree with the reported preference for giving the answer (e.g., if the R originally indicated that he or she would give an answer in terms of a total number of days per year, answered the days per year question as DK/REF, and then gave an answer in average days per month). This is a function of the computer-assisted interviewing (CAI) logic, which keeps routing Rs through the series until they give a nonmissing answer or answer all 12-month frequency questions as DK/REF; this applied to the alcohol through psychotherapeutics sections. Similarly, the R could change his or her preferred way of reporting the 12-month frequency in response to a consistency check with the 30-day frequency (e.g., starting out as reporting in terms of the total number of days in the past 12 months but changing to average number of days per month in response to a consistency check); this applied to alcohol through methamphetamine inhalants (but not to the child hallucinogens).	The preferred way of reporting the 12-month frequency that best agreed with the final answer to the 12-month frequency was assigned to the "best way" variable (e.g., ALBSTWAY for alcohol), pertaining to the R's preferred way of answering the 12-month frequency questions. Thus, if an R started out as preferring to give an answer in total days per year but the final answer came from an average number of days per month, the preferred method of reporting would be consistent with the reporting in days per month. Any other 12-month frequency questions that the R answered that did not correspond with the final result were overwritten with LEGITIMATE SKIP Logically assigned codes. In the above example, the initial answer in terms of the total number of days used in the past 12 months would be overwritten, and the final answer of average number of days per month would be retained.

Table B.4 Edits Pertaining to Parent/Child 12-Month and 30-Day Frequency Variables for Cocaine and Crack Cocaine

Issue	Edits Implemented
The 30-day or 12-month frequency of crack use is greater than the corresponding frequency of any cocaine use.	The higher value from the crack frequency was assigned to the cocaine frequency. For the 12-month frequency, this also involved moving over the source variables from crack that were associated with the overall crack frequency. In addition, the cocaine variable associated with the preferred method of reporting the 12-month frequency (CCBSTWAY) was bumped by a value of 20 to indicate that data had been moved over from the crack cocaine variables.
Respondents (Rs) indicated that their month of first use for any cocaine was in the same month they were interviewed, and they used cocaine on only 1 day in the past 12 months. However, the Rs also were users of crack cocaine, and they had missing data for the corresponding 12-month frequency variables pertaining to crack cocaine.	Logically, these Rs also would have used crack cocaine, on 1 day in the past 12 months. Therefore, the corresponding 12-month frequency variable for crack cocaine was set to 1 day in this situation.
The R was skipped out of the 12-month or 30-day frequency questions for cocaine because of his or her original answer to the cocaine recency, but nonmissing values existed in the corresponding frequency variables for crack. For example, if an R reported last using any cocaine more than 30 days ago but within the past 12 months, but reported last using crack within the past 30 days, 30-day frequency data could exist for crack but not for any cocaine.	No editing was done to the 12-month or 30-day cocaine frequency that had been skipped. In the imputation stage, the cocaine 12-month or 30-day frequency was subsequently edited or imputed for consistency with the corresponding crack frequency data.
The R estimated the frequency of use of both cocaine and crack in the past 30 days, with the estimated frequency for crack being greater than the estimated frequency for any cocaine.	The "estimated" value for the crack 30-day frequency (CK30EST) was assigned as the estimate for the number of days that the R used any cocaine in that period. The code assigned to the "estimated" cocaine 30-day frequency variable CC30EST was bumped by a value of 10 to indicate that data had been moved over from CK30EST.
The "estimated" 30-day frequency of crack use is greater than the total number of days that the R reported using any cocaine in that same period.	The "estimated" value for the crack 30-day frequency (CK30EST) was assigned as the estimate for the number of days that the R used any cocaine in that period. The code assigned to the "estimated" cocaine 30-day frequency variable CC30EST was bumped by a value of 10 to indicate that data had been moved over from CK30EST. The edited cocaine 30-day frequency variable COCUS30A was assigned a bad data code to wipe out the R's original answer. For example, if the R reported using any cocaine on exactly 5 days in the past 30 days but reported using crack on 6 to 9 days in that period (i.e., a code of 3 in question CK06DKRE), the R was logically inferred to have used any cocaine on 6 to 9 days in the past 30 days. A code of 13 was assigned to CC30EST, corresponding to the code of 3 in CK30EST. In this example, it would be impossible for the R to have used crack on at least 6 days in the past 30 days but to have used any cocaine on only 5 days in that same period.

Table B.4 Edits Pertaining to Parent/Child 12-Month and 30-Day Frequency Variables for Cocaine and Crack Cocaine (continued)

Issue	Edits Implemented
The value for the cocaine 30-day frequency is somewhere within the range given by the "estimated" crack 30-day frequency.	Examination of preliminary data from 2000 indicated that all cocaine use in the past 30 days could not be automatically inferred to be crack use. Therefore, if the value for the cocaine 30-day frequency was the minimum possible number of days that the R could have used crack in that period, the crack 30-day frequency CRKUS30A was set to be equal to the cocaine 30-day frequency COCUS30A. For example, if the R reported using cocaine on exactly 6 days in the past 30 days and estimated using crack on 6 to 9 days in that period, it could be logically inferred that the R used crack on exactly 6 days. If the value for the cocaine 30-day frequency was greater than the minimum value for the estimated crack frequency, no further editing was done. However, information was provided to the statistical imputation team regarding the narrower range in which to assign a final value for the crack 30-day frequency. For example, if the R reported using cocaine on 7 days and crack on 6 to 9 days, the allowable range for the final crack 30-day frequency would be 6 or 7 days.

Table B.5 Edits Pertaining to the 30-Day Alcohol Frequency Variables Related to Binge Alcohol Use

Issue	Edits Implemented
Frequency of binge alcohol use (i.e., consumption of four or more drinks for females or five or more drinks for males) in the past 30 days (ALCBNG30D, from AL08) is greater than the overall frequency of consumption of any alcohol in the past 30 days (ALCDAYS, from AL06), despite the presence of a consistency check.	If the respondent (R) reported binge alcohol use on all 30 days, then the edits inferred that the R drank on all 30 days. If the R reported binge alcohol use on fewer than 30 days, then the value from the binge alcohol frequency ALCBNG30D was retained and the overall 30-day frequency ALCDAYS was set to missing (i.e., bad data). This approach conserved data reported by the R that could indicate heavy alcohol use (i.e., consumption of four or more drinks for females or five or more drinks for males on 5 or more days in the past 30 days).
Frequency of binge alcohol use in the past 30 days (ALCBNG30D) is greater than the maximum possible value for the estimated frequency of consumption of alcohol in the past 30 days (AL30EST, from AL06DKRE).	If the R reported binge alcohol use on all 30 days, then the edits inferred that the R drank on all 30 days (i.e., the variable ALCDAYS was assigned a value of 30). This edit also was done if the R reported having five or more drinks per occasion on all 30 days and the R continued to answer the follow-up question AL06DKRE as DK or REF. If the R reported binge alcohol use on fewer than 30 days, then the value from ALCBNG30D was retained, and the estimated 30-day frequency variable AL30EST was assigned a bad data value. For example, if an R reported having five or more drinks per occasion on 10 days in the past 30 days but estimated drinking alcohol on 6 to 9 days in that period, it would be impossible for these two answers to be consistent.
R drank on 1 day in the past month and the usual number of drinks was four or more (for females) or five or more (for males) on this 1 day, but frequency of binge alcohol use is 0 (or is answered as DK/REF).	A value of 1 day was assigned to ALCBNG30D. If the R met the threshold for binge alcohol use on the 1 day when he or she drank, then the R logically had to engage in binge alcohol use on exactly 1 day in the past month.
R drank on 1 day in the past month and the usual number of drinks was fewer than four (for females) or fewer than five (for males), but the frequency of binge alcohol use is answered as DK/REF.	A code of 80 was assigned to ALCBNG30D, where 80 = NO OCCASION OF 4+/5+ DRINKS IN PAST 30 DAYS Logically assigned. If the R reported having a number of drinks on the 1 day when he or she drank that was below the threshold for binge alcohol use, then the R logically had no occasions of binge alcohol use in the past month.

Table B.5 Edits Pertaining to the 30-Day Alcohol Frequency Variables Related to Binge Alcohol Use (continued)

Issue	Edits Implemented
Frequency of binge alcohol use in the past 30 days (ALCBNG30D) falls within the range of the estimated frequency of consumption of alcohol in the past 30 days (AL30EST).	No further editing was done because the data were not necessarily inconsistent. For example, if the R reported drinking any alcohol on 6 to 9 days in the past 30 days and reported binge alcohol use on 7 days in that period, these data would be consistent as long as the R drank any alcohol on 7, 8, or 9 days in the past 30 days. In this situation, minimum and maximum possible values were created for use by the statistical imputation team in assigning a final 30-day frequency for alcohol. This procedure narrowed the allowable range of the alcohol 30-day frequency variable ALCDAYS for consistency with the five-drink frequency variable ALCBNG30D. The final 30-day frequency for alcohol was picked from within that narrowed range. In the above example where the R reported binge alcohol use on 7 days in the past 30 days and 6 to 9 days in that period, the allowable range for the 30-day frequency would be 7 to 9 days, not 6 to 9, in order for ALCDAYS to be consistent with ALCBNG30D.
The R drank on more than 1 day in the past month and the usual number of drinks per day was five or more, but the reported frequency of five or more drinks per occasion is 0.	<i>Both</i> the usual number of drinks (i.e., ALCUS30D, from AL07) and ALCBNG30D (from AL08) were set to bad data. This approach was consistent with cognitive testing results that suggested that Rs sometimes incorrectly answer AL07 by indicating the total number of drinks they had <i>over the entire 30-day period</i> instead of the usual number of drinks they had per day. Consequently, these edits did not automatically assume that these Rs were binge alcohol users. If ALCBNG30D had a value of DK/REF and the usual number of drinks reported was five or more, then ALCBNG30D retained the corresponding DK/REF value.
The number of days that the R reported binge alcohol use equals the total number of days that the R reported drinking any alcohol (ALCBNG30D = ALCDAYS), but the reported usual number of drinks is below the sex-specific threshold (or is answered as DK/REF).	A code of 975 was assigned to ALCUS30D (i.e., the edited version of AL07) to indicate usual consumption of at least four drinks (for females) or at least five drinks (for males).
The R answered the binge alcohol frequency question AL08 as DK or REF, but the usual number of drinks on those days when the R drank in the past 30 days (from AL07) exceeded the threshold for binge alcohol use.	No editing was done to ALCBNG30D to replace the missing data. However, a flag was set to restrict the imputation to donors who reported binge alcohol use on at least 1 day in the past 30 days.

Table B.6 Edits Pertaining to Initiation Variables Based on Consistency Checks

Issue	Edits Implemented
The respondent (R) indicated in the final verification check (e.g., CGCC22 for cigarettes) that the age at first use (AFU) or age at first misuse (AFM) based on the new month and year of first use (MFU and YFU) or month and year of first misuse (MFM and YFM) was correct (e.g., CGCC22 = 4). (For brevity, the initiation data subsequently refer only to "first use" and to the AFU, YFU, and MFU data but also apply to first misuse for prescription drugs.) The computer-assisted interviewing (CAI) program updated the AFU (such as AGE1STCG for cigarettes) with the value of the age calculated from the MFU and YFU (referred to subsequently as the MYR1ST age, such as MYR1STCG, for cigarettes). However, the new value for the AFU indicated that the R was more than 1 year younger than his or her current age at the time the R first used the drug (e.g., the R was 16, reported first use of the drug at age 15, but then confirmed an MFU and YFU that meant the R was 14 when the R first used the drug). Had the R initially reported this AFU, the R would not have been routed to the MFU and YFU questions.	The updated value was retained for the AFU (e.g., first use at age 14 for a 16-year-old R in this example). Based on this updated AFU, it was logically inferred that the R should have skipped the MFU and YFU items. A code of 9989 was assigned to the YFU variable (e.g., CIGYFU for cigarettes), and a code of 89 was assigned to the MFU variable (e.g., CIGMFU). Consistent with procedures in prior years, therefore, the values for the MFU and YFU were statistically imputed.
The final verification check (e.g., CGCC22) was skipped because the R entered revised data for the MFU and YFU that made them consistent with the AFU.	No editing was done because the R was considered to have resolved the inconsistency.
The final verification check (e.g., CGCC22) was skipped because the R entered a new MFU that was the same as the R's birth month.	The new MFU could be consistent with the AFU, depending on whether the use in that month occurred before or after the R's birthday. No editing was done to the AFU, MFU, and YFU, as long as the revised MFU and YFU were potentially consistent with the AFU. However, the MFU and YFU were set to bad data if they could never be consistent with the AFU. Suppose, for example, that a hypothetical R was born in June 1998, was interviewed in March 2016 (age 17 at the time of the interview), reported first use of a drug at age 17, and initially reported first use in May 2015. First use in May 2015 would have meant that the R was 16 when he or she first used the drug because the R's 17th birthday was not until June 2015. If the R changed the month and year to June 2015, that could be consistent with first use at age 17, if the use occurred after the R's birthday. However, if the R changed the month and year to June 2014, it would never be possible for the R to have first used at age 17 and also to have first used in June 2014. In this latter situation, the MFU and YFU would be set to bad data.

Table B.6 Edits Pertaining to Initiation Variables Based on Consistency Checks (continued)

Issue	Edits Implemented
The R entered a new MFU or YFU that differed from what the R previously reported. The MYR1ST age based on the revised MFU and YFU still mismatched the AFU, but the R indicated in the final verification check that this MYR1ST age was correct.	No editing was done in this situation. The CAI program automatically updated the AFU to be consistent with the updated values reported for the MFU and YFU.
The consistency check was triggered between the AFU and the MYR1ST age calculated from the MFU and YFU. However, the R answered the first consistency check (e.g., CGCC19 for cigarettes) as "don't know" or "refused." The R then exited the consistency check loop without having resolved the inconsistency.	The AFU was retained, but the MFU and YFU were set to bad data. Consequently, the MFU and YFU needed to be imputed, but the AFU did not.
The R entered values in the consistency checks for the MFU and YFU that again yielded a nonmissing MYR1ST age based on the MFU and YFU. However, the R failed to resolve the inconsistency between the AFU and the MYR1ST age. The R also reported either in the first verification check that the MYR1ST age was not correct (e.g., CGCC19 = 6 for cigarettes) or reported in the second verification check that the AFU was correct (e.g., CGCC20 = 2 for cigarettes).	No editing was done to the AFU. The following edits were implemented for the MFU and YFU: • The default edit was to set the MFU and YFU to bad data. • As an exception to the default edit, if the final verification check was answered as "don't know" or "refused," the MFU and YFU were assigned the code that corresponded to the answer in the final verification check.
The R entered values in the consistency checks for the MFU and YFU that again yielded a nonmissing MYR1ST age based on the MFU and YFU. However, the R failed to resolve the inconsistency between the AFU and the MYR1ST age. The R also reported in the second verification check that neither the AFU nor original MYR1ST age was correct (e.g., CGCC20 = 3 for cigarettes).	The following edits were implemented for the AFU, MFU, and YFU: • The default edit was to set the AFU, MFU, and YFU to bad data. • As an exception to the default edit, if the final verification check was answered as "don't know" or "refused," the AFU, MFU, and YFU were assigned the code that corresponded to the answer in the final verification check.
The R triggered the initial consistency check between the AFU, MFU, and YFU and reported that the MYR1ST age calculated from the MFU and YFU was correct (e.g., CGCC19 = 4). However, the R answered the second consistency check (e.g., CGCC20 for cigarettes) as "don't know" or "refused." Consequently, the R did not have an opportunity to correct the inconsistency between the AFU, MFU, and YFU.	The relevant codes for "don't know" or "refused" were assigned to the AFU, MFU, and YFU. The rationale for this edit is that conclusive information did not exist regarding whether the AFU indicated the R's correct age when he or she first used a drug, or whether the MFU and YFU indicated the R's correct age at initiation. Therefore, the AFU, MFU, and YFU all were set to missing values.

Table B.7 Edits Pertaining to Initiation Variables Other than for Parent/Child Relationships

Issue	Edits Implemented
The age at first use (AFU) was greater than the respondent's (R's) reported age; this applied to all drugs except pipe tobacco (which does not have an AFU question) and to the prescription psychotherapeutic drugs (pain relievers, tranquilizers, stimulants, and sedatives). See Section 6.3.6.2 in Chapter 6 for discussion of the logic for comparing the age at first misuse (AFM) for prescription drugs with the R's age.	The final age was accepted as the standard. For tobacco through methamphetamine, the inconsistent AFU was set to bad data. Any month and year data associated with the AFU also were set to bad data. For prescription psychotherapeutic drugs, the R could not proceed with the interview unless the R made the AFM consistent with his or her age or else changed the answer for the AFM to "don't know" (DK) or "refused" (REF).
The AFU had a value of DK or REF including situations where this assignment has been made from the consistency check data (or hard error for the prescription psychotherapeutic drugs).	For drugs other than prescription psychotherapeutic drugs: • If the month of first use (MFU) and year of first use (YFU) questions were skipped because the R answered the AFU as DK or REF to begin with, the DK or REF value from the AFU was propagated onto the skipped MFU and YFU variables. This edit was designed to indicate the reason that the MFU and YFU variables had been skipped. In addition, because the R may have first used the drug within 1 year of his or her current age, the MFU and YFU questions may have been relevant to the R. • If the R had answered the MFU and YFU questions but the final AFU had a value of DK or REF (i.e., due to a consistency check response), the MFU and YFU data were overwritten with the corresponding DK or REF value from the edited AFU. Retaining the MFU and YFU data in this situation would imply that the R used within 1 year of his or her current age. Responses of "don't know" and "refused" were not "propagated" for the prescription drug initiation variables because of the volume of variables. However, the skipped initiation variables still had a missing value (i.e., 98, for blank).
The MFU and/or YFU questions were answered because the initial answer to an AFU was within 1 year of the R's current age. However, the final, edited AFU is more than 1 year younger than the R's current age based on the consistency checks between the AFU, YFU, and MFU that were described in Table B.6 .	The original answers to the MFU and YFU were overwritten with logically assigned legitimate skip codes (Section 2.4.2). If the R originally had answered the relevant AFU as being more than 1 year younger than his or her current age, the CAI program would have skipped the R past the questions about the MFU and YFU.

Table B.7 Edits Pertaining to Initiation Variables Other than for Parent/Child Relationships (continued)

Issue	Edits Implemented
The MFU has been skipped because the R answered the YFU as DK or REF.	For drugs other than prescription psychotherapeutic drugs, the DK or REF value from the YFU was propagated onto the skipped MFU. That is, if the R did not know in what year he or she first used the drug, it was assumed that the R would not know the month either. Similarly, a refusal to answer the YFU was interpreted to be a blanket refusal to answer the month as well as the year. Responses of "don't know" and "refused" were not "propagated" from the year of first misuse (YFM) to the month of first misuse (YFM) for the prescription drug initiation variables because of the volume of variables. However, the skipped MFM variables still had a missing value (i.e., 98, for blank).
The R had missing data for the MFU. However, one of the following occurred: • The YFU was the current calendar year, and the R was interviewed in January. • The R first used at his or her current age, the R first used in the current calendar year, and the R's most recent birthday occurred in the interview month. • The R first used at his or her current age, the R first used in the prior calendar year, and the R's most recent birthday occurred in December.	The MFU could be logically inferred, as indicated below. • The MFU was logically inferred to be January of the interview year. That would be the only month in which the R could have initiated use in the current year. • The MFU was logically inferred to be the interview month. If the R first used in the current calendar year and attained his or her current age in the interview month, the R logically could not have initiated use in any month other than the interview month. • The MFU was logically inferred to be December. If the R first used in the prior calendar year and attained his or her current age in December of that year, the R logically could not have initiated use in any month other than December of the prior calendar year.

Table B.8 Edits Pertaining to Parent/Child Initiation Variables

Issue	Edits Implemented
The age at first use (AFU) for a child drug (e.g., LSDAGE for LSD) indicated first use at age 1 or 2, but the AFU for the parent drug (e.g., HALLUCAGE for any hallucinogen) had a missing value. This applied to first use of cigarettes and age at initiation of daily smoking, and to AFUs for cocaine/crack and hallucinogens/LSD/PCP/Ecstasy or Molly.	Ages of 1 or 2 in the AFU variables for child drugs (e.g., LSDAGE) were set to bad data. The imputation procedures for AFU variables did not allow Rs with AFUs of 1 or 2 to be donors for imputation. Logically, however, an AFU of 1 or 2 for a child drug would preclude the AFU for the parent drug from being imputed to any value other than 1 or 2. For example, if LSDAGE was 1 or 2 and the AFU for any hallucinogen (HALLUCAGE) had a missing value, the only potential imputation donors for HALLUCAGE would be those Rs with HALLUCAGE of 1 or 2. By definition, however, no one could be a donor for imputation. By setting LSDAGE to bad data, the AFUs for both LSD and any hallucinogen would be imputed.
The AFU for a child drug (e.g., LSDAGE) indicated first use at age 3, but the AFU for the parent drug (e.g., HALLUCAGE) had a missing value. This applied to first use of cigarettes and age at initiation of daily smoking, and to AFUs for cocaine/crack and hallucinogens/LSD/PCP/Ecstasy or Molly.	No editing was done when this pattern occurred. Logically, however, the AFU for the parent drug had to have a value of 3 because respondents with AFUs of 1 or 2 (e.g., HALLUCAGE of 1 or 2) were not eligible to be donors in the imputations. The final imputed AFU (e.g., IRHALLUCAGE for any hallucinogen) was set to a value of 3.
The AFU for a child drug (e.g., crack cocaine) was younger than AFU for the parent (e.g., any cocaine). This applied to first use of cigarettes and age at initiation of daily smoking, and to AFUs for cocaine/crack and hallucinogens/LSD/PCP/Ecstasy or Molly.	This edit required the AFUs to be defined (i.e., not missing) for both the parent and child drugs. The younger AFU from the child drug was assigned to the parent AFU. For example, if the AFU of crack cocaine was younger than the AFU of any cocaine, then the edits assigned the value from the crack AFU to the cocaine AFU.
Rs in the hallucinogens section were lifetime users only of LSD, only of PCP, or only of Ecstasy or Molly. This included situations in which the only reported use of a hallucinogen was "some other hallucinogen," but only LSD, only PCP, or only Ecstasy or Molly was specified. The CAI logic skipped these respondents out of questions related to the age, month, and year that they first used the specific hallucinogen.	Data on initiation of hallucinogen use were moved over to the corresponding specific hallucinogen variables that had been skipped. For example, if the R had used only LSD, the hallucinogen age (from HALLUCAGE) was assigned to the LSD AFU (LSDAGE). Similarly, if the R had provided data on the month and year that he or she had first used any hallucinogen, these data were assigned to the LSD month and year variables LSDMFU and LSDYFU. Similar edits were done when Rs had been skipped out of questions related to specific hallucinogens but were logically inferred to be lifetime users only of LSD, only of PCP, or only of Ecstasy or Molly, based on the "OTHER, Specify" responses.

Table B.8 Edits Pertaining to Parent/Child Initiation Variables (continued)

Issue	Edits Implemented
Rs in the hallucinogens section were lifetime users of two or more of the following: LSD, PCP, or Ecstasy or Molly. No other hallucinogen use was indicated. However, the R answered the AFU question for any hallucinogen (edited variable HALLUCAGE) as "don't know" or "refused." In addition, the specific hallucinogens that the R had used had legitimate values for his or her AFU variables (i.e., LSDAGE, PCPAGE, or ECSTMOAGE, for LSD, PCP, or Ecstasy or Molly, respectively).	Logically, the AFU for any hallucinogen (HALLUCAGE) had to be the minimum of the ages indicated in LSDAGE, PCPAGE, or ECSTMOAGE. Therefore, HALLUCAGE was assigned the minimum value from LSDAGE, PCPAGE, or ECSTMOAGE. If this edit was done and month and year data were available for a specific hallucinogen, the month and year variables for any hallucinogen (HALLUCMFU and HALLUCYFU) were edited accordingly. For example, if a missing value in HALLUCAGE was replaced with the value from LSDAGE, and the month of first use (MFU) and year of first use (YFU) for LSD (LSDMFU and LSDYFU) had valid values, the values from LSDMFU and LSDYFU were assigned to HALLUCMFU and HALLUCYFU, respectively. This edit was done <i>only</i> to replace missing data in HALLUCAGE. No editing was done if HALLUCAGE had a value that was lower than the AFU values reported in LSDAGE, PCPAGE, or ECSTMOAGE.
The Rs in the cocaine/crack or hallucinogens sections indicated first use of any drug in that category (e.g., any hallucinogen) at their current age but had missing data for the age when they first used specific drugs in that category (e.g., LSD).	The Rs were logically inferred to have first used the specific drug at their current age.
The AFU the parent drug was at the R's current age and the AFU for the child drug was at 1 year below the R's current age. Consequently, the child AFU had been assigned to the parent AFU.	The associated MFU and YFU from the child drug (e.g., crack) were assigned to the parent's drug's MFU and YFU data (e.g., for any cocaine). However, this edit was not done in the hallucinogens section when the sole hallucinogen use involved use of LSD, PCP, or Ecstasy or Molly. Consequently, some inconsistencies could remain in the edited hallucinogens data.
The Rs in the cocaine/crack or hallucinogens sections indicated first use for the parent drug category (e.g., any hallucinogen) in the current survey year but had missing data for the year when they first used a child drug (e.g., LSD).	The Rs were logically inferred to have first used the child drug in the current year. If the Rs also reported first use of drugs in the parent category (e.g., any hallucinogen) in the same month that they were interviewed and they had missing MFU data for child drugs within that category, the Rs also were logically inferred to have first used the child drug in the same month that they were interviewed. Otherwise, if the Rs originally had reported a month when they first used a child drug but that month had been assigned a code of bad data due to the "flag and impute" edits, the MFU (e.g., LSDMFU) also was reset to the originally reported value if it matched the MFU for any drug in the parent category (e.g., HALLUCMFU).

Table B.8 Edits Pertaining to Parent/Child Initiation Variables (continued)

Issue	Edits Implemented
The Rs in the cocaine/crack or hallucinogens sections indicated first using any drug in that category (e.g., any hallucinogen) in the year immediately prior to the current one (i.e., 2012) but that point of initiation was within 12 months of the interview date. These Rs also had missing data for the year and month when they first used specific drugs in that category (e.g., LSD).	If the Rs originally had reported year and month data for when they first used a specific drug but these data had been assigned codes for bad data due to the "flag and impute" edits, this information was restored <i>only</i> if the original answers matched the MFU and YFU for any drug in that category. For example, if LSDMFU and LSDYFU had been set to bad data but the original answers for the LSD month and year matched the month and year in HALLUCMFU and HALLUCYFU, LSDMFU was equated to HALLUCMFU, and LSDYFU was equated to HALLUCYFU.
The MFU and/or YFU questions for the parent drug were answered because the initial answer to the parent AFU was within 1 year of the R's current age. However, the final, edited AFU for the parent drug was more than 1 year younger than the R's current age (e.g., if the unedited AFU for cocaine was within 1 year of the R's current age, but the AFU of crack was more than 1 year younger than the R's current age).	The original answers to the parent MFU and YFU were overwritten with logically assigned legitimate skip codes (see Section 2.4.2). If the R originally had answered the parent AFU as being more than 1 year younger than his or her current age, the CAI program would have skipped the R past the questions about the parent MFU and YFU.
The AFUs for the parent and child drugs were the same, but the MFU or YFU was earlier for the child drug (e.g., crack).	The earlier month and year data from the child drug (e.g., crack) were assigned to the parent drug's month and year data (e.g., cocaine). However, this edit was not done in the hallucinogens section when the sole hallucinogen use involved use of LSD, PCP, or Ecstasy or Molly but the YFU and MFU data for the parent and child drugs were inconsistent. Consequently, some inconsistencies could remain in the edited hallucinogens data.
For cigarettes, the AFU for any cigarette use (CIGTRY) equaled the age at initiation of daily cigarette use (CIGAGE). However, the R reported initiating daily cigarette use in a year or month that was earlier than what the R reported for first use of any cigarette.	The month and year data for daily cigarette use (CIGDLMFU and CIGDLYFU, respectively) were assigned to the month and year variables for any cigarette use (CIGMFU and CIGYFU, respectively).
For hallucinogens, the respondent was known never to have used any drugs other than the child drugs, the AFU for the parent drug was nonmissing, only one child AFU was missing, and the minimum of the nonmissing child AFUs was greater than the parent AFU.	The missing child AFU variable was assigned to the parent AFU value.
For hallucinogens, the respondent was known never to have used any drugs other than the child drugs, the parent AFU was nonmissing, only one child AFU was missing, the minimum of the nonmissing child AFUs was equal to the parent AFU, and the earliest of the nonmissing child MFUs and/or YFUs was later than the parent MFU and/or YFU.	The missing child AFU variable was assigned to the parent AFU value.

Table B.8 Edits Pertaining to Parent/Child Initiation Variables (continued)

Issue	Edits Implemented
The parent AFU was equal to the respondent's current age, and some or all of the AFUs for the child drugs were missing.	All missing child AFU variables were assigned to the respondent's current age.
The parent AFU was equal to 1 less than the respondent's current age, the child recency ¹ was lifetime but not past year (or, for cigarettes, past 3 years but not past year), and the child AFU value was missing.	The child AFU variable was assigned to 1 less than the respondent's current age.
The AFU for cigarettes was equal to 3 less than the respondent's current age, the cigarette recency was lifetime but not past 3 years, and the AFU for daily cigarettes was missing.	The AFU for daily cigarettes was assigned to 3 less than the respondent's current age.

¹ Because no recency question was associated with daily cigarettes, the overall cigarette recency was used instead.

Table B.9 Miscellaneous Edits Pertaining to Tobacco Variables

Issue	Edits Implemented
The R reported smoking cigarettes on all 30 days in the past 30 days, but the R reported initiating daily cigarette use within 30 days of the interview date. The latter could occur in one of the following ways: (a) the reported month and year for when the R started smoking cigarettes daily was within 30 days of the interview date; or (b) the R reported starting to smoke cigarettes every day at his or her current age (i.e., CIGAGE = AGE), but the R had been at this age for fewer than 30 days.	The following edits were implemented: • If only the month and year indicated that the R had been smoking daily for fewer than 30 days (i.e., but the R had been at the age given in CIGAGE for at least 30 days), the month and year (CIGDLMFU and CIGDLYFU) were set to bad data. • If the R had been at the age reported in CIGAGE for fewer than 30 days and reported starting to smoke daily in the same month that he or she was interviewed, CIGAGE, CIGDLMFU, and CIGDLYFU all were set to bad data. • If the R had been at the age reported in CIGAGE for fewer than 30 days and CIGDLMFU and CIGDLYFU had codes of DK or REF, CIGAGE, CIGDLMFU, and CIGDLYFU all were set to bad data.
The R reported smoking cigarettes in the past 30 days. The R did not report smoking on all 30 days in the past 30 days but reported initiating daily cigarette use within 30 days of the interview date.	No editing was done when this pattern occurred, even though question CG15 asked respondents if they ever smoked cigarettes every day for at least 30 days. Consequently, data were preserved that indicated that the R recently initiated daily cigarette use, even if the R could not have smoked daily for at least 30 days.
The CAI program had skipped lifetime cigarette questions pertaining to periods where the R had smoked cigarettes daily for at least 30 days (CG15) or whether the R had smoked 100 or more cigarettes (CG16A). Answers to prior questions already indicated that the R had engaged in these behaviors. For example, if the R had reported smoking every day in the past 30 days, the R had to have had a point in his or her life when he or she smoked cigarettes every day for at least a month. Similarly, it could be inferred from answers to questions about the number of days that Rs smoked in the past 30 days and the number of cigarettes they usually smoked per day that they had smoked at least 100 cigarettes in their lifetime.	It was logically inferred in CIGDLYMO (the edited version of CG15) or in CIG100LF (the edited version of CG16A) that the R had engaged in the behavior of interest (i.e., ever smoked cigarettes daily for at least 30 days or smoked 100 or more cigarettes in the R's lifetime). A code of 5 was assigned. This code had the following meaning: 5 = Yes LOGICALLY ASSIGNED (from skip pattern) This code was used instead of a legitimate skip code to signify that the CAI logic had skipped the R out of the question because the R had already provided data to indicate that the question should have been answered affirmatively.
The R was routed to the lifetime 100-cigarette question (CG16A) because the R estimated the number of days that he or she smoked cigarettes in the past 30 days. The R did not answer CG16A as "yes," but the minimum number of days that the R could have smoked in the past 30 days and the minimum usual number of cigarettes smoked per day suggest that the R smoked 100 or more cigarettes in his or her lifetime.	It was logically inferred in CIG100LF (i.e., the edited version of CG16A) that the R had smoked 100 or more cigarettes in his or her lifetime. For example, if an R reported smoking 10 to 19 days and reported smoking 16 to 25 cigarettes per day, then the R would have smoked at least 100 cigarettes in the past 30 days, even if he or she smoked on only 10 days and smoked only 16 cigarettes per day.

Table B.10 Edits Pertaining to the Source of Prescription Psychotherapeutic Drug Variables

Issue	Edits Implemented
Respondents (Rs) reported that they got the last drug they misused in "some other way," but they specified getting it in one of the ways that they had been asked about previously.	The code from the "OTHER, Specify" variable plus a value of 10 was added to the variable for how Rs obtained the drug the last time they misused it (i.e., "bump" by 10). For example, if Rs reported getting the last pain reliever they misused in some other way (PRYGOT = 8), but they specified getting them from a friend or relative for free (i.e., PNRGOTOTH = 4), then PNRGOTLAS was coded as 14. Similar edits were implemented if Rs reported that friends or relatives obtained a drug that the Rs last used in "some other way" and then the Rs specified one of the ways they had asked about previously. For example, if PNRFFGTSP was coded as 4 (i.e., the friend or relative got the pain reliever from another friend or relative for free), then PNRFFGOT (how the friend or relative obtained the last pain reliever that the R misused) was coded as 14.
Rs reported that they got their last prescription drug through a prescription, either from one doctor or from more than one doctor. However, the only way that the Rs reported misusing any prescription drug in that category in the past 12 months was without a prescription of their own (e.g., in question PRYWAYS for pain relievers).	No editing was done to the report of obtaining the last prescription drug that was misused through a doctor's prescription or to the report of use without a prescription as the only way that the R misused any prescription drug in that category in the past 12 months. However, a flag was set (e.g., PNRNORXFG) to alert analysts to the occurrence of this pattern.
Rs were logically inferred to have gotten the last medication they misused from friends or relatives for free (e.g., PNRGOTLAS = 14 for pain relievers). Consequently, Rs were not asked how the friends or relatives obtained the medication that the Rs last used.	The corresponding edited variable pertaining to how the friends or relatives obtained the medication that the Rs last used retained a code of 98 (i.e., blank).
Rs reported that they or friends/relatives got a drug for the Rs' last misuse in "some other way." However, the Rs explicitly specified that the drug they or friends/relatives obtained was an over-the-counter (OTC) drug.	No editing was done when Rs specified that they or their friends or relatives had obtained OTC drugs.

B.4 Detailed Documentation of Edits for Variables in Additional Self-Administered Sections

This section presents tables that detail the edits that were applied to variables in additional self-administered sections. Because the content of self-administered sections following the initial substance use sections differs considerably from section to section, many of the edits are specific to a single section of the interview.

- [Tables B.11](#) and [B.12](#) present edits pertaining to variables in the special drugs section, corresponding to the discussion of procedures in Section 7.2.
- [Tables B.13](#) and [B.14](#) present edits pertaining to variables in the blunts section, corresponding to the discussion of procedures in Section 8.2.
- [Tables B.15](#) and [B.16](#) present edits pertaining to variables in the prior substance use section, corresponding to the discussion of procedures in Section 8.3.
- [Tables B.17](#) to [B.19](#) present edits pertaining to variables in the substance treatment section, corresponding to the discussion of procedures in Section 8.4.
- [Tables B.20](#) and [B.21](#) present edits pertaining to variables in the consumption of alcohol section, corresponding to the discussion of procedures in Section 8.5.
- [Table B.22](#) presents edits pertaining to variables in the special topics section, corresponding to the discussion of procedures in Section 9.3.
- Finally, in this section, [Table B.23](#) presents edits pertaining to editing of variables in the adult mental health service utilization section, corresponding to the discussion of procedures in Section 10.3.

Table B.11 Specific Skip Logic Edits for Heroin Use and Needle Use Variables in the Special Drugs Section

Response Pattern	Edit
Variables were skipped because the respondent (R) never used the drug of interest, and there were no other indications elsewhere in the special drugs section that the respondent ever used this drug.	A code of 91 was assigned to the edited variables. For example, if the R never used heroin, the edited variables HERSMOKE, HRSMKREC, HERSNIFF, HRSNFREC, HERNEEDL, and HRNDLREC were assigned a code of 91.
Variables were skipped because the R refused to indicate in the corresponding substance use section whether he or she ever used the drug of interest, and there were no other indications elsewhere in the special drugs section that the respondent ever used this drug.	A code of 97 (i.e., refused) was assigned to the edited variables. Thus, for example, a refusal from the heroin recency-of-use variable was propagated onto the heroin variables in the special drugs section.
Variables were skipped because the R did not know in the corresponding substance use section whether he or she ever used the drug of interest, and there were no other indications elsewhere in the special drugs section that the R ever used this drug.	The skipped special drugs variables pertaining to this drug retained a value of 98 (i.e., blank).
For cocaine and methamphetamine, there were no indications elsewhere in the special drugs section of the R reporting use of the drug as "some other drug" that he or she used with a needle. The R was a lifetime user of the drug of interest, but the corresponding needle recency variable was skipped because of the following: • The R never used the drug with a needle. • The R refused to indicate whether he or she had ever used the drug with a needle. • The R did not know whether he or she had ever used that drug with a needle.	A code of 93 was assigned to the corresponding needle recency variable (e.g., CONDLREC) to indicate that the R used the drug but never with a needle. A code of 97 was assigned to the edited needle recency variable (i.e., the refusal was propagated). The edited needle recency variable retained a code of 98 (i.e., blank).

Table B.11 Specific Skip Logic Edits for Heroin Use and Needle Use Variables in the Special Drugs Section (continued)

Response Pattern	Edit
The R was a lifetime user of heroin, but relevant recency variables for smoking heroin (HRSMKREC), sniffing heroin (HRSNFREC), or using it with a needle (HRNDLREC) were skipped because of the following: - The R never used heroin via the route of interest. - The R refused to indicate whether he or she had ever used heroin via the route of interest. - The R did not know whether he or she had ever used heroin via the route of interest. (In the case of heroin use with a needle, there were no other indications elsewhere in the special drugs section of heroin use with a needle.)	A code of 93 was assigned to the relevant heroin recency variable(s) (e.g., HRSMKREC for smoking heroin) to indicate that the R used heroin but not in that particular way. A code of 97 was assigned to the relevant heroin recency variable(s) (i.e., the refusal was propagated). The edited heroin recency variable(s) retained a code of 98 (i.e., blank). The special situation in which Rs reported lifetime use of heroin in the heroin use section but reported that they never smoked, sniffed, or used it with a needle is discussed in the text.
General needle use variables were skipped because the R reported never using cocaine, heroin, methamphetamine, or any other drug (question SD15; edited variable OTDGNEDL) with a needle.	A code of 99 (i.e., legitimate skip) was assigned to all of the general needle use variables that had been skipped (GNNDREUS, GNNDLSH1, GNNDCLN, GNNDLSH2, and GNNDGET).
General needle use variables were skipped because question SD15 was answered as "no" (OTDGNEDL = 2); there were no affirmative reports of the use of cocaine, heroin, or methamphetamine with a needle, but one or more of the lifetime needle use variables for these drugs was answered as "don't know" or "refused."	When there was no affirmative report of the use of cocaine, heroin, or methamphetamine, with a needle, question SD15 was worded as follows: "Have you ever, even once, used a needle to inject <i>any drug</i> that was not prescribed for you?" (wording not italicized in the interview). Therefore, a code of 99 (i.e., legitimate skip) was assigned to all of the general needle use variables that had been skipped (GNNDREUS, GNNDLSH1, GNNDCLN, GNNDLSH2, and GNNDGET) because it could be inferred that the response of "no" in question SD15 pertained to use of any drug with a needle. However, no editing was done to any responses of "don't know" or "refused" in the lifetime needle use variables pertaining to cocaine (COCNEEDL), heroin (HERNEEDL), or methamphetamine (METHNEEDL).
General needle use variables were skipped because question SD15 was refused (OTDGNEDL = 97), and there were no affirmative reports of heroin, methamphetamine, other stimulants, or cocaine with a needle.	The refusal from OTDGNEDL was propagated to the general needle use variables GNNDREUS, GNNDLSH1, GNNDCLN, GNNDLSH2, and GNNDGET.

Table B.12 Edits in the Special Drugs Section (Other than Skip Patterns) Pertaining to Heroin Use and Needle Use

Issue	Edits Implemented
The respondent (R) reported injecting other drugs in question SD15, and the R specified injecting only cocaine, only heroin, or only methamphetamine in the "OTHER, Specify" questions SD15SP to SD15D. However, the R reported more recent use of the "other" drug with a needle in question SD16 than was reported in the corresponding needle recency question for cocaine, heroin, or methamphetamine.	The value from SD16 (edited variable OTDGNDLRC) plus a value of 10 (i.e., "bump" by 10) was assigned to the needle recency variable for the specific drug. For example, if the heroin recency indicated that the R last used heroin with a needle more than 12 months ago (HRNDLREC = 3), the R reported the injection of only heroin as some other in his or her lifetime, and the most recent use of "other" drugs with a needle was more than 30 days ago but within the past 12 months (OTDGNDLRC = 2), then HRNDLREC was set to 12 (More than 30 days ago but within the past 12 months LOGICALLY ASSIGNED). This edit was not implemented if the R had responses of "don't know" or "refused" in the "OTHER, Specify" questions SD15SP to SD15D because these responses potentially indicated the injection of drugs other than cocaine, heroin, or methamphetamine.
The R reported injecting other drugs in question SD15, and the R specified injecting only cocaine, only heroin, or only methamphetamine in the "OTHER, Specify" questions SD15SP to SD15D. However, the R reported more recent use with a needle in the corresponding needle recency question for cocaine, heroin, or methamphetamine than the R reported in the needle recency question SD16 for other drugs.	The value from the specific needle recency variable (i.e., CONDLREC for cocaine, HRNDLREC for heroin, or METHNDLRC for methamphetamine) plus a value of 10 (i.e., "bump" by 10) was assigned to the needle recency variable OTDGNDLRC (corresponding to question SD16) for other drugs. For example, if the heroin recency indicated that the R last used heroin with a needle more than 30 days ago but within the past 12 months (HRNDLREC = 2), the R reported the injection of only heroin as some other in his or her lifetime, and the most recent use of "other" drugs with a needle was more than 12 months ago (OTDGNDLRC = 3), then OTDGNDLRC was set to 12 (More than 30 days ago but within the past 12 months LOGICALLY ASSIGNED). This edit was not implemented if the R had responses of "don't know" or "refused" in the "OTHER, Specify" questions SD15SP to SD15D because these responses potentially indicated the injection of drugs other than cocaine, heroin, or methamphetamine.

Table B.12 Edits in the Special Drugs Section (Other than Skip Patterns) Pertaining to Heroin Use and Needle Use (continued)

Issue	Edits Implemented
The R was a lifetime user of heroin and reported smoking, sniffing, or using heroin with a needle at least once in his or her lifetime, but did not know or refused to indicate when he or she last smoked, sniffed, or injected heroin.	The edits typically depended on the most recent use of heroin reported in the corresponding heroin recency variable, except where noted. • In general, the edited heroin recency variables in the special drugs section (HRSMKREC, HRSNFREC, HRNDLREC) were assigned a code of 9 (i.e., used at some point in the lifetime). Exceptions are noted below. • If the only "other" drug that the R injected in his or her lifetime in SD15SP to SD15D was heroin and the needle recency for other drugs (OTDGNDLRC, corresponding to question SD16) indicated last use in a definite period, then the value from OTDGNDLRC plus a value of 10 (i.e., "bump" by 10) was assigned to HRNDLREC. In this situation, for example, if OTDGNDLRC = 2 (More than 30 days ago but within the past 12 months), then HRNDLREC was set to 12 (More than 30 days ago but within the past 12 months LOGICALLY ASSIGNED). • If the heroin recency indicated that the R last used heroin more than 12 months ago and there was no other indication in the special drugs section that the R had used heroin more recently, then the edited variables pertaining to smoking, sniffing, or injection of heroin were assigned a code of 13 (i.e., More than 12 months ago LOGICALLY ASSIGNED). This edit did not apply if the R answered "did not know" or refused to report when he or she last used heroin in a particular way (e.g., smoking it) but reported last using it a different way in the past 12 months (e.g., with a needle).
The R was a lifetime user of cocaine and reported using it with a needle at least once in his or her lifetime. However, the R did not know or refused to indicate when he or she last used cocaine with a needle.	The edits typically depended on the most recent use reported in the corresponding recency variable COCREC, except where noted. • In general, the edited needle recency variable CONDLREC for cocaine was assigned a code of 9 (i.e., used at some point in the lifetime). Exceptions are noted below. • If the only "other" drug that the R injected in his or her lifetime in SD15SP to SD15D was cocaine and the needle recency for other drugs (OTDGNDLRC, corresponding to question SD16) indicated last use in a definite period, then the value from OTDGNDLRC plus a value of 10 (i.e., "bump" by 10) was assigned to CONDLREC. In this situation, for example, if OTDGNDLRC = 2 (More than 30 days ago but within the past 12 months), then CONDLREC was set to 12 (More than 30 days ago but within the past 12 months LOGICALLY ASSIGNED). • If COCREC indicated that the R last used cocaine more than 12 months ago and there was no other information to indicate unambiguously that the R used cocaine with a needle more recently, then CONDLREC was assigned a code of 13 (i.e., More than 12 months ago LOGICALLY ASSIGNED).

Table B.12 Edits in the Special Drugs Section (Other than Skip Patterns) Pertaining to Heroin Use and Needle Use (continued)

Issue	Edits Implemented
The R was a lifetime user of methamphetamine and reported using it with a needle at least once in his or her lifetime. However, the R did not know or refused to indicate when he or she last used methamphetamine with a needle.	The edits typically depended on the most recent use reported in the corresponding recency variable METHAMREC, except where noted. • In general, the edited needle recency variable METHNDLRC for methamphetamine was assigned a code of 9 (i.e., used at some point in the lifetime). Exceptions are noted below. • If the only "other" drug that the R injected in his or her lifetime in SD15SP to SD15D was methamphetamine and the needle recency for other drugs (OTDGNDLRC, corresponding to question SD16) indicated last use in a definite period, then the value from OTDGNDLRC plus a value of 10 (i.e., "bump" by 10) was assigned to METHNDLRC. In this situation, for example, if OTDGNDLRC = 2 (More than 30 days ago but within the past 12 months), then METHNDLRC was set to 12 (More than 30 days ago but within the past 12 months LOGICALLY ASSIGNED). • If METHAMREC indicated that the R last used methamphetamine more than 12 months ago and there was no other information to indicate unambiguously that the R used methamphetamine with a needle more recently, then METHNDLRC was assigned a code of 13 (i.e., More than 12 months ago LOGICALLY ASSIGNED).
The R reported in the corresponding substance use sections that he or she never used heroin (HERREC = 91), never used cocaine (COCREC = 91), or never used methamphetamine (METHAMREC = 91). However, the R specified use of one or more of these drugs as "some other drug" that he or she had ever injected.	No editing was done to the substance use data that appeared earlier in the interview. However, the R was logically inferred in the special drugs data to be a lifetime user of that drug with a needle, even though the earlier substance use data indicated that the R never used that drug. If the R reported the injection of more than one other drug in his or her lifetime, then the corresponding needle recency variable was set to a value of 9 (Used at some point in the lifetime LOGICALLY ASSIGNED). For example, if the R reported in the heroin section that he or she never used heroin but specified injection of heroin as "some other drug," the lifetime heroin needle use variable HERNEEDL was set to 3 (Yes LOGICALLY ASSIGNED), and the heroin needle recency HRNDLREC was set to 9. Similar edits were done for the needle use variables pertaining to cocaine and methamphetamine. However, the edits that were described previously applied if the R reported the injection of only heroin, only cocaine, or only methamphetamine in the "OTHER, Specify" questions SD15SP to SD15D. For example, if HERREC = 91, the R specified the injection of only heroin in SD15SP to SD15D, and the R reported last injecting other drugs with a needle more than 30 days ago but within the past 12 months (OTDGNDLRC = 2), then HERNEEDL was set to 3, as described above. However, HRNDLREC was set to 12 (More than 30 days ago but within the past 12 months LOGICALLY ASSIGNED) instead of being set to 9.

Table B.12 Edits in the Special Drugs Section (Other than Skip Patterns) Pertaining to Heroin Use and Needle Use (continued)

Issue	Edits Implemented
The R specified lifetime use of more than five drugs with a needle.	Priority was given to retaining as many unique mentions as possible for other drugs that the R used with a needle. Thus, multiple mentions of the same drug would be counted only once. Priority also would be given to retaining mentions of drugs that were covered in the special drugs section that the R had not previously reported using with a needle (e.g., if the question corresponding to METHNEEDL had been answered as "no" but methamphetamine had been specified as "some other drug" that the R used with a needle). Conversely, retention of "OTHER, Specify" codes corresponding to drugs that the R had already reported using with a needle were given lower priority. If there were still more than five mentions of unique drugs after the above steps, priority was given to retaining the most serious drugs according to the Drug Enforcement Administration (DEA) drug schedule (e.g., first priority to retaining mention of Schedule I drugs that have no approved medical use in the United States, second priority to retaining Schedule II drugs). Finally, after the drugs had been ranked according to their severity based on the DEA drug schedule, if mention of more than five drugs still remained, the codes were retained in the order they appeared in the data.
The R reported using a needle to inject a drug (SD05 = 1), but the R previously reported never using marijuana, cocaine, heroin, hallucinogens, inhalants, methamphetamine, prescription pain relievers, prescription tranquilizers, prescription stimulants, or prescription sedatives.	No editing was done if the R - specified needle use involving a drug that could be misused or that had psychoactive properties (e.g., steroids, one or more categories of drugs covered in the earlier substance use sections that were not covered elsewhere in special drugs, such as prescription pain relievers), or - reported one or more "risky" needle use behaviors (reusing a needle, needle sharing, or cleaning a needle with bleach). The R was inferred not to be a lifetime nonmedical needle user (OTDGNEDL = 4) if the R - specified use of a drug that was typically not misused and did not have psychoactive properties (e.g., if injection of antibiotics was specified), and - reported never reusing a needle, sharing a needle (before or after someone had used it), or cleaning a needle with bleach (i.e., "risky" needle use behaviors). When OTDGNEDL = 4, any data in the general needle use variables GNNDREUS, GNNDLSH1, GNNDCLN, GNNDLSH2, and GNNDGET were replaced with a code of 89 (LEGITIMATE SKIP Logically assigned). Note, however, that this edit was not implemented in 2016 for a small number of respondents (fewer than five) who had patterns that would have qualified OTDGNEDL to be set to 4. This issue was corrected for the 2016 NSDUH editing.

Table B.12 Edits in the Special Drugs Section (Other than Skip Patterns) Pertaining to Heroin Use and Needle Use (continued)

Issue	Edits Implemented
The general needle use items corresponding to questions SD12 through SD16 had been answered. However, the only indications of lifetime needle use had been set to bad data because of patterned responses in the corresponding substance use section.	The edited general needle use variables GNNDREUS through GNNDGET were set to bad data.
The R reported getting his or her last needle "some other way" and specified a meaningful response for how he or she last got the needle.	The final, edited variable pertaining to how the R got his or her last needle (GNNDGET) was a composite of the response categories that were offered to the R (i.e., bought the needle from a pharmacy, got the needle from a needle exchange, bought the needle on the street, got the needle in a shooting gallery, got the needle some other way). This was done because the computer-assisted interviewing (CAI) logic did not allow Rs to specify an "other" way that they got the needle if they reported getting the needle in one of the first four ways. If Rs reported getting the needle "some other way" and specified a meaningful way they got the needle, then the "OTHER, Specify" response was assigned to GNNDGET.
The R reported getting his or her last needle "some other way" and did not know what that other way was, refused to specify what that other way was, or gave a response that was coded as bad data (e.g., a nonsensical response).	The final, edited variable pertaining to how the R got his or her last needle (GNNDGET) retained a nonspecific code of "some other way." Stated another way, the response of "some other way" was given precedence over the missing value in the "OTHER, Specify" response. The edit was done in this manner to provide a nonmissing value for analysts to use.
The R answered "don't know" or "refused" at the outset, when asked how he or she got the last needle that he or she used.	The response of "don't know" or "refused" was retained in the final, edited variable (GNNDGET).

Table B.13 Edits in the Blunts Section Pertaining to Consistency Checks for Initiation

Issue	Edits Implemented
The respondent (R) indicated in the final verification check (BLCC06) that the age at first use (AFU) based on the new month and year of first use (MFU and YFU) was correct (i.e., BLCC06 = 4). The computer-assisted interviewing (CAI) program updated the AFU (AGE1STBL) with the value of the age calculated from the MFU and YFU (i.e., MYR1STBL). However, the new value for the AFU indicated that the R was more than 1 year younger than his or her current age at the time the R first used blunts (e.g., the R was 16, reported first use of blunts at age 15, but then confirmed an MFU and YFU that meant the R was 14 when the R first used blunts). Had the R initially reported this AFU, the R would not have been routed to the MFU and YFU questions.	The updated value was retained in the edited AFU variable BLNTAGE (e.g., first use at age 14 for a 16-year-old R in this example). Based on this updated AFU, it was logically inferred that the R should have skipped the MFU and YFU items. A code of 9989 was assigned to the YFU variable BLNTYFU, and a code of 89 was assigned to the MFU variable BLNTMFU.
The final verification check (BLCC06) was skipped because the R entered revised data for the MFU and YFU that made BLNTMFU and BLNTYFU consistent with the AFU in BLNTAGE.	No editing was done because the R was considered to have resolved the inconsistency.
The final verification check (BLCC06) was skipped because the R entered a new MFU that was the same as the R's birth month.	The new MFU could be consistent with the AFU, depending on whether the use in that month occurred before or after the R's birthday. No editing was done to the AFU, MFU, and YFU, as long as the revised MFU and YFU were potentially consistent with the AFU. However, BLNTMFU and BLNTYFU were set to bad data if they could never be consistent with the AFU. Suppose, for example, that a hypothetical R was born in June 1998, was interviewed in March 2016 (age 17 at the time of the interview), reported first use of blunts at age 17, and initially reported first use in May 2015. First use in May 2015 would have meant that the R was 16 when he or she first used blunts because the R's 17th birthday was not until June 2015. If the R changed the month and year to June 2015, that could be consistent with first use at age 17, if the use occurred after the R's birthday. However, if the R changed the month and year to June 2014, it would never be possible for the R to have first used at age 17 and also to have first used in June 2014. In this latter situation, BLNTMFU and BLNTYFU would be set to bad data.
The R entered a new MFU or YFU that differed from what the R previously reported. The age based on the revised MFU and YFU (updated in MYR1STBL) still mismatched the AFU, but the R indicated in the final verification check that the new value from MYR1STBL was correct.	No editing was done in this situation. The CAI program automatically updated AGE1STBL to be consistent with the updated values reported for the MFU and YFU.

Table B.13 Edits in the Blunts Section Pertaining to Consistency Checks for Initiation (continued)

Issue	Edits Implemented
The consistency check was triggered between AGE1STBL and MYR1STBL. However, the R answered the first consistency check (BLCC03, regarding whether the value in MYR1STBL was correct) as "don't know" or "refused." The R then exited the consistency check loop without having resolved the inconsistency.	The AFU from AGE1STBL was retained in BLNTAGE, but BLNTMFU and BLNTYFU were set to bad data.
The consistency check was triggered between AGE1STBL and MYR1STBL. However, the R entered the same values for the YFU and MFU that triggered the inconsistency with the AFU in the first place.	No editing was done to BLNTAGE, but BLNTYFU and BLNTMFU were set to bad data.
The R entered new values in the consistency checks for the MFU and YFU that again yielded a nonmissing value for MYR1STBL. However, the R failed to resolve the inconsistency between AGE1STBL and the updated value in MYR1STBL. The R also reported either in the first verification check that MYR1STBL was not correct (i.e., BLCC03 = 6) or reported in the second verification check that the AFU from AGE1STBL was correct (i.e., BLCC04 = 2).	No editing was done to the AFU. The following edits were implemented for the MFU and YFU: • The default edit was to set BLNTMFU and BLNTYFU to bad data. • As an exception to the default edit, if the final verification check was answered as "don't know" or "refused," BLNTMFU and BLNTYFU were assigned the code that corresponded to the answer in the final verification check.
The R entered new values in the consistency checks for the MFU and YFU that again yielded a nonmissing value for MYR1STBL. However, the R failed to resolve the inconsistency between AGE1STBL and the updated value in MYR1STBL. The R also reported in the second verification check that neither answer was correct for what was originally captured in AGE1STBL and MYR1STBL (i.e., BLCC04 = 3).	The following edits were implemented for BLNTAGE, BLNTMFU, and BLNTYFU: • The default edit was to set BLNTAGE, BLNTMFU, and BLNTYFU to bad data. • As an exception to the default edit, if the final verification check was answered as "don't know" or "refused," BLNTAGE, BLNTMFU, and BLNTYFU were assigned the code that corresponded to the answer in the final verification check.
The R triggered the initial consistency check between the AFU, MFU, and YFU and reported that the age from MYR1STBL was correct (i.e., BLCC03 = 4). However, the R answered the second consistency check (BLCC04) as "don't know" or "refused." Consequently, the R did not have an opportunity to correct the inconsistency between the AFU, MFU, and YFU.	The relevant codes for "don't know" or "refused" were assigned to BLNTAGE, BLNTMFU, and BLNTYFU. The rationale for this edit is that conclusive information did not exist regarding whether the AFU indicated the R's correct age when he or she first used blunts, or whether the MFU and YFU indicated the R's correct age at initiation. Therefore, BLNTAGE, BLNTMFU, and BLNTYFU all were set to missing values.

Table B.14 Edits in the Blunts Section Other than Those Pertaining to Consistency Checks for Initiation

Issue	Edits Implemented
The respondent (R) reported lifetime use of blunts but did not know or refused to report in BL02 when he or she last used them.	The blunt recency variable BLNTREC (corresponding to question BL02) was set to 9 (Used at some point in the lifetime LOGICALLY ASSIGNED). The 30-day frequency BLNT30DY was left as blank.
The R reported using blunts (BLNTEVER = 1, corresponding to question BL01), but the R previously reported never using marijuana (MJEVER = 2). The consistency check BLNT30C1 (corresponding to question BL03) was triggered because the R reported using blunts in the past 30 days in question BL02 and the R never used marijuana. BLNT30C1 was not triggered for less recent reports of blunt use.	• If BLNT30C1 = 1 (i.e., R has verified that he or she has never used marijuana): – BLNTEVER was set to 4 (No LOGICALLY ASSIGNED). – BLNTREC was set to 81 (NEVER USED BLUNTS Logically assigned). – BLNT30DY was to 91 (if blank) or 81 (if a value needed to be overwritten). – Initiation variables (BLNTAGE, BLNTYFU, BLNTMFU) were assigned codes of 981, 9981, and 81, respectively. – Other assignment of legitimate skip codes applies, as described in the main text. • If BLNT30C1 = 2 (i.e., R has verified past month use of blunts), no editing was done to BLNTEVER or BLNTREC. Also, these data in blunts were not used to edit the MJEVER variable for lifetime marijuana use (see Section 2.4.1).
The R reported using blunts, but the R previously reported never using marijuana. BLNT30C1 was not triggered because the R reported use of blunts that was less recent than the past 30 days.	• BLNTEVER was set to a value of 11 (i.e., "bump" by 10) to signify that the R reported lifetime use of cigars with marijuana in them but reported never using marijuana in the marijuana section, and the R was not asked to verify which answer was correct. • No editing was done to the recency values in BLNTREC.
The R reported using blunts in the past 30 days, and the R is a marijuana user. However, the R reported last using marijuana more than 30 days ago but within the past 12 months or more than 12 months ago. The consistency check BLNT30C2 (corresponding to question BL04) was triggered for these cases.	• If BLNT30C2 = 1 (i.e., R has verified that the marijuana recency is correct): – BLNTREC was set to 19 (More than 30 days ago LOGICALLY ASSIGNED). – BLNT30DY was set to 93 if blank and to 83 (DID NOT USE BLUNTS IN THE PAST 30 DAYS Logically assigned) if the 30-day frequency data needed to be overwritten. – Note, however, that no editing was done to BLNTREC when BLNT30C2 = 1, but the marijuana recency MJREC had been set to a value of 8 or 9 (Used at some point in the past 12 months or Used at some point in the lifetime, respectively), based on edits in the marijuana section (see Section 6.2.2.3 in Chapter 6). The rationale for doing no editing to BLNTREC was that values of 8 or 9 in MJREC indicated potential use of marijuana in the past month. • If BLNT30C2 = 2 (i.e., R has verified past month use of blunts), no editing was done to BLNTEVER or BLNTREC. Also, these data in blunts were not used to edit the marijuana recency MJREC (see Section 2.4.1).

Table B.14 Edits in the Blunts Section Other than Those Pertaining to Consistency Checks for Initiation (continued)

Issue	Edits Implemented
The R reported last using blunts more than 30 days ago but within the past 12 months (BLNTREC = 2), and the R reported last using marijuana more than 12 months ago (MJREC = 3).	BLNT30C2 (BL04) was not triggered in this situation because the R did not report use of blunts in the past 30 days. • BLNTREC was set to a value of 12 (i.e., "bump" by 10), to signify that the R reported last using blunts more than 30 days ago but within the past 12 months but reported last using marijuana more than 12 months ago, and the R was not asked to verify which answer was correct. • The editing procedures continued to set the 30-day frequency BLNT30DY to 93 because the R was not a past month user of blunts.
The R reported last using blunts more than 30 days ago but within the past 12 months (BLNTREC = 2), but the marijuana recency MJREC was assigned codes of 8 or 9 (Used at some point in the past 12 months or lifetime, respectively).	No editing was done to the blunts variables when MJREC had values of 8 or 9 because BLNTREC = 2 was not necessarily inconsistent with when the R last used marijuana.
Question BL07 (edited variable BLNTNOMJ) was skipped. The R reported in the tobacco section that he or she never smoked a cigar and reported in the blunts section that he or she never smoked a blunt.	BLNTNOMJ was set to a value of 91 (NEVER USED CIGARS/BLUNTS). This included situations in which BLNTEVER had been set to 4 (No LOGICALLY ASSIGNED) and BLNTREC to 81 (NEVER USED BLUNTS Logically assigned) because the R confirmed in BLNT30C2 that he or she never used marijuana.
BLNTNOMJ was skipped, and the R reported lifetime use of blunts but definitely did not smoke blunts in the past 30 days. The R also reported in the tobacco section that he or she did not smoke a cigar in the past 30 days (including situations in which the R reported never smoking a cigar).	BLNTNOMJ was set to value of 93 (DID NOT USE CIGARS/BLUNTS IN THE PAST 30 DAYS). This edit also was implemented if the R reported use of blunts less recently than the past 30 days but reported in the tobacco section that he or she never smoked a cigar. Even though the cigar and blunts data were inconsistent, BLNTNOMJ still was not applicable.
BLNTNOMJ was skipped, and the R reported lifetime use of blunts but definitely did not smoke blunts in the past 30 days. Unlike the situation above, the R reported in the tobacco section that he or she smoked cigars in the past 30 days.	BLNTNOMJ was set to a value of 5, where 5 = Yes LOGICALLY ASSIGNED (from skip pattern). For instance, if the R has reported last smoking cigars with marijuana in them more than 30 days ago, it could logically be inferred that all past month cigar use had to involve use of cigars that did not have marijuana in them.
BLNTNOMJ was skipped, even though the R reported use of blunts in the past 30 days (BLNTREC = 1). However, the R had reported lifetime use of cigars in the tobacco section but reported not smoking cigars in the past 30 days.	Rather than assign a code of 99 (LEGITIMATE SKIP), BLNTNOMJ was set to a value of 14, where 14 = USED BLUNTS PAST 30 DAYS/DIDN'T USE CIGARS PAST 30 DAYS.
BLNTNOMJ was skipped, even though the R reported use of blunts in the past 30 days (BLNTREC = 1). However, the R had reported in the tobacco section that he or she had never smoked part or all of a cigar.	Rather than assign a code of 99, BLNTNOMJ was set to a value of 24, where 24 = USED BLUNTS PAST 30 DAYS/NEVER USED CIGARS.

Table B.14 Edits in the Blunts Section Other than Those Pertaining to Consistency Checks for Initiation (continued)

Issue	Edits Implemented
BLNTNOMJ was skipped because the cigar recency variable CIGARREC or the blunts recency variable BLNTREC had missing values or "indefinite" values (i.e., used at some point in the past 12 months, or used at some point in the lifetime).	BLNTNOMJ was left as 98 (blank).
Question BL07 (corresponding to BLNTNOMJ) was answered but was not answered as "yes," and the R had not smoked blunts in the past 30 days.	BLNTNOMJ was set to a value of 3, where 3 = Yes LOGICALLY ASSIGNED. This edit affected cases where BL07 had been answered (but not as "yes"), and the Rs confirmed in BLNT30C2 that they did not use marijuana in the past 30 days.
The age at first use (AFU) was inconsistent with the R's current age despite the R being prompted to change the AFU.	The final age was accepted as the standard, and BLNTAGE was set to bad data.
BLNTAGE had a value of "don't know" (DK) or "refused" (REF), including situations where this assignment has been made from the consistency check data.	If the questions pertaining to BLNTYFU and BLNTMFU were skipped because the R answered the AFU as DK or REF, the DK or REF value from BLNTAGE was propagated onto BLNTYFU and BLNTMFU. This edit was designed to indicate the reason that BLNTYFU and BLNTMFU had been skipped. In addition, because the R may have first used blunts within 1 year of his or her current age, the month- and year-of-first-use (MFU and YFU) questions may have been relevant to the R. If the R had answered the MFU and YFU questions but BLNTAGE had a final code for DK or REF (i.e., due to a consistency check response), data in BLNTYFU and BLNTMFU were overwritten with the corresponding DK or REF value from BLNTAGE. Retaining the MFU and YFU data in this situation would imply that the R first smoked blunts within 1 year of his or her current age.
The MFU has been skipped because the R answered the YFU as DK or REF.	The DK or REF value from BLNTYFU was propagated onto BLNTMFU. That is, if the R did not know in what year he or she first used blunts, it was assumed that the R would not know the month either. Similarly, a refusal to answer the YFU was interpreted to be a blanket refusal to answer the month as well as the year.
The R had missing data for the MFU. However, one of the following occurred: • The YFU was the current calendar year, and the R was interviewed in January. • The R first used at his or her current age, the R first used in the current calendar year, and the R's most recent birthday occurred in the interview month. • The R first used at his or her current age, the R first used in the prior calendar year, and the R's most recent birthday occurred in December.	The MFU could be logically inferred, as indicated below. • The MFU was logically inferred to be January of the interview year. That would be the only month in which the R could have initiated use of blunts in the current year. • The MFU was logically inferred to be the interview month. If the R first used in the current calendar year and attained his or her current age in the interview month, the R logically could not have initiated use of blunts in any month other than the interview month. • The MFU was logically inferred to be December. If the R first used blunts in the prior calendar year and attained his or her current age in December of that year, the R logically could not have initiated use in any month other than December of the prior calendar year.

Table B.15 Edits in the Prior Substance Use Section Pertaining to Consistency Checks for Age, Year, and Month of Last Use Variables

Issue	Edits Implemented
The respondent (R) indicated in the final verification check (e.g., LUCG08 for cigarettes) that the age of last use (ALU) based on the new month and year of last use (MLU and YLU) was correct (e.g., LUCG08 = 4). The computer-assisted interviewing (CAI) program updated the ALU (such as AGELSTCG for cigarettes) with the value of the age calculated from the MLU and YLU (referred to subsequently as the MYRLST age, such as MYRLSTCG, for cigarettes). However, the new value for the ALU indicated that the R was more than 1 year younger than his or her current age at the time the R last used the drug (e.g., the R was 16, reported last use of the drug at age 15, but then confirmed an MLU and YLU that meant the R was 14 when the R last used the drug). Had the R initially reported this ALU, the R would not have been routed to the MLU and YLU questions.	The updated value was retained for the ALU (e.g., last use at age 14 for a 16-year-old R in this example). Based on this updated ALU, it was logically inferred that the R should have skipped the MLU and YLU items. A code of 9989 was assigned to the YLU variable (e.g., CIGYLU for cigarettes), and a code of 89 was assigned to the MLU variable (e.g., CIGMLU).
The final verification check (e.g., LUCG08 for cigarettes) was skipped because the R entered revised data for the MLU and YLU that made them consistent with the ALU.	No editing was done because the R was considered to have resolved the inconsistency.
The final verification check (e.g., LUCG08) was skipped because the R entered a new MLU that was the same as the R's birth month.	The new MLU could be consistent with the ALU, depending on whether the use in that month occurred before or after the R's birthday. No editing was done to the ALU, MLU, and YLU, as long as the revised MLU and YLU were potentially consistent with the ALU. However, the MLU and YLU were set to bad data if they could never be consistent with the ALU. Suppose, for example, that a hypothetical R was born in June 1998, was interviewed in March 2016 (age 17 at the time of the interview), reported last use of cigarettes at age 17, and initially reported last use in May 2015. Last use in May 2015 would have meant that the R was 16 when he or she last smoked cigarettes because the R's 17th birthday was not until June 2015. If the R changed the month and year to June 2015, that could be consistent with last use at age 17, if the use occurred after the R's birthday. However, if the R changed the month and year to June 2014, it would never be possible for the R to have last used at age 17 and also to have last used in June 2014. In this latter situation, CIGMLU and CIGYLU would be set to bad data.
The consistency check was triggered between the ALU and the MYRLST age. However, the R answered the first consistency check (e.g., LUCG05 for cigarettes, regarding whether the value in the MYRLST age was correct) as "don't know" or "refused." The R then exited the consistency check loop without having resolved the inconsistency.	The value for the ALU was retained, but the MLU and YLU were set to bad data.

Table B.15 Edits in the Prior Substance Use Section Pertaining to Consistency Checks for Age, Year, and Month of Last Use Variables (continued)

Issue	Edits Implemented
The final verification check (e.g., LUCG08 for cigarettes) was skipped because the R entered the same values for the MLU and YLU that triggered the inconsistency with the ALU in the first place.	The value for the ALU was retained, but the MLU and YLU were set to bad data.
The R entered a new MLU or YLU that differed from what the R previously reported. The MYRLST age based on the revised MLU and YLU still mismatched the ALU, but the R indicated in the final verification check that this MYRLST age was correct.	No editing was done in this situation. The CAI program automatically updated the ALU to be consistent with the updated values reported for the MLU and YLU.
The consistency check was triggered between the ALU and the MYRLST age calculated from the MLU and YLU. However, the R answered the first consistency check (e.g., LUCG05 for cigarettes) as "don't know" or "refused." The R then exited the consistency check loop without having resolved the inconsistency.	The ALU was retained, but the MLU and YLU were set to bad data.
The R entered values in the consistency checks for the MLU and YLU that again yielded a nonmissing MYRLST age based on the MLU and YLU. However, the R failed to resolve the inconsistency between the ALU and the MYRLST age. The R also reported either in the first verification check that the MYRLST age was not correct (e.g., LUCG05 = 6 for cigarettes) or reported in the second verification check that the ALU was correct (e.g., LUCG06 = 2 for cigarettes).	No editing was done to the ALU. The following edits were implemented for the MLU and YLU: The default edit was to set the MLU and the YLU to bad data. As an exception to the default edit, if the final verification check was answered as "don't know" or "refused," the MLU and YLU were assigned the code that corresponded to the answer in the final verification check.
The R entered values in the consistency checks for the MLU and YLU that again yielded a nonmissing MYRLST age based on the MLU and YLU. However, the R failed to resolve the inconsistency between the ALU and the MYRLST age. The R also reported in the second verification check that neither the ALU nor original MYRLST age was correct (e.g., LUCG06 = 3 for cigarettes).	The following edits were implemented for the ALU, MLU, and YLU: The default edit was to set the ALU, MLU, and YLU to bad data. As an exception to the default edit, if the final verification check was answered as "don't know" or "refused," the ALU, MLU, and YLU were assigned the code that corresponded to the answer in the final verification check.
The R triggered the initial consistency check between the ALU, MLU, and YLU and reported that the MYRLST age calculated from the MLU and YLU was correct (e.g., LUCG05 = 4). However, the R answered the second consistency check (e.g., LUCG06 for cigarettes) as "don't know" or "refused." Consequently, the R did not have an opportunity to correct the inconsistency between the ALU, MLU, and YLU.	The relevant codes for "don't know" or "refused" were assigned to the ALU, MLU, and YLU. The rationale for this edit is that conclusive information did not exist regarding whether the ALU indicated the R's correct age when he or she last used a drug or whether the MLU and YLU indicated the R's correct ALU. Therefore, the ALU, MLU, and YLU all were set to missing values.

Table B.16 Edits in the Prior Substance Use Section Pertaining to the Age, Year, and Month of Last Use (Other than for Consistency Checks)

Issue	Edits Implemented
The age of last use (ALU) for a given drug or behavior (MRJAGLST for marijuana, CIGDLLST for daily cigarette smoking) was greater than the respondent's (R's) current age.	The ALU was set to bad data.
The ALU was lower than the age at first use (AFU) from the corresponding earlier substance use section (e.g., MJAGE for marijuana). If the ALU for drugs, cigarettes, or daily cigarette use was lower than the corresponding AFU, a consistency check was triggered. The Rs were not allowed to go back and change the AFU from the earlier substance use section; the only way they could resolve the inconsistency was by making the ALU consistent with the AFU. Thus, the Rs may indicate that the ALU was correct as reported, or they could enter a new ALU that still was inconsistent with the AFU.	The ALU was set to bad data, even if the R reported in the consistency check that "yes," what the R had previously reported for the ALU was correct. Thus, as a <i>partial</i> exception to the principle of not editing across sections (Section 2.4.1), AFU data to edit the ALU in the prior substance use section, but not vice versa.
The AFU from the corresponding earlier substance use section (e.g., MJAGE for marijuana) had a missing value, but the ALU (e.g., MRJAGLST) was defined. That included situations in which the AFU was set to bad data as part of the "flag and impute" edits (see Section 6.2.2.4).	No editing was done to the ALU, as long as it was consistent with the R's age (see above). In addition, no editing was done to the ALU if the imputed AFU (e.g., IRMJAGE for marijuana) was given a value greater than the ALU. (Imputation of the AFU variables in 2016 did not take into account data from the corresponding ALU variables.)
The edited recency from the earlier substance use section (e.g., CIGREC for cigarettes) indicated that the last use was more than 12 months ago, but the ALU indicated last use at the R's current age. The answer to the ALU would suggest use in the past 12 months. (For cigarettes, either CIGAGLST or CIGDLLST could be inconsistent with the recency.)	No editing was done to the ALU. However, a standard codebook footnote was included for these variables to alert analysts to the fact that values in this section could be inconsistent with values from earlier sections of the interview.
The recency from the earlier substance use section was set to an "indefinite" value of 8 or 9 (used at some point in the past 12 months or used at some point in the lifetime, respectively). These indefinite recency values could be imputed to past month use. If the unedited recency in the corresponding substance use section indicated use more than 30 days ago, the R would be routed to the prior substance use section and could provide data for the ALU. The latter would suggest that the R is not a past month user, even though the final imputed recency might indicate that the R is a past month user.	No editing was done to the ALU. The standard codebook footnote described above also applied to this issue.

Table B.16 Edits in the Prior Substance Use Section Pertaining to the Age, Year, and Month of Last Use (Other than for Consistency Checks) (continued)

Issue	Edits Implemented
For Rs who last used drugs more than 30 days ago but within the past 12 months, the number of days that the R could have used a drug (e.g., marijuana) in the past 12 months based on MRJAGLST was less than the number of days in the edited or imputed 12-month frequency for marijuana (MJYRTOT or IRMJFY, respectively).	No editing was done to MRJAGLST. The standard codebook footnote described above also applied to this issue.
The ALU variable had a missing value, but the corresponding AFU indicated that the R first used (or started smoking cigarettes daily) at his or her current age.	• The ALU was set to equal the AFU. Logically, last use cannot be any later than the R's current age. • As part of this same edit, if the year of first use (YFU) indicated that the R first used in the year of the interview, then the R was logically inferred to have last used in the current year.
The year of last use (YLU) had a missing value. In addition, the R reported first using a drug (or first smoking cigarettes daily) at an age that was 1 year younger than his or her current age. The R also reported first using the drug (or first smoking cigarettes daily) in the current year.	No editing was done to the ALU, but the YLU was set to equal the YFU. Logically, if the first use was in the current year, then the last use had to be in the current year as well.
The month of last use (MLU) had a missing value, but the corresponding month of first use (MFU, or first daily cigarette use) was in the calendar month prior to in the interview month.	The MLU was set to equal the MFU. Logically, if the R was routed to the prior substance use section because he or she had not used a given drug in the past 30 days, the R also had to have last used that drug in the month prior to the interview month.
The R was interviewed in January and reported first use of a drug (or first daily cigarette use) in December of the previous year.	If the YLU had a missing value or had a value that was inconsistent with the YFU, then the YLU was set to equal the YFU. Concurrently, the MLU was set to equal the MFU under similar constraints.
The ALU had a missing value, but the calculated ALU variable (MYR; e.g., MYRLSTMJ for marijuana) from the YLU, MLU, and birth date was consistent with the AFU and the R's current age. In addition, none of the above edits had been applied.	The ALU was set equal to the calculated ALU (MYR).
The R gave a valid value for the YLU and reported last use in the same month when he or she had a birthday, or else the R answered the MLU question as "don't know" or "refused."	Two possible values for the ALU were calculated, based on the YLU that the R reported and the R's birth year. If neither of these two possible values matched the ALU that the R reported, then the YLU and MLU were set to bad data.
The R reported an ALU (and also may have reported values for the YLU and MLU) for when he or she last used any hallucinogen. However, the ALU, YLU, and MLU variables were skipped for LSD, PCP, or Ecstasy or Molly because the R had used only LSD, PCP, or Ecstasy or Molly, respectively.	Values from the ALU, YLU, and MLU for any hallucinogen (HALLAGLST, HALLYRLST, and HALLMOLST, respectively) were transferred to the edited variables that had been skipped for LSD, PCP, or Ecstasy or Molly. For example, if the R was a lifetime user only of Ecstasy or Molly, data from HALLAGLST, HALLYRLST, and HALLMOLST were transferred to the corresponding variables ECSAGLST, ECSYLU, and ECSMLU for Ecstasy or Molly.

Table B.16 Edits in the Prior Substance Use Section Pertaining to the Age, Year, and Month of Last Use (Other than for Consistency Checks) (continued)

Issue	Edits Implemented
The R reported last using any cocaine and crack cocaine in the same year. However, the R also reported last using any cocaine in a month that was earlier than the month when the R reported <i>first</i> using crack cocaine. This issue also could occur for daily cigarette use (relative to any cigarette use).	The MLU for any cocaine was set to bad data. Similar edits were implemented for the MLU variables for any cigarette use as applicable.
The ALU for crack cocaine was later than the corresponding ALU for any cocaine. This issue also could occur for daily cigarette use (relative to any cigarette use).	The ALU for cocaine was set to equal the ALU for crack cocaine. Similar edits were implemented for the ALU any cigarette use as applicable.
The ALU for any cocaine had a missing value. However, the R reported last using crack cocaine at his or her current age. This issue also could occur for daily cigarette use (relative to any cigarette use).	• The ALU for any cocaine was set to equal the ALU for crack cocaine. Logically, if the R last used crack at his or her current age, then that also had to be the age at which the R last used any cocaine. • As part of this same edit, if the crack YLU variable indicated that the R last used crack in the current year, then the cocaine YLU was set to the current year. Similar edits were implemented for the ALU for any cigarette use, as applicable.
The YLU for crack cocaine was later than the YLU for any cocaine. This issue also could occur for daily cigarette use (relative to any cigarette use).	The YLU for any cocaine was set to equal the YLU for crack cocaine. Values in the MLU for crack cocaine also were carried over to the MLU for any cocaine. Similar edits were implemented for the YLU and MLU variables for any cigarette use, as applicable.
The YLU for cocaine had a missing value. However, the R also indicated that he or she last used crack cocaine at an age that was 1 year younger than his or her current age. The R also reported last using crack cocaine in the current year. This issue also could occur for daily cigarette use (relative to any cigarette use).	No editing was done to the ALU for any cocaine, but the YLU for any cocaine was set to the current year. Logically, if the last use of crack cocaine was in the current year, then the last use of any cocaine had to be in the current year as well. Similar edits were implemented for the YLU and MLU variables for any cigarette use, as applicable.
The R last used any cocaine and crack cocaine at the same age and in the same year. However, the R reported last using crack cocaine in a month that was later than what the R reported for last use of any cocaine. This issue also could occur for daily cigarette use (relative to any cigarette use).	The MLU for any cocaine was set to equal the MLU for crack cocaine. Similar edits were implemented for the MLU variables for any cigarette use, as applicable.
The cocaine MLU had a missing value, but the R reported last using crack cocaine in the calendar month prior to the interview month. This issue also could occur for daily cigarette use (relative to any cigarette use).	The MLU for any cocaine was set to the calendar month prior to the interview month. Logically, if the R was asked the questions about any cocaine and crack cocaine in the prior substance use section because he or she had not used these drugs in the past 30 days, the R also had to have last used any cocaine in the month prior to the interview month. Similar edits were implemented for the MLU variables for any cigarette use, as applicable.

Table B.16 Edits in the Prior Substance Use Section Pertaining to the Age, Year, and Month of Last Use (Other than for Consistency Checks) (continued)

Issue	Edits Implemented
The ALU for any cocaine had a missing value, but the calculated value in the MYR variable for crack cocaine was consistent with both the crack ALU and the R's current age. In addition, none of the above edits was applied. This issue also could occur for daily cigarette use (relative to any cigarette use).	The ALU for any cocaine was set to equal the MYR value for crack cocaine. Similar edits were implemented for the ALUs for any cigarette use, as applicable.
The R reported last using any hallucinogen and LSD, PCP, or Ecstasy or Molly in the same year. However, the R also reported last using any hallucinogens in a month that was earlier than the month when the R reported <i>first</i> using LSD, PCP, or Ecstasy or Molly.	The MLU for any hallucinogens was set to bad data.
At least one of the ALUs for LSD, PCP, and Ecstasy or Molly was greater than the ALU for any hallucinogen.	• The hallucinogen ALU was set to the latest ALU from the ALUs for LSD, PCP, or Ecstasy. • As part of this same edit, if a given ALU for LSD, PCP, or Ecstasy was moved over to the ALU for any hallucinogens and the corresponding YLU for LSD, PCP, or Ecstasy was later than the YLU for any hallucinogen, then the later YLU was moved over to the YLU for any hallucinogen. If two or all three of these drugs had the same latest ALU value relative to the hallucinogen ALU, then the latest YLU from these drugs was moved over to the YLU for any hallucinogen. • If the YLU for LSD, PCP, or Ecstasy was moved over to the YLU for any hallucinogen and the hallucinogen MLU had a nonmissing value that was earlier than the corresponding MLU for LSD, PCP, or Ecstasy, then the MLU for any hallucinogen was set to the latest MLU from LSD, PCP, or Ecstasy.
The ALU for any hallucinogens had a missing value, but one or more ALUs for LSD, PCP, or Ecstasy indicated that the R last used at his or her current age.	• The hallucinogen ALU was set to be equal to the R's current age. • As part of this same edit, if a given YLU for LSD, PCP, or Ecstasy indicated that the R last used in the current year, then the hallucinogen YLU was set to the current year.
The ALU for any hallucinogens had a missing value, and one or more ALUs for LSD, PCP, or Ecstasy indicated that the R last used at a year younger than his or her current age. At least one YLU for LSD, PCP, or Ecstasy indicated that the R last used any of these drugs in the current year.	The hallucinogen YLU was set to the current year. No editing was done to the hallucinogen ALU.

Table B.16 Edits in the Prior Substance Use Section Pertaining to the Age, Year, and Month of Last Use (Other than for Consistency Checks) (continued)

Issue	Edits Implemented
The MLU for any hallucinogens had a missing value, but the R reported last using LSD, PCP, or Ecstasy in the calendar month prior to the interview month.	The MLU for any hallucinogens was set to the calendar month prior to the interview month. Logically, if the R was asked the questions about any hallucinogens and LSD, PCP, or Ecstasy because he or she had not used these drugs in the past 30 days, the R also had to have last used any hallucinogens in the month prior to the interview month.

Table B.17 Edits Pertaining to the Receipt of Substance Treatment Variables That Existed Prior to 2004

Issue	Edits Implemented
The respondent's (R's) only report(s) of drug use in the earlier substance use sections that routed the R into question TX01 about lifetime substance treatment had been set to bad data as part of the editing of the previous substance use sections.	Nonblank values in the edited variables pertaining to receipt of substance treatment were replaced with bad data codes.
Responses to the questions on the receipt of treatment in the past 12 months and the last time that the R received treatment were inconsistent (e.g., if the R reported that he or she did not receive treatment in the past 12 months but subsequently reported last receiving treatment during that period).	The edits favored responses that indicated more recent receipt of treatment: • If an R responded affirmatively that he or she had received treatment in the past 12 months but reported last receiving treatment "more than 12 months ago," the edits logically inferred that the R last received treatment at some point in the past 12 months (i.e., TXRCVDREC = 8). • If an R reported that he or she did not receive treatment in the past 12 months but reported last receiving treatment in the past 12 months, the edits logically inferred that the R had received treatment in that period (i.e., TXYRRECVD = 3).
The question on the receipt of treatment in the past 12 months had missing data (e.g., a response of "don't know" or "refused"), but the question on the last time that the R received treatment did not. Alternatively, the question on the last time that the R received treatment had missing data, but the question on receipt of treatment in the past 12 months did not.	Where possible, data were used to replace the missing value with a nonmissing value. Suppose, for example, that the R did not know or refused to report whether he or she had received treatment in the past 12 months. • If the R reported last receiving treatment in this period, the ambiguous response was replaced with a value to indicate that the R had received treatment in this period (i.e., TXYRRECVD = 3). • If the R reported last receiving treatment more than 12 months ago, it was logically inferred that the question about receipt of any treatment in the past 12 months should have been answered as "no" (i.e., TXYRRECVD = 4).
The question about the most recent receipt of treatment had missing data.	Data from the past year treatment variable (TXYRRECVD) or the lifetime treatment variable (TXEVRRECVD) were used to replace missing values in the edited treatment recency variable TXRCVDREC. • If the R had received treatment in the past 12 months (TXYRRECVD = 1), TXRCVDREC was assigned a code of 8 to indicate treatment at some point in the past 12 months. • If the R had not received treatment in the past 12 months (TXYRRECVD = 2), TXRCVDREC was assigned a code of 13 to indicate that the last treatment episode was more than 12 months ago. • If the question about treatment in the past 12 months (TX02) was answered as "don't know" or "refused" and the R did not report currently being in treatment in question TX07, TXRCVDREC was assigned a code of 9 to indicate treatment at some point in the R's lifetime. (The R had to have answered the lifetime treatment question TX01 as "yes" in order to have been routed to TX02.)

Table B.17 Edits Pertaining to the Receipt of Substance Treatment Variables That Existed Prior to 2004 (continued)

Issue	Edits Implemented
The R reported currently being in treatment in question TX07, so the question about the most recent time that the R had been in treatment was skipped.	The edited variable corresponding to question TX07 (TXCURRENT) continued to be coded as 1 (i.e., "yes"). Instead of a legitimate skip code being assigned, the edited treatment recency variable (TXRCVDREC) was assigned a code of 7, where 7 = Still in treatment LOGICALLY ASSIGNED. A code of 21 (still in treatment) also was assigned to the treatment outcome variable TXLTYOCOM.
The R reported currently being in treatment in question TX07 but did not know or refused to report in question TX02 whether he or she had received treatment in the past 12 months.	The R was logically inferred to have received treatment in the past 12 months (TXYRRECVD = 3).
The R was routed to question TX07 but did not know or refused to report whether he or she was still in treatment.	Data from the treatment outcome variable (TXLTYOCOM) and treatment recency variable (TXRCVDREC) were used to replace missing values in the edited variable TXCURRENT, corresponding to TX07. • If the R had received treatment in the past 30 days (TXRCVDREC = 1) and reported still being in treatment (TXLTYOCOM = 1 or 21), TXCURRENT was assigned a code of 3 to indicate that the R logically was still in treatment. • Otherwise, if the R definitely had not received treatment in the past 30 days, TXCURRENT was assigned a code of 4 to indicate that the R logically was no longer in treatment.
The R reported that he or she was not currently in treatment (TXCURRENT = 2 or 4), but the R reported still being in treatment when asked about the outcome of the last treatment episode.	The treatment outcome variable (TXLTYOCOM) was assigned a bad data code.
The R specified receiving treatment for an over-the-counter (OTC) psychotherapeutic medication (e.g., aspirin).	This information on OTC drugs was not used to infer treatment for any of the psychotherapeutic drugs because the questions about receipt of treatment for psychotherapeutic drugs referred specifically to treatment for prescription medications (i.e., not to OTCs).
The R did not report receiving treatment for a particular drug during his or her last (or current) treatment episode, but treatment for this drug was specified as treatment for "some other drug." In the case of the psychotherapeutics, the "other" drug specified was not an OTC drug.	The R was inferred to have received (or be receiving) treatment for the use of that drug. A code of 3 (Yes LOGICALLY ASSIGNED) was assigned to the corresponding edited drug variable. If the R specified receiving treatment for the use of methamphetamine and question TX31A was not answered as yes, then the edited variable TXLTYMETH was assigned a code of 3. Because of the addition of the new methamphetamine section, however, no editing was done to TXLTYSTIM for treatment for the use of prescription stimulants.

Table B.17 Edits Pertaining to the Receipt of Substance Treatment Variables That Existed Prior to 2004 (continued)

Issue	Edits Implemented
The R did not report receiving treatment for a particular drug during his or her last (or current) treatment episode but indicated that this drug was the primary drug for which he or she last received treatment (or was currently receiving treatment).	The R was inferred to have received (or be receiving) treatment for the use of that drug.
The R reported receiving treatment only for alcohol in the past 12 months, but questions about treatment for specific drugs during the last or current treatment episode had missing values (i.e., "don't know," "refused," bad data, or blank).	The R was logically inferred not to have received treatment for that drug during the last or current episode. The missing value in that drug's variable was replaced with a special code of 4 (No LOGICALLY ASSIGNED).
The R was routed to questions about the last or current treatment episode but did not have any indication of treatment for any of the drugs that he or she ever used.	The following edits were implemented: • If the R reported receiving treatment only for alcohol in the past 12 months, a special logically inferred "yes" code of 5 was assigned to the variable for alcohol treatment during the last treatment episode (TXLTYALCO). • If the R reported receiving treatment only for drugs other than alcohol in the past 12 months, a special code of 5 was assigned to the "some other drug" variable (TXLTYOTHR) to indicate that the drug for which the R received treatment was unknown. • Otherwise, a special code of 7 was assigned to TXLTYOTHR, the "some other drug" variable, to indicate that treatment for alcohol or other drugs was unknown. (Prior to 2002, these edits required Rs to have denied receiving treatment for all drugs they had ever used. Since 2002, the above edits also have been implemented if Rs did not report treatment for any specific drugs, and missing data existed in the questions about treatment for specific drugs.)
The R was logically inferred to have received treatment for alcohol during the last or current treatment episode (TXLTYALCO = 5), and question TX36 about treatment for any other drug (TXLTYOTHR) was answered as "no."	The edited "OTHER, Specify" variables TXLTYA through TXLTYE were assigned legitimate skip codes.
The R refused to report in question TX36 whether he or she received treatment for any other drug.	The refusal was propagated onto the edited "OTHER, Specify" variables TXLTYA through TXLTYE. Since 2002, this edit has been implemented regardless of whether the R had reported treatment for at least one drug in questions TX26 through TX35. (Prior to 2002, this edit required at least one response of "yes" in TX26 through TX35.)

Table B.17 Edits Pertaining to the Receipt of Substance Treatment Variables That Existed Prior to 2004 (continued)

Issue	Edits Implemented
The R reported treatment only for "some other drug," but the only substances specified were tobacco products (i.e., cigarettes, chewing tobacco, snuff, cigars, pipe tobacco).	The variables specifying treatment for tobacco products were assigned bad data codes. In addition, other variables pertaining to the last (or current) treatment episode were assigned bad data codes if the items had been answered. The following variables were affected: TXLYTMNPL (i.e., main place where the R was last treated); TXLYYOCOM (i.e., outcome of the last treatment episode); variables beginning with TXPY (i.e., payment sources for the last [or current] treatment episode); and TXLYTDAYS (i.e., length of the last or current treatment). The rationale for these edits was that anything pertaining to the last treatment (e.g., payment sources for the last or current treatment) would logically be assumed to pertain to treatment only for tobacco.
If the R reported in the alcohol section that he or she never used alcohol (AL01 = 2), the R would be skipped out of question TX26, pertaining to receipt of treatment for alcohol. However, the R also could report in question TX03 that he or she received treatment for "alcohol only" or "alcohol and drugs" in the past 12 months.	The edited variable pertaining to receipt of alcohol treatment during the last or current episode (TXLYALCO) retained a code of 98 (blank).
Question TX37 pertaining to the main drug for which the R last received (or was currently receiving) treatment was skipped because the R reported receipt of treatment for only one drug during the last or current treatment episode. That includes situations in which the only drug for which the R reported receiving treatment was "some other drug."	The edited variable TXLYTMAIN was assigned a legitimate skip code, provided that none of the edited variables about treatment for alcohol through prescription sedatives (TXLYALCO through TXLYYSEDV) had a code of 98 (blank). Otherwise, TXLYTMAIN retained a code of 98.
The R reported being in treatment for 366 days in the past 12 months.	The edited variable TXLYTDAYS was edited to 365 days.
The length of time that the R reported currently being in treatment or being in treatment the last time translated to a number of years greater than the R's age.	The edited variable TXLYTDAYS was assigned a bad data code.

Table B.17 Edits Pertaining to the Receipt of Substance Treatment Variables That Existed Prior to 2004 (continued)

Issue	Edits Implemented
The R reported receiving treatment in the past 12 months and reported receiving treatment in the past 12 months for alcohol only or drugs only. However, this response was inconsistent with the responses to questions on the drugs for which the R was treated (or was being treated) during the last (or current) treatment episode. For example, the R reported being treated in the past 12 months only for alcohol but reported last being treated for use of one or more other drugs.	Logically, the last or current treatment episode would fall within the 12-month period prior to the interview. Therefore, the variable pertaining to receipt of treatment for alcohol, other drugs, or both in the past 12 months (TXYRALDGB) was edited as follows: • If the R originally indicated treatment for alcohol only (i.e., a code of 1 in question TX03), with treatment for other drugs also having been indicated during the last episode, a special code of 11 was assigned to TXYRALDGB. • If the R originally indicated treatment for drugs only (i.e., a code of 2 in question TX03), with treatment for alcohol also having been indicated during the last episode, a special code of 12 was assigned to TXYRALDGB. The edits were done in this manner because the subsequent fill pattern for specific locations where the R received treatment in the past 12 months was based on the R's original answer for receipt of treatment only for alcohol, only for other drugs, or both.
The R reported receiving treatment in the past 12 months but did not know or refused to report whether he or she received treatment only for alcohol, only for other drugs, or for both. However, data were provided on the drugs for which the R was treated during his or her last (or current) treatment episode.	Data on the drugs for which the R was last treated (or was currently being treated) were used to indicate the <i>minimum</i> for which the R could have been treated in the past 12 months: • If the R indicated last (or currently) being treated for alcohol but did not indicate treatment for other drugs during the last (or current) treatment episode, it was possible to infer in TXYRALDGB that the R was at least treated for alcohol in the past 12 months in TXYRALDGB (but the R also may have been treated for other drugs at some point during that period). A special code of 4 was assigned to TXYRALDGB. • If the R indicated last (or currently) being treated for one or more drugs other than alcohol but did not indicate treatment for alcohol, it was possible to infer in TXYRALDGB that the R was at least treated for drugs other than alcohol in the past 12 months. A special code of 5 was assigned to TXYRALDGB. • If the R reported last (or currently) being treated both for alcohol and for other drugs, it was possible to infer in TXYRALDGB that the R was treated for both alcohol and other drugs in the past 12 months. A special code of 6 was assigned to TXYRALDGB.
The R reported receiving treatment in the past 12 months and did not report receiving treatment in a particular location in the past 12 months, but this location was specified as treatment in "some other place" in the past 12 months.	The R was logically inferred to have received treatment in that location in the past 12 months. A code of 3 (Yes LOGICALLY ASSIGNED) was given to the edited treatment location variable in this situation.

Table B.17 Edits Pertaining to the Receipt of Substance Treatment Variables That Existed Prior to 2004 (continued)

Issue	Edits Implemented
The R reported receiving treatment in the past 12 months (or was inferred to have received treatment in the past 12 months) and did not report receiving treatment in a particular location that he or she subsequently reported was the main place that he or she received treatment the last time (or the main place where he or she was currently receiving treatment).	The R was logically inferred to have received treatment in that location in the past 12 months. A special logically assigned "yes" code of 5 was assigned to indicate that the affirmative response came from the data on the main location where the R last received (or was currently receiving) treatment. If the R reported that the main location where he or she received treatment was "some other place" and specified a valid response in question TX25SP, that "OTHER, Specify" response also was moved over to the "OTHER, Specify" variable TXYROTHSP pertaining to treatment locations in the past 12 months.
The R reported receiving treatment in the past 12 months but answered "no" to every item about particular locations for treatment in that period, including "some other place."	The edited variable pertaining to "some other place" (TXYROTHER) was assigned a special code of 7 to indicate that the treatment location was unknown.
The R reported receiving treatment in the past 12 months and did not initially indicate receiving treatment in a hospital emergency room in that period. However, the R subsequently reported receiving treatment in the past 12 months in an emergency room for use of marijuana, cocaine, heroin, LSD, PCP, or methamphetamine.	The variable that did not indicate treatment in an emergency room (TXYREMRGN) was edited to infer that the R had received treatment in that location in the past 12 months.
The R reported receiving treatment in the past 12 months in every specific location that was asked about (i.e., except for treatment in "some other place").	No editing was done if the R reported being or having been in treatment for 15 days or more. If the R reported being or having been in treatment for fewer than 15 days, however, responses of "yes" in the entire list of edited past year treatment location variables were replaced with bad data codes. If treatment in "some other place" also was reported, the edited variable TXYROTHER was assigned a bad data code. In the variable TXYROTHSP (i.e., the other treatment location that was specified), any responses were replaced with bad data codes. If the R also reported that he or she was still in treatment (TX07 = 1), the edited variable TXCURRENT also was assigned a bad data code.
Rs could report still being in treatment in question TX07 but may report that they received treatment only in jail in the past 12 months.	When Rs reported receiving treatment only in jail in the past 12 months, they were logically inferred not to be currently in treatment (TXCURRENT = 4). This edit also applied when Rs reported that the main treatment location was jail, and this was the only specific location reported for treatment in the past 12 months.
The R did not report a particular payment source for his or her last episode of treatment but specified this payment source as "some other source."	The R was inferred to have used that particular payment source for treatment. A code of 3 (Yes LOGICALLY ASSIGNED) was assigned to the edited variable for that payment source.
The R answered all items about payment sources for treatment as "no," including the item indicating that the last treatment was free.	A special code of 5 was assigned to the edited "some other source" variable (TXPAYOTHR) to indicate that the payment source was unknown.

Table B.17 Edits Pertaining to the Receipt of Substance Treatment Variables That Existed Prior to 2004 (continued)

Issue	Edits Implemented
The R reported that every specific payment source that was asked about paid for his or her last episode of treatment (i.e., except for "some other source" and free payment, the latter of which would have been skipped).	All source of payment variables that the R had answered as "yes" were assigned a bad data code. That included situations in which "some other source" of payment also was reported. In the variable TXPAYOTSP (i.e., the other payment source that was specified), any responses were replaced with bad data codes.
The R reported <i>all</i> of the following: • receipt of treatment in every specific location in the past 12 months (i.e., except for treatment in "some other place"), and • payment of the last treatment by every specific payment source (i.e., except for "some other source" and free treatment).	When this specific pattern occurred, data from additional variables also were assumed to be questionable. Responses entered for the following variables were replaced with bad data codes: TXYRALDGB (i.e., treatment for alcohol, drugs, or both in the past 12 months); TXYRERDRG (treatment in an emergency room for marijuana, cocaine, heroin, LSD, PCP, or methamphetamine in the past 12 months); TXYRERNUM (number of times the R visited an emergency room for treatment of the above drugs); TXLTYMNPL (the main place the R received treatment the last time); drugs that the R was asked about for the last treatment episode (including the main drug for which the R received treatment, if applicable); and TXLTYDAYS (length of time in treatment currently or the last time).
The R indicated that "some other source" paid for the last treatment, but then specified that this treatment was free.	If no other payment source was indicated, then it was logically inferred that the R's last treatment was free (i.e., TXPAYFREE = 3). Otherwise, if one or more payment sources had been indicated previously (e.g., private health insurance, the R's own funds), then it was inferred that "some other source" had not paid for the last treatment. In this situation, the response of free treatment that had been specified also was wiped out in the edited "OTHER, Specify" variable (TXPAYOTSP).
The R reported in question TX44 that the only treatment he or she received in the past 12 months was for detoxification (or answered TX44 as "don't know" or "refused"), but the R also reported attending self-help groups in the past 12 months. Self-help groups typically are not places where people go to receive detoxification.	The response was accepted that the R received treatment in a self-help group in the past 12 months, and the R was logically inferred to have received treatment other than detoxification in that period. The edited variable corresponding to question TX44 (TXYRONDTX) was assigned a code of 4 (No LOGICALLY ASSIGNED).

Table B.18 Edits Pertaining to the Receipt of Substance Treatment Variables That Were Added in 2004

Issue	Edits Implemented
The respondent (R) reported receiving substance treatment in the past 12 months (i.e., TXYRRECVD = 1).	The following lifetime treatment variables were assigned legitimate skip codes: TXALCONLY, TXALCONAG, TXDRGONLY, TXDRGONAG, TXALCDRGU, TXALCDAGE, TXDRGALCU, and TXDRGAAGE. These variables corresponded to questions TX45, TX45A, TX46, TX46A, TX47, TX47A, TX48, and TX48A, respectively.
The R was logically inferred to have received treatment in the past 12 months (i.e., TXYRRECVD = 3). Consequently, the R was routed into the lifetime treatment questions. Had the R answered the past year treatment question TX02 as "yes," however, the R would have been skipped out of the lifetime treatment questions.	Responses of 1 and 2 (i.e., "yes" and "no," respectively) in TXALCONLY, TXDRGONLY, TXALCDRGU, and TXDRGALCU were bumped by a value of 10 (i.e., set to values of 11 or 12). No editing was done to these variables if they had values of 94 ("don't know") or 97 ("refused").
The R received treatment in his or her lifetime but not in the past 12 months, and the R never used alcohol (i.e., ALCEVER = 2). The R also reported using at least one other drug in the marijuana through sedatives sections or in the special drugs section for methamphetamine.	The following variables pertaining to lifetime alcohol treatment were assigned legitimate skip codes: TXALCONLY, TXALCONAG, TXALCDRGU, TXALCDAGE, TXDRGALCU, and TXDRGAAGE.
The R received treatment in his or her lifetime but not in the past 12 months, and the R indicated unambiguously in the marijuana through sedatives sections or in the special drugs section for methamphetamine that he or she never used any of these drugs. The R also reported using alcohol.	All the variables pertaining to drugs are skip filled, namely, TXDRGONLY, TXALCDRGU, TXDRGALCU and the age variables TXDRGONAG, TXALCDAGE, and TXDRGAAGE.
The R reported first receiving substance treatment for alcohol or other drugs (or both) at an age that was later than his or her current age.	Consistent with standard editing practice, any ages for first receipt of treatment that were inconsistent with the R's current age were set to bad data.

Table B.18 Edits Pertaining to the Receipt of Substance Treatment Variables That Were Added in 2004 (continued)

Issue	Edits Implemented
The R reported first receiving treatment for alcohol or other drugs at an age that was earlier than the age-at-first-use (AFU) data from the previous substance use sections. - For alcohol, the R reported first receiving treatment for alcohol at an age that earlier than ALCTRY, the AFU for alcohol from the alcohol section. - For other drugs, the R reported first receiving treatment for his or her use of drugs at an age that was earlier than the minimum AFU for the drugs that the R reported using.	No editing was done to the age at first treatment data that were inconsistent with the AFU data from the previous substance use sections. However, flags were created to indicate the degree of deviation from (a) ALCTRY (TXFGALAGE), (b) the minimum AFU for other drugs (TXFGDGAGE), and (c) the minimum AFU from both ALCTRY and the AFUs for other drugs (TXFGADAGE). Values in these flags had the following meanings: 0 = First treatment age was consistent with minimum AFU 1 = First treatment age differs from minimum AFU by 1 year 2 = First treatment age differs from minimum AFU by 2 years 3 = First treatment age differs from minimum AFU by 3 or 4 years 4 = First treatment age differs from minimum AFU by 5 or more years. The flag for both alcohol and drugs (TXFGADAGE) was created from the values in TXFGALAGE and TXFGDGAGE. The final value in TXFGADAGE was chosen according to whatever value in TXFGALAGE and TXFGDGAGE indicated the greatest inconsistency between treatment age and the corresponding AFU data. If both TXFGALAGE and TXFGDGAGE were blank, then TXFGADAGE was set to blank as well.
The R could report receiving treatment for alcohol in the past 12 months and could give an age when he or she first received alcohol treatment. However, the R previously reported in the alcohol section that he or she never used alcohol. Similarly, the R could report receiving treatment for other drugs in the past 12 months and could give an age when he or she first received drug treatment. However, the R previously reported never using drugs in the marijuana through sedatives sections.	Again, no editing was done to the inconsistent treatment age data. However, the flag variables mentioned above were assigned a code of 5 when this situation occurred. This code of 5 had the following meanings: 5 = Never used alcohol but reported an alcohol treatment age 5 = Never used drugs but reported a drug treatment age. If TXFGALAGE or TXFGDGAGE had a value of 5, then TXFGADAGE also was assigned a value of 5, where 5 = Any lifetime nonuse with a related treatment age.

Table B.18 Edits Pertaining to the Receipt of Substance Treatment Variables That Were Added in 2004 (continued)

Issue	Edits Implemented
The R answered the lifetime treatment question TX01 (edited variable TXEVRRCVD) as "don't know" or "refused."	The following edits were implemented for the lifetime treatment variables TXALCONLY through TXDRGAAGE according to what the R reported in the earlier substance use sections for lifetime use (or nonuse) of alcohol or drugs. • If the R was a lifetime user only of alcohol (i.e., had definitely never used other drugs that were covered in the earlier substance use sections), the lifetime variables pertaining to treatment for drugs (TXDRGONLY, TXDRGONAG, TXALCDRGU, TXALCDAGE, TXDRGALCU, and TXDRGAAGE) were assigned legitimate skip codes; even if the R had reported lifetime receipt of treatment, these variables still would have been skipped because the R never used any of the drugs covered in the earlier substance use sections. In addition, if TXEVRRCVD was refused, that refusal was propagated to TXALCONLY and TXALCONAG. • If the R was a lifetime user of other drugs but had never used alcohol, the lifetime variables pertaining to treatment for alcohol (TXALCONLY, TXALCONAG, TXALCDRGU, TXALCDAGE, TXDRGALCU, and TXDRGAAGE) were assigned legitimate skip codes; even if the R had reported lifetime receipt of treatment, these variables still would have been skipped because the R never used alcohol. In addition, if TXEVRRCVD was refused, that refusal was propagated to TXDRGONLY and TXDRGONAG.
The R had the following pattern in his or her data: 1. Lifetime use of alcohol (i.e., ALCEVER = 1). 2. Initially skipped out of past year treatment questions but was logically inferred to have received treatment in the past 12 months (i.e., TXYRRECVD = 3; see Table B.17). 3. Never used marijuana through sedatives, or methamphetamine (based on answers from special drugs). 4. Question TX45 (edited variable TXALCONLY) was answered as "no," meaning that the R had never received treatment or counseling for alcohol. Logically, the R had to have received treatment for something, and alcohol was the only substance that the R reported using.	TXALCONLY was bumped by 20, such that TXALCONLY would show a value of 22. This was done instead of inferring that the R received alcohol treatment. Nevertheless, a value of 22 in TXALCONLY would alert analysts to an inconsistency in the data and give them the option of deciding how to handle these cases in an analysis.

Table B.18 Edits Pertaining to the Receipt of Substance Treatment Variables That Were Added in 2004 (continued)

Issue	Edits Implemented
The R had the following pattern in his or her data: 1. Lifetime use of at least one drug in the marijuana through sedatives sections or methamphetamine use from the special drugs section. 2. Initially skipped out of past year treatment questions but was logically inferred to have received treatment in the past 12 months (i.e., TXYRRECVD = 3). 3. Never used alcohol. 4. Question TX46 (edited variable TXDRGONLY) was answered as "no," meaning that the R had never received treatment or counseling for drugs other than alcohol. Logically, the R had to have received treatment for something, and illicit drugs (but not alcohol) were the only substances that the R reported using.	TXDRGONLY was bumped by 20, such that TXDRGONLY would show a value of 22. This was done instead of inferring that the R received drug treatment. Nevertheless, a value of 22 in TXDRGONLY would alert analysts to an inconsistency in the data and give them the option of deciding how to handle these cases in an analysis.
The R reported lifetime use of both alcohol and other drugs. However, both TXALCDRGU = 2 and TXDRGALCU = 2, meaning that the R never got treatment for either alcohol or drugs. Logically, if the R answered the lifetime treatment question TX01 (edited variable TXEVRRCD) as "yes," the R had to have gotten treatment for something.	Both TXALCDRGU and TXDRGALCU were bumped by 20, such that the edited value would be 22. Again, this would give analysts the option of deciding how to handle these cases in an analysis.
A lifetime lead variable (i.e., TXALCONLY, TXDRGONLY, TXALCDRGU, and TXDRGALCU) was coded as 2 (i.e., "no") or was coded as 12 (based on the second issue described in this table).	The corresponding age-at-first-treatment variables TXALCONAG, TXDRGONAG, TXALCDAGE, and TXDRGAAGE were assigned legitimate skip codes.
A past year lead variable (i.e., TXYALODRG, TXYDROALC) was coded as 2 (i.e., "no").	The corresponding age-at-first-treatment variables TXYALODAG and TXYDROAAG were assigned legitimate skip codes.

Table B.18 Edits Pertaining to the Receipt of Substance Treatment Variables That Were Added in 2004 (continued)

Issue	Edits Implemented
The past year treatment variable TXYRALDGB (corresponding to question TX03) indicated that the R was treated in the past year only for alcohol or only for drugs.	If TXYRALDGB = 1 (i.e., treatment in the past year only for alcohol), the following variables were assigned legitimate skip codes: TXYDRONAG (age at first treatment for drugs, if the R had been treated in the past year only for drugs), TXYDROALC (whether the R ever was treated for alcohol use, if the R had ever used alcohol and had been treated in the past year only for drugs), TXYDROAAG (the treatment age variable corresponding to TXYDROALC), TXYALDAAG (age at first treatment for alcohol, if the R had been treated in the past year for both alcohol and drugs), and TXYALDDAG (age at first treatment for drugs, if the R had been treated in the past year for both alcohol and drugs). Similarly, if TXYRALDGB = 2 (i.e., treatment in the past year only for drugs), the following variables were assigned legitimate skip codes: TXYALONAG (age at first treatment for alcohol, if the R had been treated in the past year only for drugs), TXYALODRG (whether the R ever had been treated for drug use, if the R had ever used marijuana through sedatives and had been treated in the past year only for alcohol), TXYALODAG (the treatment age variable corresponding to TXYALODRG), TXYALDAAG (age at first treatment for alcohol, if the R had been treated in the past year for both alcohol and drugs), and TXYALDDAG (age at first treatment for drugs, if the R had been treated in the past year for both alcohol and drugs).
TXYRALDGB indicated that the R was treated in the past year for both alcohol and drugs, including situations in which the R was logically inferred to have received treatment for both alcohol and drugs in the past year.	The following variables were assigned legitimate skip codes: TXYALONAG, TXYALODRG, TXYALODAG, TXYDRONAG, TXYDROALC, and TXYDROAAG. If TXYRALDGB = 6 (i.e., logically inferred to have received treatment for both alcohol and drugs in the past year; see Table B.17), the skipped variables TXYALDAAG and TXYALDDAG retained codes of blank.
TXYRALDGB had been assigned a code of 4 (i.e., received treatment for alcohol in the past year but treatment for drugs during this period was unknown; see Table B.17).	The following variables pertaining to receipt of treatment only for drugs were assigned legitimate skip codes: TXYDRONAG, TXYDROALC, and TXYDROAAG. The skipped variables pertaining to treatment for alcohol in the past year, with or without treatment for drugs (i.e., TXYALONAG, TXYALODRG, TXYALODAG, TXYALDAAG, and TXYALDDAG), retained codes of blank.
TXYRALDGB had been assigned a code of 5 (i.e., received treatment for alcohol in the past year but treatment for drugs during this period was unknown; see Table B.17).	The following variables pertaining to receipt of treatment only for alcohol were assigned legitimate skip codes: TXYALONAG, TXYALODRG, and TXYALODAG. The skipped variables pertaining to treatment for drugs in the past year, with or without treatment for alcohol (i.e., TXYDRONAG, TXYDROALC, TXYDROAAG, TXYALDAAG, and TXYALDDAG), retained codes of blank.

Table B.18 Edits Pertaining to the Receipt of Substance Treatment Variables That Were Added in 2004 (continued)

Issue	Edits Implemented
TXYRALDGB had been assigned a code of 11 (i.e., the R had reported receiving treatment only for alcohol in the past 12 months but some treatment for drugs also was indicated; see Table B.17). TXYALODRG did not indicate that the R had ever received treatment for his or her use of drugs.	TXYALODRG was assigned a code of 3 (i.e., Yes LOGICALLY ASSIGNED).
TXYRALDGB had been assigned a code of 12 (i.e., the R had reported receiving treatment only for drugs in the past 12 months but some treatment for alcohol also was indicated; see Table B.17). TXYDROALC did not indicate that the R had ever received treatment for his or her use of alcohol.	TXYDROALC was assigned a code of 3 (i.e., Yes LOGICALLY ASSIGNED).
TXYRALDGB had been set to bad data.	Any nonblank values in variables that were dependent on TXYRALDGB were set to bad data. This edit was relevant to the following questions (corresponding edited variables shown in parentheses): TX49 (TXYALONAG), TX49A (TXYALODRG), TX49B (TXYALODAG), TX50 (TXYDRONAG), TX50A (TXYDROALC), TX50B (TXYDROAAG), TX51 (TXYALDAAG), and TX51A (TXYALDDAG).
TXYRALDGB was answered as "don't know" or "refused."	All of the variables TXYALONAG through TXYALDDAG that had been skipped retained codes of blank.

Table B.19 Edits Pertaining to the Perceived Need for Treatment Variables

Issue	Edits Implemented
The respondent (R) specified the need for treatment for an over-the-counter (OTC) psychotherapeutic medication (e.g., aspirin).	This information on OTC drugs was not used to infer need for treatment for any of the psychotherapeutic drugs because the questions about perceived need for treatment for psychotherapeutic drugs referred specifically to prescription medications (i.e., and not OTCs).
The R did not report needing treatment for a particular drug in the past 12 months, but need for treatment for this drug was specified as a treatment need for "some other drug." In the case of the psychotherapeutics, the other drug specified was not an OTC drug.	The R was inferred to perceive the need for treatment for the use of that drug. The edited variable for the perceived need for treatment for that drug was assigned a code of 3 (Yes LOGICALLY ASSIGNED). This code of 3 could be edited further, as discussed below.
The R reported needing treatment in the past 12 months for the use of alcohol or other drugs, but questions about the perceived need for treatment for all specific drugs that the R had ever used were answered as "no."	A special code was assigned to the "some other drug" variable (NDTXYROTH) to indicate that the specific drug for which the R thought that he or she needed treatment was unknown.
Question TX10, pertaining to the perceived need for additional treatment, is an "enter all that apply" type of question. That is, Rs could report needing additional treatment for more than one drug shown in the list in TX10. However, Rs could report needing additional treatment for drugs that they had reported never using in the corresponding substance use section (e.g., reported never using heroin but reported needing additional treatment for heroin). In contrast, Rs would not get asked questions TX11 through TX21 (regarding perceived need for treatment for specific drugs) unless they were lifetime users of a particular drug.	No editing was done when this pattern occurred. Consequently, these data would be inconsistent with the earlier substance use data.

Table B.19 Edits Pertaining to the Perceived Need for Treatment Variables (continued)

Issue	Edits Implemented
The R was logically inferred to have received treatment in the past 12 months (TXYRRECVD = 3). Because the R did not originally answer question TX02 as "yes," the computer-assisted interviewing (CAI) program routed the Rs to questions about whether they thought they needed treatment for their use of alcohol or specific drugs (i.e., question TX08 and questions TX11 through TX22).	The following edits were done when TXYRRECVD = 3: • If a question was originally answered as "yes," then the corresponding edited variable was assigned a code of 11 (Yes [TXYRRECVD = 3]). For example, if the R reported needing treatment for alcohol or other drugs (TX08 = 1), then the edited variable NDTXYRADG was assigned a code of 11. Similarly, if the R reported needing treatment for a specific drug (e.g., prescription pain relievers), then the edited variable (e.g., NDTXYRPNR) was assigned a code of 11. • If a question was originally answered as "no," then the corresponding edited variable was assigned a code of 12 (No [TXYRRECVD = 3]). For example, if TX08 had been answered as "no" (TX08 = 2), then NDTXYRADG was assigned a code of 12. (If NDTXYRADG was set to 12, then subsequent variables continued to be assigned legitimate skip codes.) Similarly, if a question about the need for treatment for a specific drug had been answered as "no," then the edited variable was assigned a code of 12. • If the R was inferred to perceive the need for treatment for a drug based on "OTHER, Specify" data, the edited variable was assigned a code of 13. Suppose, for example, that NDTXYRPNR had already been coded as 3 because the R had specified prescription pain relievers as "some other drug" for which the R needed treatment (but question TX17 had not been answered as "yes"). If the R was logically inferred to have received treatment in the past 12 months (TXYRRECVD = 3), then NDTXYRPNR was subsequently coded as 13 (Yes LOGICALLY ASSIGNED [TXYRRECVD = 3]). • If the R was a lifetime nonuser of a drug, the edits continued to assign a legitimate skip code. For example, if the R had never misused prescription pain relievers, then NDTXYRPNR continued to receive a code of 99 when TXYRRECVD = 3. • Codes for any reasons that respondents reported for why they did not get treatment were bumped by 10; the resulting codes were 11 or 13. The rationale for these edits was that Rs would not have been asked questions about their perceived need for treatment for alcohol or specific other drugs if they had originally reported that they received treatment in the past 12 months. The above edits were done to conserve respondents' answers, as opposed to wiping out the data.
The R reported making an effort to get treatment (question TX22 answered as "yes"), but the R reported not needing treatment for every specific drug that he or she was asked about.	The edited variable NDTXEFFRT was assigned a code of 11. The same edits described above for other variables that applied when TXYRRECVD = 3 also were performed when NDTXEFFRT was assigned a code of 11 due to this issue.

Table B.20 Prioritization of Specific Skip Logic Edits for the Consumption of Alcohol Section

Response Pattern	Edit
Data from the alcohol section indicated that the respondent (R) never used alcohol.	A code of 91 (or 991, 9991) was assigned to the edited variables in this section. This edit took precedence over situations in which questions may not have been applicable for other reason (e.g., if Rs were male).
Even if the R had used alcohol, data from remaining substance use sections indicated that the R had never used drugs other than alcohol.	A code of 91 was assigned to the edited variables corresponding to questions CA08 and CA09, pertaining to alcohol use in combination with other drugs in the past 30 days (edited variables CADRKDRUG through CADRKMETH). The code of 91 had the following meaning in these variables: NEVER USED ALCOHOL OR NEVER USED OTHER DRUGS. Similarly, if the R had not used illicit drugs and the R had been logically inferred to be a nonuser of at least one drug, then these edited variables were assigned a code of 81, where 81 = NEVER USED ALCOHOL OR NEVER USED DRUGS Logically assigned.
Data from the alcohol section indicated that the R had used alcohol, but definitely not in the past 30 days.	A code of 93 (or 993, 9993) was assigned to the edited variables in this section to indicate that the R did not use alcohol in the past 30 days.
Even if the R had used alcohol in the past 30 days, data from remaining substance use sections indicated that the R had used some other drugs in the lifetime but not in the past 30 days.	A code of 93 was assigned to the edited variables corresponding to questions CA08 and CA09. The code of 93 had the following meaning in these variables: DID NOT USE ALCOHOL OR DID NOT USE DRUGS IN THE PAST 30 DAYS.
The R had used alcohol in the past 30 days (or was potentially a past month user, based on ALCREC = 8 from the alcohol section), but the R was aged 21 or older.	The edited variables corresponding to questions CA02 through CA07SP about underage alcohol use were assigned a code of 99 (LEGITIMATE SKIP) if they were blank or otherwise were assigned a code of 89 (LEGITIMATE SKIP Logically assigned).

Table B.20 Prioritization of Specific Skip Logic Edits for the Consumption of Alcohol Section (continued)

Response Pattern	Edit
The R had used alcohol in the past 30 days (or was potentially a past month user) and was aged 12 to 20.	• If the R did not indicate in question CA02B that he or she drank the alcohol in some other place the last time, then the edited variable CADROTHSP (corresponding to question CA02SP) was assigned a legitimate skip code (i.e., 99 if blank or 89 otherwise). • If the R indicated in question CA03 (edited variable CABUYFRE) that the R paid for the last alcohol that he or she drank: — Edited variables corresponding to questions CA07 (edited variable CAFREWHO) and CA07SP (edited variable CAFRESP) were assigned a legitimate skip code. — If the R paid for the last alcoholic beverage that he or she drank but the R gave money to someone else who bought it (question CA04 = 2, corresponding to the edited variable CAGVMONY), then the edited variables corresponding to questions CA05 (CABUYWHO), CA05A (CABPLACE), and CA05B (CABUNDAG) were assigned a legitimate skip code. — If the R paid for the last alcoholic beverage that he or she drank and also bought it (CAGVMONY = 1), then the edited variable corresponding to question CA06 (CAGVWHO) was assigned a legitimate skip code. In addition, if the R reported in CA05 that he or she bought the alcohol in a store, in a restaurant, in a club, or at an event (CA05 = 1), then CABUNDAG was assigned a legitimate skip code. If the R reported in question CA05 that he or she bought the alcohol from another person (CA05 = 2), then CABPLACE was assigned a legitimate skip code. • If the R indicated in question CA03 that he or she did not pay for the last alcohol that he or she drank: — CAGVMONY, CABUYWHO, CABPLACE, CABUNDAG, and CAGVWHO were assigned a legitimate skip code. — If the R did not report in question CA07 that he or she got the last alcoholic beverage some other way, then the edited variable CAFRESP (corresponding to CA07SP) was assigned a legitimate skip code.

Table B.20 Prioritization of Specific Skip Logic Edits for the Consumption of Alcohol Section (continued)

Response Pattern	Edit
The R had used alcohol in the past 30 days and also had used other drugs in that period.	• If the R reported in question CA08 (edited variable CADRKDRUG) that he or she did not use other drugs at the same time or within a couple of hours of the last time when the R drank alcohol in the past 30 days, then the edited variables corresponding to the drugs in question CA09 (i.e., edited variables CADRKMARJ through CADRKMETH) were assigned legitimate skip codes. • Otherwise (e.g., if CA08 was answered as "yes"), if the corresponding data from earlier illicit drug sections indicated that the R had not used a particular drug in the past 30 days but the R had used some other drug in the past 30 days, then the variable corresponding to the drug that the R had not used was assigned a legitimate skip code. Suppose, for example, that the R used other drugs at the same time or within a couple of hours of when the R last used alcohol and the R had used marijuana but not cocaine in the past 30 days. In this situation, CADRKCOCN (i.e., use of cocaine in combination with alcohol) was assigned a legitimate skip code.
The R reported using other drugs at the same time or within a couple of hours of when he or she last used alcohol. However, question CA09 about use of alcohol in combination with specific drugs was skipped because the R reported use of only one illicit drug in the past 30 days.	The edited variable corresponding to the one illicit drug that the R had used in the past 30 days was assigned a code of 5, where 5 = Response entered LOGICALLY ASSIGNED (from skip pattern). For all other drugs that the R had definitely not used in the past 30 days, the edited variables were assigned legitimate skip codes. For example, if the R answered question CA08 as "yes," and the only drug that the R reported using in the past 30 days was marijuana, then it could be logically inferred that the R used marijuana at the same time or within a couple of hours of when he or she last used alcohol, and the edited variable CADRKMARJ was assigned a code of 5.
The lifetime binge alcohol question CA10 (for consumption of four or more drinks on a single occasion for females or consumption of five or more drinks on a single occasion for males) had been skipped because either of the following occurred: • The alcohol use data indicated binge alcohol use on 1 or more days in the past 30 days based on the thresholds described above for males and females. • Question CA01 (edited variable CADRLAST) indicated that the R binge drank the last time that he or she drank in the past 30 days.	The edited variable CABINGEVR (corresponding to question CA10) was set to a value of 5, where 5 = Yes LOGICALLY ASSIGNED (from skip pattern).
Question CA10 was answered as "no," indicating that the R had never engaged in binge drinking (see above).	The edited variables CABINGAGE (corresponding to question CA11), CABINGYFU (derived from questions CA11a through CA11c), and CABINGMFU (derived from questions CA11c and CA11d) were assigned legitimate skip codes.

**Table B.20 Prioritization of Specific Skip Logic Edits for the Consumption of Alcohol Section
(continued)**

Response Pattern	Edit
The R was inferred to have used alcohol at some point in the past 12 months (based on ALCREC = 8 from the alcohol section) because the R reported use on 0 days in the past 30 days (Table 6.2). However, the R was asked questions CA01 through CA09 (where applicable), as though the R used alcohol in the past 30 days.	If data existed in questions CA01 through CA09 (i.e., if Rs were not legitimately skipped out of questions because of other patterns described in this table, such as if Rs were aged 21 or older), then these data were retained in the corresponding edited variables. Beginning in 2017, Rs will not be asked these consumption of alcohol questions if they reported in the alcohol section of the interview that they used alcohol on 0 days in the past 30 days.

Table B.21 Edits in the Consumption of Alcohol Section Pertaining to the Binge Alcohol Use Variables

Issue	Edits Implemented
The respondent (R) indicated in the final verification check for the age when he or she first engaged in binge alcohol use (BACC06) that the age at first binge alcohol use based on the new month and year of first use (MFU and YFU) was correct (i.e., BACC06 = 4). The computer-assisted interviewing (CAI) program updated the age at first use (AFU; AGE1STBA) with the value of the age calculated from the MFU and YFU (i.e., MYR1STBA). However, the new value for the AFU indicated that the R was more than 1 year younger than his or her current age at the time the R first engaged in binge alcohol use. Had the R initially reported this AFU, the R would not have been routed to the relevant MFU and YFU questions.	The updated value was retained in the edited AFU variable CABINGAGE. Based on the updated value for the AFU, it was logically inferred that the R should have skipped the MFU and YFU items. A code of 9989 was assigned to the relevant YFU variable (CABINGYFU), and a code of 89 was assigned to the corresponding MFU variable (CABINGMFU).
The final verification check (BACC06) was skipped because the R entered revised data for the relevant MFU and YFU that made MFU and YFU for first binge alcohol use consistent with the AFU.	No editing was done because the R was considered to have resolved the inconsistency.
The final verification check (BACC06) was skipped because the R entered a new MFU that was the same as the R's birth month.	The new MFU could be consistent with the AFU, depending on whether the use in that month occurred before or after the R's birthday. No editing was done to the AFU, MFU, and YFU, as long as the revised MFU and YFU were potentially consistent with the AFU. However, the revised MFU and YFU were set to bad data if they could never be consistent with the AFU.
The R entered a new MFU or YFU that differed from what the R previously reported. The age based on the revised MFU and YFU (updated in MYR1STBA) still mismatched the AFU, but the R indicated in the final verification check that the new value from the relevant MYR variable was correct.	No editing was done in this situation. The CAI program automatically updated the value for the AFU to be consistent with the updated values reported for the MFU and YFU.
The consistency check was triggered between the AGE1STBA variable stored by the CAI program and the corresponding MYR1STBA variable. However, the R answered the first consistency check (i.e., BACC03, regarding whether the value in MYR1STBA was correct) as "don't know" or "refused." The R then exited the consistency check loop without having resolved the inconsistency.	The AFU value that had been stored in the AGE1STBA variable was retained, but the corresponding MFU and YFU variables were set to bad data.
The consistency check was triggered between the AGE1STBA and corresponding MYR1STBA variable. However, the R entered the same values for the YFU and MFU that triggered the inconsistency with the AFU in the first place.	No editing was done to the AFU, but the YFU and MFU variables were set to bad data.

Table B.21 Edits in the Consumption of Alcohol Section Pertaining to the Binge Alcohol Use Variables (continued)

Issue	Edits Implemented
The R entered new values in the consistency checks for the MFU and YFU that again yielded a nonmissing value for the MYR1STBA variable. However, the R failed to resolve the inconsistency between the AGE1STBA variable and the updated value in the MYR1STBA variable. The R also reported either in the first verification check that MYR1STBA was not correct (i.e., BACC03 = 6) or reported in the second verification check that the AFU from the AGE1STBA variable was correct (i.e., BACC04 = 2).	No editing was done to the AFU. The following edits were implemented for the MFU and YFU: • The default edit was to set the MFU and YFU to bad data. • As an exception to the default edit, if the final verification check was answered as "don't know" or "refused," the MFU and YFU were assigned the code that corresponded to the answer in the final verification check.
The R entered new values in the consistency checks for the MFU and YFU that again yielded a nonmissing value for the MYR1STBA variable. However, the R failed to resolve the inconsistency between the AGE1STBA variable and the updated value in the MYR1STBA variable. The R also reported in the second verification check that neither answer was correct for what was originally captured in the AGE1STBA and MYR1STBA variables (i.e., BACC04 = 3).	The following edits were implemented for the AFU, MFU, and YFU: • The default edit was to set the AFU, MFU, and YFU to bad data. • As an exception to the default edit, if the final verification check was answered as "don't know" or "refused," the AFU, MFU, and YFU were assigned the code that corresponded to the answer in the final verification check.
The AFU was inconsistent with the value from the MYR1STBA variable (based on the MFU and YFU). The R also reported that the age from the MYR1STBA variable was correct (i.e., BACC03 = 4). However, the R answered the second consistency check (BACC04) as "don't know" or "refused." Consequently, the R did not have an opportunity to correct the inconsistency between the AFU, MFU, and YFU.	The relevant codes for "don't know" or "refused" were assigned to the AFU, MFU, and YFU. The rationale for this edit is that conclusive information did not exist regarding whether the AFU indicated the R's correct age when he or she first engaged in binge alcohol use, or whether the MFU and YFU indicated the R's correct age at initiation of binge alcohol use. Therefore, the AFU, MFU, and YFU all were set to missing values.
The R reported first binge alcohol use at an age that was earlier than what was reported for his or her first use of any alcohol or at an age that was greater than the R's current age.	The edited AFU variable (i.e., CABINGAGE) was set to bad data. Any nonmissing values in the YFU and MFU also were set to bad data.
The R reported first binge alcohol use in a year that was earlier than what the R reported for the year in which he or she first used any alcohol.	The edited YFU variable was set to bad data. Any nonmissing values in the MFU also were set to bad data.
The R reported first binge alcohol use in the same year in which the R first used any alcohol. However, the R reported first binge alcohol use in a month that was earlier than the month in which the R first used any alcohol.	The edited MFU variable was set to bad data.
The R reported first binge alcohol use in the same month as his or her birth month. However, the R could never have first engaged in binge alcohol use at the age he or she reported and also have first engaged in binge alcohol use in the year that the R reported.	The relevant YFU and MFU variables were set to bad data.

Table B.21 Edits in the Consumption of Alcohol Section Pertaining to the Binge Alcohol Use Variables (continued)

Issue	Edits Implemented
An edited AFU variable for binge alcohol use had a missing value. However, the R first used alcohol at his or her current age.	• The binge alcohol AFU variable with the missing value was set to the R's current age. • The corresponding editing indicator (EI) variable was set to a value of 2, where 2 = Logically assigned data. • If the YFU and MFU had been set to bad data but the unedited values that the R reported for the YFU and MFU were the same as those in ALCYFU and ALCMFU, then the relevant binge YFU and MFU variables were reset to their unedited values.
An edited YFU variable for binge alcohol use had a missing value. However, the R first used alcohol in the current survey year.	• The binge alcohol YFU variable with the missing value was set to the current year. • If the edited YFU was not the same as the unedited value, then the corresponding EI variable was set to a value of 2, where 2 = Logically assigned data. • If the MFU had a missing value and ALCMFU indicated that the R first used alcohol in the interview month, then the binge MFU was set to the same value as ALCMFU. If the edited MFU value was not the same as the unedited value, then the corresponding EI variable for the MFU (i.e., EIBINGMFU) was set to a value of 2.
An edited MFU variable for binge alcohol use had a missing value. However, the YFU data and interview date supplied information on when the R logically first engaged in binge alcohol use.	• The missing value in the MFU was replaced with the definite month in which the R logically first engaged in binge alcohol use. For example, if the R reported first use of any alcohol and first binge alcohol use in the same year and the R first used any alcohol in December of that year, then December logically was the only month in that year in which the R could have initiated binge alcohol use. • If the edited MFU value was not the same as the unedited value, then the corresponding EI variable for the MFU (i.e., EIBINGMFU) was set to a value of 2.

Table B.22 Edits Pertaining to the Special Topics Section

Issue	Edits Implemented
The respondent (R) reported never having been arrested or answered the lifetime arrest question as "don't know" or "refused" but reported being on probation or parole in the past 12 months.	The R was logically inferred to have been arrested at least once in his or her lifetime (i.e., BOOKED = 3). The rationale for this edit was that someone could not be on probation or parole without first having been arrested for a crime. The skipped variables pertaining to arrests in the past 12 months retained a value of blank.
The R reported being arrested in the past 12 months, did not report being arrested for a specific crime in that period, but reported being arrested for this crime as "some other offense."	The R was logically inferred to have been arrested for that crime. No further editing was done to the affirmative answer where the R reported being arrested for "some other offense" (BKOTH). Similarly, no further editing was done to the "OTHER, Specify" variable (BKOTHOFF) that indicated the crime for which the R was arrested (see Section 2.4.5).
The R reported being arrested at least once in the past 12 months and answered all specific past year arrest questions as "no." However, the R specified an offense in the series of follow-up questions SP03R, SP03RSP, SP03S, and SP03SSP.	The variable for "some other offense" (BKOTH) was logically inferred to be "yes." A code of 3 (i.e., Yes LOGICALLY ASSIGNED) was assigned to BKOTH. The offense that was specified in SP03RSP or SP03SSP was assigned to the edited "OTHER, Specify" variable BKOTHOFF.
The R reported being arrested at least once in the past 12 months but answered all specific past year arrest questions as "no" and reported nothing in the "OTHER, Specify" variable BKOTHOFF to support the indication of being arrested.	The response was retained to indicate that the R had been arrested in the past 12 months. A code of 5 (i.e., Offense unknown LOGICALLY ASSIGNED) was assigned to the "some other offense" variable (BKOTH).
The R reported being arrested for every offense in the past 12 months that was asked about in the section. (For youths aged 12 to 17, that included reports of being arrested for possession of tobacco; this question was skipped for adults.)	The edits differed, depending on what Rs specified for their "other" offense: • If a valid "other" offense was not specified, the entire series of past year offense variables was assigned a bad data code. • If the R gave a valid response for some other offense for which he or she was arrested in the past 12 months, the data were retained to indicate that the R was arrested for this other offense. However, the variables pertaining to arrests for all other offenses were set to bad data. • For adults, the variable pertaining to arrests for possession of tobacco (BKPOSTOB) continued to be assigned a legitimate skip code.
The R reported being arrested only one time in the past 12 months, did not report being arrested for some other offense (BKOTH = 2), but reported being arrested for every other offense in that same period.	Not including BKOTH or its associated "OTHER, Specify" variable (BKOTHOFF), the variables pertaining to arrests for specific offenses in the past 12 months were assigned a bad data code. For adults, the BKPOSTOB variable continued to be assigned a legitimate skip code.
The R reported being arrested 80 or more times in the past 12 months.	The variable pertaining to the number of arrests in the past 12 months (NOBOOKYR) was set to bad data.
The R had alternating "yes/no" or "no/yes" patterns to all questions about arrests for specific offenses in the past 12 months (e.g., SP03a = 1, SP03b = 2, SP03c = 1).	All variables pertaining to arrests for specific offenses in the past 12 months were set to bad data.

Table B.23 Edits Pertaining to the Adult Mental Health Service Utilization Variables

Issue	Edits Implemented
Respondents (Rs) did not choose an outpatient treatment location from the list of locations in question ADMT14, but that location was specified as a source of outpatient mental health treatment in the past 12 months in AUOPYRSP.	The edited variable corresponding to receipt of outpatient treatment at that location was assigned a code of 3 (Response entered LOGICALLY ASSIGNED). For example, if an R did not report receiving outpatient mental health counseling at the office of a private therapist, reported receiving outpatient counseling in "some other place," and specified something to indicate that he or she received counseling from a private therapist, the edited variable AUOPTHER was assigned a code of 3.
Rs reported receiving mental health services in every inpatient or outpatient location in a list.	For inpatient treatment, all of the variables corresponding to the service locations were set to bad data, including the numbers of nights that Rs reported spending at these various inpatient treatment settings. For outpatient treatment, the edits depended on what Rs specified for the "other" outpatient location where they received mental health treatment. • If a valid "other" outpatient location was not specified, the entire series of outpatient variables (including the reported numbers of visits) was set to bad data. • If the R reported a valid "other" outpatient location where he or she received mental health services in the past 12 months, the data were retained to indicate that the R received services in this location. However, the remaining variables pertaining to receipt of outpatient mental health treatment were set to bad data.
Rs reported at least one of the following: (a) they stayed overnight as an inpatient for mental health treatment in a particular type of facility for 365 or 366 days in the past 12 months, or (b) they stayed overnight as an inpatient in more than one type of facility, and the total number of nights that they stayed as inpatients summed to 365 or more.	If Rs reported inpatient treatment in a particular location for 366 days in the past 12 months, the corresponding edited variable (e.g., AUNMPSYH for the number of nights hospitalized in a psychiatric hospital) was reset to 365. No other editing was done when these patterns occurred.
Rs did not choose a payment source for their mental health treatment but subsequently indicated that this was (or would be) the principal payment source.	The edited payment source variable was assigned a code of 3 (Response entered LOGICALLY ASSIGNED). For example, if an R did not report that private health insurance paid or would pay for outpatient treatment but then reported that private insurance was (or would be) the principal source of payment, the edited variable AUPOPINS (private health insurance paid/will pay for any outpatient mental health treatment) was assigned a code of 3.
Rs reported a specific source of payment for their services but also reported that "No one paid because the treatment was free."	No editing was done because these responses were not necessarily inconsistent. Rs could have received services in more than one setting or from more than one provider, with some services being free and other services requiring payment.

Table B.23 Edits Pertaining to the Adult Mental Health Service Utilization Variables (continued)

Issue	Edits Implemented
Rs did not report a specific reason in question ADMT27 or ADMT27A for why they did not receive mental health treatment in the past 12 months, but they specified this as "some other reason."	The edited variable associated with that particular reason for not receiving mental health treatment was assigned a code of 3 (Response entered LOGICALLY ASSIGNED). For example, if Rs specified that they did not get mental health treatment because they could not afford the cost and they had not chosen that reason in question ADMT27, the edited variable AUUNCOST (no mental health treatment because the R could not afford the cost) was assigned a code of 3. Similarly, if Rs specified that they did not get treatment because they thought they could handle the problem without treatment but had not chosen that reason in ADMT27A, the edited variable AUUNHNDL (no mental health treatment because the R thought he or she could handle the problem without treatment) was assigned a code of 3.
Rs did not choose a particular alternative service provider from the list of providers in question ADMT29b, but that provider was specified as a source of alternative mental health treatment or support in the past 12 months.	The edited variable corresponding to receipt of alternative treatment from that type of provider was assigned a code of 3 (Response entered LOGICALLY ASSIGNED). For example, if an R did not report receiving treatment or support from a chiropractor in question ADMT29b, reported receiving treatment from some other provider, and specified something to indicate that he or she received treatment or support from a chiropractor, the edited variable AUALCHIR was assigned a code of 3.
Rs did not report receiving inpatient mental health treatment, or they reported receiving inpatient treatment but not at a location listed in question ADMT02. However, the Rs also specified treatment in a particular inpatient location as some other source of "alternative" treatment in AUALOTSP.	If a specific inpatient treatment location had not been reported in AUINPSYH through AUINRESL but the Rs specified treatment in that location in AUALOTSP, the edited inpatient variable was assigned a code of 5, where 5 = Response entered LOGICALLY ASSIGNED (from AUALOTSP). In addition, if AUINPYR, pertaining to receipt of any inpatient mental health treatment in the past 12 months, was not answered as "yes," AUINPYR was assigned a code of 3, where 3 = Yes LOGICALLY ASSIGNED.
Rs did not report receiving outpatient mental health treatment, or they reported receiving outpatient treatment but not at a location listed in question ADMT14. However, the Rs also specified treatment in a particular outpatient location as some other source of "alternative" treatment in AUALOTSP.	If a specific outpatient treatment location had not been reported in AUOPMENT through AUOPDTMT but the Rs specified treatment in that location in AUALOTSP, the edited outpatient variable was assigned a code of 5, where 5 = Response entered LOGICALLY ASSIGNED (from AUALOTSP). This code of 5 was designed to allow analysts to distinguish between logical inferences based on AUOPYRSP (see above) and those based on AUALOTSP. In contrast, if the R did not report receiving mental health treatment in an outpatient medical clinic, for example, a code of 3 in AUOPCLNC would mean that the R had not reported treatment in an outpatient medical clinic but reported receiving treatment in that location as some other outpatient location, from AUOPYRSP. In addition, if AUOPTYR, pertaining to receipt of any outpatient mental health treatment in the past 12 months, was not answered as "yes," AUOPTYR was assigned a code of 3, where 3 = Yes LOGICALLY ASSIGNED.

Table B.23 Edits Pertaining to the Adult Mental Health Service Utilization Variables (continued)

Issue	Edits Implemented
Rs did not report taking medication in the past 12 months that was prescribed for a mental health condition. However, the Rs also specified that they took medication as some other source of "alternative" treatment in AUALOTSP.	The edited variable AURXYR, pertaining to taking prescribed medication in the past 12 months, was assigned a code of 3, where 3 = Yes LOGICALLY ASSIGNED.
Rs did not report receiving treatment from any alternative providers in the past 12 months, or they reported receiving alternative treatment but not from a provider listed in question ADMT29B. However, the Rs also specified treatment from a particular alternative provider as some other source of "outpatient" treatment in AUOPYRSP.	If treatment from a specific alternative provider had not been reported in AUALACUP through AUALMASG but the Rs specified treatment from that location in AUOPYRSP, the edited alternative treatment variable was assigned a code of 5, where 5 = Response entered LOGICALLY ASSIGNED (from AUOPYRSP). This code of 5 was designed to allow analysts to distinguish between logical inferences based on AUALOTSP (see above) and those based on AUOPYRSP. For example, if the R did not report receiving mental health treatment from a chiropractor but reported receiving outpatient treatment from a chiropractor in AUOPYRSP, the edited variable AUALCHIR (mental health treatment from a chiropractor in the past 12 months) was assigned a code of 5. In comparison, a code of 3 in AUALCHIR would mean that the R had not reported treatment from a chiropractor but reported receiving that treatment from some other provider in AUALOTSP. In addition, if AUALTYR, pertaining to receipt of any alternative mental health treatment in the past 12 months, was not answered as "yes," AUALTYR was assigned a code of 3, where 3 = Yes LOGICALLY ASSIGNED.

Appendix C: Race, Hispanic/Latino Group, and Drug Codes

This page intentionally left blank

Appendix C: Race, Hispanic/Latino Group, and Drug Codes

C.1 Introduction

For the 2016 National Survey on Drug Use and Health (NSDUH), it was not uncommon for a respondent to feel that the categories for race or Hispanicity given in the questionnaire did not apply to him or her. In these situations, interviewers were given the opportunity to manually enter (type) a category that the respondent felt best described him or her. The manually entered responses were called "OTHER, Specify" or "alpha-specify" responses because they were typed in a part of the question that asked the interviewer to specify an alphabetic response. These alpha-specify responses were then matched to codes to describe the responses, which were collected and maintained in a file known as a "dictionary." Other-specify responses from each survey year were matched against this file, and any responses without codes were given new codes and added to the dictionary. Consequently, the size of the dictionary file has increased each survey year.

In most cases, new unmatched responses were just new misspellings of an already established category, such as a response of "Porto Rican" instead of "Puerto Rican." If an interviewer entered both a geographical entity and a race in the "OTHER, Specify" response, such as "Japanese Peruvian," the geographical entity was ignored and the respondent was coded as "Japanese." The geographical entity was recorded only if no other information was available, either in the "OTHER, Specify" response or in the non-OTHER, Specify response. As discussed in Chapter 4, many respondents provided a race in the alpha-specify response to the Hispanic/Latino group question and vice versa, so responses to both questions were examined in the creation of each variable. This appendix summarizes the procedures that were implemented to assign race and Hispanic/Latino values to respondents based on alpha-specify responses to the questionnaire.

Once a racial category was selected that represented the "OTHER, Specify" response, this was combined with information that was provided in the non-OTHER, Specify categories. If the information provided in the "OTHER, Specify" response was so general that formal imputation seemed to be required, and more specific information was available in the non-OTHER, Specify categories, then the final assignment of a racial category was made using only the information from the non-OTHER, Specify category(ies) and the "OTHER, Specify" information was ignored.

In addition, the data dictionary and coding of "OTHER, Specify" data for illicit drugs were revised to reflect changes to the prescription drug sections that were discussed in Section 1.4.1.5. These changes are discussed further in Section C.4 of this appendix. [Table C.5](#) also shows changes for prescription drug codes between 2014 and 2016.

C.2 Race

In the 2016 questionnaire, two questions (QD05 and QD05ASIA) focused on the respondent's race. Respondents were permitted to select more than one race in QD05. If they selected "Asian" as one of their races, they were routed to QD05ASIA, where they were

permitted to select more than one answer. Respondents had the opportunity to direct the interviewer to select "Other" as the race in both QD05 and (if applicable) QD05ASIA, whereby the interviewer then typed the alphabetic response given by the respondent. The alpha-specify responses to these two questions were considered together. Two new levels, Guamanian or Chamorro and Samoan, were added to QD05 in 2013. The race questions used in the 2016 survey were as follows:

QD05: Which of these groups describes you?

- 1 White
- 2 Black/African American
- 3 American Indian/Alaska Native (American Indian includes North American, Central American, and South American Indians)
- 4 Native Hawaiian
- 5 Guamanian or Chamorro
- 6 Samoan
- 7 Other Pacific Islander
- 8 Asian (Including: Asian Indian, Chinese, Filipino, Japanese, Korean and Vietnamese)
- 9 Other (Specify)

QD05ASIA: (Asked only if level 8 of QD05 was selected.) Which of these Asian groups describes you?

- 1 Asian Indian
- 2 Chinese
- 3 Filipino
- 4 Japanese
- 5 Korean
- 6 Vietnamese
- 7 Other (Specify)

The Hispanic/Latino group question (QD04), discussed in Section C.3, also has an "OTHER, Specify" response. Whenever race information was not available from QD05 or QD05ASIA, the response to QD04 was examined to determine whether any race information was available.

C.2.1 Race Alpha Responses

The four types of race "OTHER, Specify" responses are described in Chapter 4. Abbreviated descriptions are repeated here for convenience.

Directly Mapped Codes. Directly mapped codes were codes mapped to one or more of the categories given in the questionnaire. There were two types of directly mapped codes: (1) racial category codes and (2) geographic category codes. Racial category codes were exactly equivalent to one or more categories in QD05 or QD05ASIA. For example, a response such as "Han" mapped directly to a category in QD05ASIA (Chinese), and a response such as "mestizo"

mapped directly to two categories in QD05 (white and American Indian/Alaska Native). Geographic category codes corresponded to a country where census data indicated a racially homogeneous society. For example, an entry of "Polish" mapped to white because the Polish census data indicated that nearly all Poles were white.

Indirectly Mapped Codes. Codes that were indirectly mapped also corresponded to countries where census data were used, but for indirect mapping, the countries were racially heterogeneous. A racial category from among the 11 categories given in the questionnaire was chosen by generating a random number and allocating the race based on a comparison of the random number with the proportions of races in the geographical entity's (country's) census. For example, an entry of "St. Thomas" would have a 76.2 percent chance of being allocated to the black/African-American category, because the latest census indicated that 76.2 percent of individuals from the U.S. Virgin Islands were black. If two or three heterogeneous countries were entered in the "OTHER, Specify" response, the final race was allocated using the following procedure: (1) randomly assign races based on the proportions for each country mentioned, and (2) combine the results. Exceptions to these rules occurred with the categories Mexican, Puerto Rican, Cuban, Central or South American, Dominican, and Spanish (from Spain).

Starting with the 2006 imputation process, the handling of indirectly mapped codes obtained from QD05ASIA has been simplified. In earlier survey years, these types of write-in responses were mapped to a race through country census information. Since the 2006 NSDUH, however, all census-based write-in responses to the Asian race question were mapped directly to the "Other Asian" racial category, unless a predominant Asian group existed within the country, in which case the response was directly mapped to this category. For example, responses of "United Kingdom" to the Asian race question are mapped to "Asian Indian."

Informative Codes for Formal Imputation Procedures. Some "OTHER, Specify" responses did not lead to definitive information about the respondent's race but were used to limit the final imputation. With these informative codes, the final imputation was restricted according to the information that was available. No imputation was required if more specific information was available from responses to the non-OTHER, Specify categories. For example, a response of "mixed" resulted in an imputation among donors with two or more races, and a response of "brown" resulted in an imputation among donors who were not single-race white.

Noninformative Codes. Finally, a noninformative response that was not accompanied by a response to one of the given (non-OTHER, Specify) categories resulted in an unrestricted imputation. Religious identifications (e.g., Muslim) were considered noninformative, even if the religion was usually associated with a particular ethnic group (e.g., Shinto is usually associated with Japanese).

Table C.1 lists all the race codes used in the 2016 survey, along with supplementary information related to race codes. Special situations associated with the four types of codes described in this section are discussed in the following sections. For most codes, the final assignment depended upon whether the response was given in QD05 or QD05ASIA. For informative codes described above, the six Hispanic/Latino codes—Mexican, Puerto Rican, Cuban, Central or South American, Dominican, and Spanish (from Spain)—were treated

differently depending upon whether they were listed in conjunction with other racial or geographical entities.

Codes with an asterisk were those that caused the Hispanic/Latino indicator to be edited to a "yes." That is, if QD03 was either missing or "no," and any of these codes appeared as an "OTHER, Specify" response to QD05 or QD05ASIA, the edited Hispanic/Latino indicator (EDHOIND) was set to 1 to denote "Hispanic/Latino," and the imputation indicator for the Hispanic/Latino indicator (IIHOIND) was set to 2 to indicate "logically assigned." See Chapter 4 for more details on the edited Hispanic/Latino indicator. Note that EDHOIND also could be edited to a "no." This is discussed in Section C.3.1.

C.2.1.1 Handling of Directly Mapped Codes

For codes that were directly mapped, the final column of [Table C.1](#) indicates to which race the code was mapped. With some exceptions, the handling of directly mapped codes that were racial categories or Asian geographic categories did not depend upon whether the response was observed in QD05 or QD05ASIA. The exceptions to this rule occurred if the response included a reference to "Indian," which was mapped to "American Indian/Alaska Native" if the response was given in QD05 and to "Asian Indian" if the response was given in QD05ASIA. On the other hand, for directly mapped codes that were non-Asian geographic categories, the final mapping always depended upon whether the response was observed in QD05 or QD05ASIA. In this case, if the code was observed in QD05ASIA, the code was always mapped to "Other Asian."

Most of the directly mapped cases were mapped directly to a single category regardless of whether the response was in QD05 or QD05ASIA. However, sometimes the category to which the code was mapped in these cases is indicated only for QD05 in the final column in [Table C.1](#). In these instances, it was assumed that the directly mapped code for QD05ASIA was "Other Asian" (this is not shown in the table for space-saving reasons). For codes that corresponded to multiple-race respondents, individual Asian categories were not tracked.

In general, if the respondent selected one or more non-OTHER, Specify categories in QD05 and/or QD05ASIA, racial category codes were recognized, but geographic category codes were ignored. This is the primary difference in the handling of the two types of directly mapped codes. For example, if the interviewer selected the category for "black/African American" for the respondent and also wrote in "Polish," it was assumed that the respondent was a black Pole, and for racial identification purposes, the respondent was considered single-race black/African American. This was true even though the Polish census did not identify significant numbers of people with a race other than white in the Polish population.

C.2.1.2 Handling of Indirectly Mapped Codes

In most cases, indirectly mapped codes refer to heterogeneous countries where census data were used. In these cases, the race was assigned by comparing a randomly generated number to the proportion of each racial category in that country's census. As with the directly mapped codes, the final mapping of the indirectly mapped codes also depended upon whether the response was in QD05 or QD05ASIA, unless the heterogeneous countries listed were all Asian.

In a similar manner to the directly mapped QD05 cases, if the code was observed in QD05ASIA, it was mapped to "Other Asian," provided none of the entries observed were Asian racial categories, Asian countries, or countries with an Asian minority. (Codes that were indirectly mapped if the response was in QD05, but were directly mapped to "Other Asian" if the response was in QD05ASIA, are denoted by "QD05ASIA: O.A." in the fourth column of the table.) Codes where there was at least one Asian minority in a specified heterogeneous country that was not all Asian, and the response was given in QD05ASIA, were handled on a case-by-case basis. The resulting strategy was either a different indirect mapping than that given if the response was in QD05 or a direct mapping.

When census data were used, it was not uncommon to find that a small proportion of the population was identified as "Other." In the rare instance that the randomly generated number indicated the respondent belonged to this "Other" group, then the selected race was determined by imputation. Codes where this was possible are identified with a superscript "I" in the third column of [Table C.1](#). Rather than an "Other" indication, the census sometimes gave general information (e.g., Asian) where more specific information needed to be determined through imputation. In the case where the imputation was limited to Asian categories, the superscript "IA" was used.

Generally, if two entries (countries or racial categories) were observed, the race for each entity was determined first (either through a direct map or a random assignment using census data), and then the two races were combined. In some cases, a racial category was listed along with a geographical entity. As stated earlier, in most cases the geographical entity was ignored because it was usually assumed that the respondent was a resident of the listed country who also happened to be identified with the given racial category. However, it became clear on occasion that the respondent had parentage that belonged to the racial category and different parentage that came from the listed country. In these instances, the racial category was treated in the same manner as a homogeneous country of that race, and the determination of a final race was conducted in the same manner as if two countries had been listed. If one of the races listed was an Asian racial category, for example, then the response was treated in the same manner whether it was observed in QD05 or QD05ASIA. If the final assignment depended upon the census data of two indirectly mapped codes or an indirectly mapped code and a racial category, "double census" is parenthetically indicated in the third column of [Table C.1](#). If three indirectly mapped codes were indicated by the respondent, "triple census" is indicated.¹

Details about how to handle census information for each indirectly mapped code are shown in [Table C.2](#). Note that the racial categories for each country listed in [Table C.2](#) have been modified to conform to the racial categories specified by the questionnaire. For example, the black racial category from other countries has been modified to the black/African-American category. Every category and restricted imputation level with a nonzero probability of selection is listed. If a code had an indirect map (using census data) for QD05, but had a direct map for

¹ When an indirectly mapped code with superscript I or IA appeared as a component in a double census or triple census code, the probability associated with the "other" category was distributed among the races appearing in the census. This was the simplest way to preserve race information from all the component indirectly mapped codes. A more complicated alternative would be to impute race information for each component country, even if the "other" category was selected at random for one or more of the race categories in the census. See the entry for "Costa Rica" (code 78) in [Table C.2](#) for an example.

QD05ASIA, this is not specified in [Table C.2](#). Instead, this information must be obtained from [Table C.1](#). Explanations of the categories that are not self-explanatory are listed below.

- White or Mestizo: Imputation was restricted to respondents who were either white or Mestizo (i.e., white and American Indian/Alaska Native only). See Chapter 4 for the explanation of level 18 of EDRACE.
- Not American Indian: Imputation was restricted to respondents who were a single race other than American Indian/Alaska Native or were multiple race and American Indian/Alaska Native was not one of their component races. See Chapter 4 for the explanation of level 19 of EDRACE.
- Multiple: Imputation was restricted to respondents who were multiple race. See Chapter 4 for the explanation of level 16 of EDRACE.

C.2.1.3 Handling of Codes Informative for Formal Imputation Procedures

For six Hispanic/Latino codes that were highly prevalent in the data, census data were not used to assign the final racial category. (These are the six categories listed in QD04.) Instead, the final racial category for respondents who responded "Mexican," "Puerto Rican," "Central or South American," "Cuban," "Dominican," or "Spanish" was determined by a restricted imputation with donors who indicated one of these categories in QD04. Furthermore, if a respondent indicated any combination of these six categories, the final racial category was determined using a restricted imputation with donors who were from the geographical entities listed. On the other hand, if any of these six Hispanic/Latino groups was listed along with a second code that was not among these six, census data were used along with the census data from the second country listed. More details about how specific restricted imputations are conducted are shown in [Table C.3](#).

If the code was observed in QD05ASIA, then the imputation was not only restricted by what was written in the "OTHER, Specify" response but also by the Asian categories that had the necessary attributes. Again, the information was ignored if more specific information was available from responses to the non-OTHER, Specify categories.

C.2.1.4 Noninformative Codes

For noninformative codes, a final race could still have been assigned based on responses to other categories in QD05. If no other categories were listed in QD05, race was imputed, where the imputation was restricted to a Hispanic/Latino group if the respondent gave Hispanic/Latino information in QD04. Otherwise, the final race was determined through an unrestricted imputation.

C.3 Hispanicity

As with the race questions, Hispanic/Latino respondents² had the opportunity to specify a Hispanic/Latino group by giving the response "Other" to QD04, the Hispanic/Latino group

² For the purposes of the instrument question-routing, Hispanic/Latino respondents were identified by their response to question QD03: "Are you of Hispanic, Latino, or Spanish origin or descent?"

question. Also, respondents were permitted to select multiple Hispanic/Latino groups in response to QD04. Below is the Hispanic/Latino group question.

QD04: Which of these Hispanic, Latino, or Spanish groups best describes you?

- 1 Mexican/Mexican American/Mexicano/Chicano
- 2 Puerto Rican
- 3 Central or South American
- 4 Cuban/Cuban American
- 5 Dominican (from Dominican Republic)
- 6 Spanish (from Spain)
- 7 Other (Specify)

Levels 5 and 6 were added to QD04 after the 2004 survey. They were included because there was a large number of "OTHER, Specify" responses for these categories in previous years.

The QD05 and QD05ASIA questions are discussed in Section C.2. They also have "OTHER, Specify" responses, which were gleaned for Hispanic/Latino group information whenever no Hispanic/Latino group information was available from QD04.

C.3.1 Hispanic/Latino Group Alpha Responses

There were only two types of Hispanic/Latino group "OTHER, Specify" responses: (1) those that mapped to one or more EDQD04xx³ variables and (2) those that were ignored. There were no census-based routines and no responses that caused the imputation to be restricted. The imputation of a Hispanic/Latino group was restricted only when race information was available.

Table C.4 lists all the Hispanic/Latino group codes used in the 2016 survey and the Hispanic/Latino groups to which they mapped. Note that these mappings utilized the arbitrary priority rule provided in Chapter 4. This rule was used to create EDHOGRP, which skipped the intermediate step of recording the Hispanic/Latino groups indicated in QD04. These are recorded in the EDQD04xx variables, which are described in Chapter 4, along with the creation of EDHOGRP. The Hispanic/Latino code 600, "Stated Clearly as Not Hispanic/Latino," was unique in that it could be used to edit the Hispanic/Latino indicator, if needed. If QD03 was missing or 1, then EDHOIND was edited to a 2 if this code appeared in QD04, QD05, or QD05ASIA.

C.4 Other Drugs

As noted in Section 2.3.1.1 in Chapter 2, codes for prescription drugs were revised in 2015 in response to changes to the prescription drug questions. Prior to 2015, prescription drug questions often could apply to more than one prescription drug. For example, question PR03 in 2014 asked whether respondents ever used the pain relievers Vicodin®, Lortab®, or Lorcet® that were not prescribed for them or that they took only for the experience or feeling that these drugs caused. Code 434 for pain relievers was assigned in 2014 if respondents reported any of these

³ See Chapter 4 for a discussion of EDQD04xx.

pain relievers in the "OTHER, Specify" data or they reported similar brand name drugs that also contain hydrocodone (e.g., Anexsia[®], Co-Gesic[®], Norco[®]).

In contrast, most prescription drug questions beginning in 2015 applied to one (and only one) prescription drug. For example, respondents in 2015 were asked separately about their use or misuse in the past 12 months of the following pain relievers that contain hydrocodone: Vicodin[®], Lortab[®], Norco[®], Zohydro[®] ER, and generic hydrocodone. Also, Lorcet[®] was dropped from the 2015 questionnaire. For purposes of editing the data since 2015 (see Section 2.4.5), code 434 needs to apply to only one specific pain reliever. Beginning in 2015, therefore, code 434 has been used for mentions of Vicodin[®], and only for Vicodin[®]. Separate codes were created for Lortab[®] (code 4501) and Norco[®] (code 4502). Although the pain relievers Lorcet[®] and Anexsia[®] were not included in the 2015 questionnaire, codes also were created for these drugs (codes 4503 and 4504, respectively) to reflect the mentions of these drugs in the past and to be prepared for possible mentions of these drugs in the future. (These two drugs were not mentioned in the "OTHER, Specify" data in 2015, but Lorcet[®] was mentioned in 2016.)

The redesign of the prescription drug sections for 2015 also provided an opportunity to reassign drugs that had been given "OTHER, Specify" codes as pain relievers, tranquilizers, stimulants, or sedatives based on where the drugs had originally been mentioned but that did not belong to the particular psychotherapeutic category where they had been coded. These drugs were given new codes in the 8000 series (i.e., other drugs) beginning in 2015. For example, the antipsychotic drug Thorazine[®] had been given a code of 227 (i.e., as a tranquilizer) in 2014 and earlier years. This drug has been given a code of 8512 since 2015. Similarly, the animal tranquilizer Ketaset[®], which contains ketamine, had been coded as a tranquilizer in 2014. Since 2015, this has been coded as a hallucinogen (code 6501).

Table C.5 shows "OTHER, Specify" codes for prescription drugs in 2016 and the corresponding codes in 2014 (i.e., prior to the redesign of the prescription drug questions). In addition, Table C.5 shows "OTHER, Specify" codes that had been classified as prescription pain relievers, tranquilizers, stimulants, or sedatives in 2014 that were given codes in the 8000 series in 2016 (shown as "Other" in the table). Rows in Table C.5 that are listed as "Not Used" refer to codes from 2014 that were dropped from the coding beginning in 2015. Codes that are listed as "None" for 2014 denote new codes that have been used since 2015, without a direct equivalent in prior years. For example, code 3501 for amphetamine or dextroamphetamine salts represents the generic equivalent of Adderall[®] and does not correspond directly to code 339 for unspecified amphetamines.

The data dictionary for illicit drugs was updated to reflect the changes to the questions for prescription drugs beginning in 2015. For example, dictionary entries for Vicodin[®] continued to be associated with code 434. Entries for Lortab[®], Norco[®], Lorcet[®], and Anexsia[®] that had been associated with code 434 in 2014 were reassigned to codes 4501, 4502, 4503, and 4504, respectively. Thus, if a respondent in 2016 gave an answer anywhere in the "OTHER, Specify" data for illicit drugs that corresponded to a dictionary entry for Norco[®], then the corresponding "OTHER, Specify" variable was assigned a code of 4502.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
21	White	Directly mapped (racial category)	White
22	Black/African American	Directly mapped (racial category)	Black/African American
23	American Indian/Alaska Native	Directly mapped (racial category)	American Indian/Alaska Native
24	Native Hawaiian	Directly mapped (racial category)	Native Hawaiian
25	Other Pacific Islander	Directly mapped (racial category)	Other Pacific Islander
26	Asian Indian	Directly mapped (racial category)	Asian Indian
27	Chinese	Directly mapped (racial category)	Chinese
28	Filipino	Directly mapped (racial category)	Filipino
29	Japanese	Directly mapped (racial category)	Japanese
30	Korean	Directly mapped (racial category)	Korean
31	Vietnamese	Directly mapped (racial category)	Vietnamese
32	Other Asian	Directly mapped (racial category)	Other Asian
33	Asian (nonspecific)	Codes informative for formal imputation procedures	Not a Direct Map
34	Guamanian/Chamorro	Directly mapped (racial category)	Guamanian or Chamorro
35	Indian (Asian or American unclear)	Directly mapped (racial category)	QD05: American Indian/Alaska Native QD05ASIA: Asian Indian
50	Belize	Indirectly mapped (QD05) ¹	QD05ASIA: O.A. ²
51	Guyana	QD05: Indirectly mapped ¹ QD05ASIA: Directly mapped (geographic category)	QD05ASIA: Asian Indian
52	Suriname	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
53	Haiti	Indirectly mapped (QD05)	QD05ASIA: O.A.
54	Trinidad and Tobago	Indirectly mapped (QD05) ¹	QD05ASIA: Asian Indian
55	Jamaica	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
56	Virgin Islands (St. Thomas, St. Croix)	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
57	Bahamas	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
58	Barbados	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
59	Grenada	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
60	St. Lucia	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
61	St. Vincent & the Grenadines	Directly mapped (geographic category)	Black/African American
62	Dominica	Directly mapped (geographic category)	Black/African American
63	Other West Indies	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
64	Brazil	Indirectly mapped (QD05)	QD05ASIA: Japanese
65	Canada	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
66	Bahamas & Haiti	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
67	Brazil & Portugal	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
70	Mexico	Codes informative for formal imputation procedures	QD05ASIA: O.A.
71	Puerto Rico	Codes informative for formal imputation procedures	QD05ASIA: O.A.
72	Cuba	Codes informative for formal imputation procedures	QD05ASIA: O.A.
73	Dominican Republic	Codes informative for formal imputation procedures	QD05ASIA: O.A.
74	Guatemala	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
75	Honduras	Indirectly mapped (QD05)	QD05ASIA: O.A.
76	El Salvador	Indirectly mapped (QD05)	QD05ASIA: O.A.
77	Nicaragua	Indirectly mapped (QD05)	QD05ASIA: O.A.
78	Costa Rica	Indirectly mapped (QD05) ¹	QD05ASIA: Chinese
79	Panama	Indirectly mapped (QD05)	QD05ASIA: O.A.
80	Colombia	Indirectly mapped (QD05)	QD05ASIA: O.A.
81	Venezuela	Indirectly mapped (QD05)	QD05ASIA: O.A.
82	Ecuador	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
83	Peru	Indirectly mapped (QD05)	QD05ASIA: Japanese
84	Bolivia	Indirectly mapped (QD05)	QD05ASIA: O.A.
85	Chile	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
86	Argentina	Indirectly mapped (QD05)	QD05ASIA: O.A.
87	Paraguay	Indirectly mapped (QD05)	QD05ASIA: O.A.
88	Uruguay	Indirectly mapped (QD05)	QD05ASIA: O.A.
89	Mexico & Puerto Rico	Codes informative for formal imputation procedures	Not a Direct Map
90	Mexico & Cuba	Codes informative for formal imputation procedures	Not a Direct Map
91	Mexico & Dominican	Codes informative for formal imputation procedures	Not a Direct Map
92	Mexico & Spain	Codes informative for formal imputation procedures	Not a Direct Map
93	Puerto Rico & Cuba	Codes informative for formal imputation procedures	Not a Direct Map
94	Puerto Rico & Dominican	Codes informative for formal imputation procedures	Not a Direct Map
95	Puerto Rico & Spain	Codes informative for formal imputation procedures	Not a Direct Map
96	Cuban & Dominican	Codes informative for formal imputation procedures	Not a Direct Map

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
97	Cuban & Spain	Codes informative for formal imputation procedures	Not a Direct Map
98	Dominican & Spain	Codes informative for formal imputation procedures	Not a Direct Map
100	Norway	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
101	Sweden	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
102	Denmark	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
103	United Kingdom	Indirectly mapped (QD05) ¹	QD05ASIA: Asian Indian
104	Ireland	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
105	Portugal	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
106	Spain	Codes informative for formal imputation procedures	QD05ASIA: O.A.
107	Germany	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
108	France	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
109	Italy	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
110	Netherlands	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
111	Belgium	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
112	Greece	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
113	Russia	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
114	Ukraine	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
115	Turkey	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
116	Other Western Europe	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
117	Other Eastern Europe	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
118	Other Southern Europe	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
119	Morocco	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
120	Algeria	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
121	Tunisia	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
122	Libya	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
123	Egypt	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
124	Other North Africa	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
125	Saudi Arabia	Indirectly mapped (QD05)	QD05ASIA: O.A.
126	Yemen	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
127	Oman	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
128	UAE	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
129	Qatar	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
130	Bahrain	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
131	Israel	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
132	Iraq	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
133	Kuwait	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
134	Iran	Directly mapped (geographic category)	Other Asian
135	Other Middle East	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
136	Armenia	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
137	Georgia	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
138	Azerbaijan	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
139	Russia Asian people groups (Tatar, Chechen, Dagestan, etc.)	Directly mapped (racial category)	Other Asian

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
140	Kazakhstan	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
141	Uzbekistan	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
142	Tadjikistan	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
143	Kyrgyzstan	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
144	Turkmenistan	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
145	Other Central Asia (includes Afghanistan)	Directly mapped (geographic category)	Other Asian
150	Sri Lanka	Directly mapped (geographic category)	Other Asian
151	India	Directly mapped (geographic category)	Asian Indian
152	Other South Asia (includes Pakistan, Bangladesh, Himalayan countries)	Directly mapped (geographic category)	Other Asian
153	Burma/Myanmar	Directly mapped (geographic category)	Other Asian
154	Laos/Hmong/Iu Mienh	Directly mapped (geographic category)	Other Asian
155	Cambodia/Kampuchea	Directly mapped (geographic category)	Other Asian
156	Indonesia/Bali/Java	Directly mapped (geographic category)	Other Asian
157	Malaysia	Indirectly mapped	Not a Direct Map
158	Malay	Directly mapped (racial category)	QD05ASIA: O.A.
159	Singapore	Indirectly mapped ¹	Not a Direct Map
160	Thailand	Directly mapped (geographic category)	Other Asian
161	Thai	Directly mapped (racial category)	Other Asian
162	Mongolia	Directly mapped (geographic category)	Other Asian
163	Tibet	Directly mapped (geographic category)	Other Asian
164	Other East Asia	Directly mapped (geographic category)	Other Asian
165	Djibouti	Indirectly mapped (QD05)	QD05ASIA: O.A.
166	Sudan	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
167	Other Eastern Africa	Directly mapped (geographic category)	QD05: Black/African American QD05ASIA: Asian Indian
168	South Africa	Indirectly mapped (QD05)	QD05ASIA: Asian Indian
169	Namibia	Indirectly mapped (QD05)	QD05ASIA: O.A.
170	Zimbabwe	Indirectly mapped (QD05)	QD05ASIA: Asian Indian
171	Zambia	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
172	Botswana	Directly mapped (geographic category)	QD05: Black/African American QD05ASIA: O.A.
173	Angola	Indirectly mapped (QD05)	QD05ASIA: O.A.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
174	Mozambique	Directly mapped (geographic category)	QD05: Black/African American QD05ASIA: O.A.
175	Mauritius	Indirectly mapped (specific mapping depended upon whether response was in QD05 or QD05ASIA)	Not a Direct Map
176	Other Southern Africa	Directly mapped (geographic category)	QD05: Black/African American QD05ASIA: O.A.
177	Cape Verde	Indirectly mapped (QD05)	QD05ASIA: O.A.
178	Sao Tome	Directly mapped (geographic category)	QD05: White & Black/African American QD05ASIA: O.A.
179	Mauritania	Indirectly mapped (QD05)	QD05ASIA: O.A.
180	Mali	Indirectly mapped (QD05)	QD05ASIA: O.A.
181	Niger	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
182	Other Western Africa	Directly mapped (geographic category)	QD05: Black/African American QD05ASIA: O.A.
183	Chad	Directly mapped (geographic category)	QD05: Black/African American QD05ASIA: O.A.
184	Other Central Africa	Directly mapped (geographic category)	QD05: Black/African American QD05ASIA: O.A.
185	African/Africa	Directly mapped (geographic category)	QD05: Black/African American QD05ASIA: Asian Indian
186	Australia	Indirectly mapped (QD05) ¹	QD05ASIA: O.A.
187	New Zealand	Indirectly mapped (QD05)	QD05ASIA: O.A.
188	Fiji	Directly mapped (geographic category)	QD05: Other Pacific Islander QD05ASIA: O.A.
189	Nauru	Directly mapped (geographic category)	QD05: Other Pacific Islander QD05ASIA: O.A.
190	Samoa	Directly mapped (racial category)	QD05: Samoan QD05ASIA: O.A.
191	Samoan	Directly mapped (racial category)	QD05: Samoan QD05ASIA: O.A.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
192	Other Oceania	Directly mapped (geographic category)	QD05: Other Pacific Islander QD05ASIA: O.A.
193	European (nonspecific)	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
194	Cape Verde & Portuguese	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
195	Cape Verde & Mexican	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
196	Cape Verde & European (not Spanish)	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
201	Biracial (nonspecific)	Codes informative for formal imputation procedures	QD05ASIA: O.A.
202	White & Black/African American	Directly mapped (racial category)	QD05: White & Black/African American QD05ASIA: O.A.
203	White & American Indian/Alaska Native (including mestizo)	Directly mapped (racial category)	QD05: White & American Indian/Alaska Native QD05ASIA: O.A.
204	White & Native Hawaiian	Directly mapped (racial category)	QD05: White & Native Hawaiian QD05ASIA: O.A.
205	White & Other Pacific Islander	Directly mapped (racial category)	QD05: White & Other Pacific Islander QD05ASIA: O.A.
206	White & Asian Indian	Directly mapped (racial category)	White & Asian Indian
207	White & Chinese	Directly mapped (racial category)	White & Chinese
208	White & Filipino	Directly mapped (racial category)	White & Filipino
209	White & Japanese	Directly mapped (racial category)	White & Japanese
210	White & Korean	Directly mapped (racial category)	White & Korean
211	White & Vietnamese	Directly mapped (racial category)	White & Vietnamese
212	White & Other Asian	Directly mapped (racial category)	White & Other Asian
213	White & Asian (nonspecific)	Directly mapped (racial category)	White & Asian Nonspecific
214	White & Indian (Asian or American unclear)	Directly mapped (racial category)	QD05: White & American Indian/Alaska Native QD05ASIA: White & Asian Indian
215	White & Guamanian/Chamorro	Directly mapped (racial category)	QD05: White & Guamanian/Chamorro QD05ASIA: O.A.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
223	Black/African American & American Indian/Alaska Native	Directly mapped (racial category)	QD05: Black/African American & American Indian/Alaska Native QD05ASIA: O.A.
224	Black/African American & Native Hawaiian	Directly mapped (racial category)	QD05: Black/African American & Native Hawaiian QD05ASIA: O.A.
225	Black/African American & Other Pacific Islander	Directly mapped (racial category)	QD05: Black/African American & Other Pacific Islander QD05ASIA: O.A.
226	Black/African American & Asian Indian	Directly mapped (racial category)	Black/African American & Asian Indian
227	Black/African American & Chinese	Directly mapped (racial category)	Black/African American & Chinese
228	Black/African American & Filipino	Directly mapped (racial category)	Black/African American & Filipino
229	Black/African American & Japanese	Directly mapped (racial category)	Black/African American & Japanese
230	Black/African American & Korean	Directly mapped (racial category)	Black/African American & Korean
231	Black/African American & Vietnamese	Directly mapped (racial category)	Black/African American & Vietnamese
232	Black/African American & Other Asian	Directly mapped (racial category)	Black/African American & Other Asian
233	Black/African American & Asian (nonspecific)	Directly mapped (racial category)	Black/African American & Asian Nonspecific
234	Black/African American & Indian (Asian or American unclear)	Directly mapped (racial category)	QD05: Black/African American & American Indian QD05ASIA: Black/African American & Asian Indian
244	American Indian/Alaska Native & Native Hawaiian	Directly mapped (racial category)	QD05: American Indian/Alaska Native & Native Hawaiian QD05ASIA: O.A.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
245	American Indian/Alaska Native & Other Pacific Islander	Directly mapped (racial category)	QD05: American Indian/Alaska Native & Other Pacific Islander QD05ASIA: O.A.
246	American Indian/Alaska Native & Asian Indian	Directly mapped (racial category)	American Indian/Alaska Native & Asian Indian
247	American Indian/Alaska Native & Chinese	Directly mapped (racial category)	American Indian/Alaska Native & Chinese
248	American Indian/Alaska Native & Filipino	Directly mapped (racial category)	American Indian/Alaska Native & Filipino
249	American Indian/Alaska Native & Japanese	Directly mapped (racial category)	American Indian/Alaska Native & Japanese
250	American Indian/Alaska Native & Korean	Directly mapped (racial category)	American Indian/Alaska Native & Korean
251	American Indian/Alaska Native & Vietnamese	Directly mapped (racial category)	American Indian/Alaska Native & Vietnamese
252	American Indian/Alaska Native & Other Asian	Directly mapped (racial category)	American Indian/Alaska Native & Other Asian
253	American Indian/Alaska Native & Asian (nonspecific)	Directly mapped (racial category)	American Indian/Alaska Native & Asian Nonspecific
265	Native Hawaiian & Other Pacific Islander	Directly mapped (racial category)	QD05: Native Hawaiian & Other Pacific Islander QD05ASIA: O.A.
266	Native Hawaiian & Asian Indian	Directly mapped (racial category)	Native Hawaiian & Asian Indian
267	Native Hawaiian & Chinese	Directly mapped (racial category)	Native Hawaiian & Chinese
268	Native Hawaiian & Filipino	Directly mapped (racial category)	Native Hawaiian & Filipino
269	Native Hawaiian & Japanese	Directly mapped (racial category)	Native Hawaiian & Japanese
270	Native Hawaiian & Korean	Directly mapped (racial category)	Native Hawaiian & Korean
271	Native Hawaiian & Vietnamese	Directly mapped (racial category)	Native Hawaiian & Vietnamese

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
272	Native Hawaiian & Other Asian	Directly mapped (racial category)	Native Hawaiian & Other Asian
273	Native Hawaiian & Asian (nonspecific)	Directly mapped (racial category)	Native Hawaiian & Asian Nonspecific
286	Other Pacific Islander & Asian Indian	Directly mapped (racial category)	Other Pacific Islander & Asian Indian
287	Other Pacific Islander & Chinese	Directly mapped (racial category)	Other Pacific Islander & Chinese
288	Other Pacific Islander & Filipino	Directly mapped (racial category)	Other Pacific Islander & Filipino
289	Other Pacific Islander & Japanese	Directly mapped (racial category)	Other Pacific Islander & Japanese
290	Other Pacific Islander & Korean	Directly mapped (racial category)	Other Pacific Islander & Korean
291	Other Pacific Islander & Vietnamese	Directly mapped (racial category)	Other Pacific Islander & Vietnamese
292	Other Pacific Islander & Other Asian	Directly mapped (racial category)	Other Pacific Islander & Other Asian
293	Other Pacific Islander & Asian (nonspecific)	Directly mapped (racial category)	Other Pacific Islander & Asian Nonspecific
307	Asian Indian & Chinese	Directly mapped (racial category)	Asian Indian & Chinese
308	Asian Indian & Filipino	Directly mapped (racial category)	Asian Indian & Filipino
309	Asian Indian & Japanese	Directly mapped (racial category)	Asian Indian & Japanese
310	Asian Indian & Korean	Directly mapped (racial category)	Asian Indian & Korean
311	Asian Indian & Vietnamese	Directly mapped (racial category)	Asian Indian & Vietnamese
312	Asian Indian & Other Asian	Directly mapped (racial category)	Asian Indian & Other Asian
328	Chinese & Filipino	Directly mapped (racial category)	Chinese & Filipino
329	Chinese & Japanese	Directly mapped (racial category)	Chinese & Japanese
330	Chinese & Korean	Directly mapped (racial category)	Chinese & Korean
331	Chinese & Vietnamese	Directly mapped (racial category)	Chinese & Vietnamese
332	Chinese & Other Asian	Directly mapped (racial category)	Chinese & Other Asian
349	Filipino & Japanese	Directly mapped (racial category)	Filipino & Japanese
350	Filipino & Korean	Directly mapped (racial category)	Filipino & Korean

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
351	Filipino & Vietnamese	Directly mapped (racial category)	Filipino & Vietnamese
352	Filipino & Other Asian	Directly mapped (racial category)	Filipino & Other Asian
360	Japanese & Korean	Directly mapped (racial category)	Japanese & Korean
361	Japanese & Vietnamese	Directly mapped (racial category)	Japanese & Vietnamese
362	Japanese & Other Asian	Directly mapped (racial category)	Japanese & Other Asian
371	Korean & Vietnamese	Directly mapped (racial category)	Korean & Vietnamese
372	Korean & Other Asian	Directly mapped (racial category)	Korean & Other Asian
382	Vietnamese & Other Asian	Directly mapped (racial category)	Vietnamese & Other Asian
383	Indian (Asian or American unclear) & Native Hawaiian	Directly mapped (racial category)	QD05: American Indian/Alaska Native & Native Hawaiian QD05ASIA: Asian Indian & Native Hawaiian
384	Indian (Asian or American unclear) & Other Pacific Islander	Directly mapped (racial category)	QD05: American Indian/Alaska Native & Other Pacific Islander QD05ASIA: Asian Indian & Other Pacific Islander
385	Indian (Asian or American unclear) & Chinese	Directly mapped (racial category)	QD05: American Indian/Alaska Native & Asian QD05ASIA: Asian Indian & Chinese
386	Indian (Asian or American unclear) & Filipino	Directly mapped (racial category)	QD05: American Indian/Alaska Native & Asian QD05ASIA: Asian Indian & Filipino
387	Indian (Asian or American unclear) & Japanese	Directly mapped (racial category)	QD05: American Indian/Alaska Native & Asian QD05ASIA: Asian Indian & Japanese

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
388	Indian (Asian or American unclear) & Korean	Directly mapped (racial category)	QD05: American Indian/Alaska Native & Asian QD05ASIA: Asian Indian & Korean
389	Indian (Asian or American unclear) & Vietnamese	Directly mapped (racial category)	QD05: American Indian/Alaska Native & Asian QD05ASIA: Asian Indian & Vietnamese
390	Indian (Asian or American unclear) & Other Asian	Directly mapped (racial category)	QD05: American Indian/Alaska Native & Asian QD05ASIA: Asian Indian & Other Asian
391	Indian (Asian or American unclear) & Jamaican	Directly mapped (racial category)	QD05: American Indian/Alaska Native & Jamaican QD05ASIA: Asian Indian & Jamaican
401	White, Black/African American, American Indian/Alaska Native	Directly mapped (racial category)	QD05: White, Black/African American, American Indian QD05ASIA: O.A.
402	White, Black/African American, Native Hawaiian	Directly mapped (racial category)	QD05: White, Black/African American, Native Hawaiian QD05ASIA: O.A.
403	White, Black/African American, Other Pacific Islander	Directly mapped (racial category)	QD05: White, Black/African American, Other Pacific Islander QD05ASIA: O.A.
404	White, Black/African American, Asian Indian	Directly mapped (racial category)	White, Black/African American, Asian Indian
405	White, Black/African American, Chinese	Directly mapped (racial category)	White, Black/African American, Chinese
406	White, Black/African American, Filipino	Directly mapped (racial category)	White, Black/African American, Filipino
407	White, Black/African American, Japanese	Directly mapped (racial category)	White, Black/African American, Japanese
408	White, Black/African American, Korean	Directly mapped (racial category)	White, Black/African American, Korean

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
409	White, Black/African American, Vietnamese	Directly mapped (racial category)	White, Black/African American, Vietnamese
410	White, Black/African American, Other Asian	Directly mapped (racial category)	White, Black/African American, Other Asian
411	White, Black/African American, Asian (nonspecific)	Directly mapped (racial category)	White, Black/African American, Asian Nonspecific
412	White, American Indian/Alaska Native, Native Hawaiian	Directly mapped (racial category)	QD05: White, American Indian/Alaska Native, Native Hawaiian QD05ASIA: O.A.
413	White, American Indian/Alaska Native, Other Pacific Islander	Directly mapped (racial category)	QD05: White, American Indian/Alaska Native, Other Pacific Islander QD05ASIA: O.A.
414	White, American Indian/Alaska Native, Asian Indian	Directly mapped (racial category)	White, American Indian/Alaska Native, Asian Indian
415	White, American Indian/Alaska Native, Chinese	Directly mapped (racial category)	White, American Indian/Alaska Native, Chinese
416	White, American Indian/Alaska Native, Filipino	Directly mapped (racial category)	White, American Indian/Alaska Native, Filipino
417	White, American Indian/Alaska Native, Japanese	Directly mapped (racial category)	White, American Indian/Alaska Native, Japanese
418	White, American Indian/Alaska Native, Korean	Directly mapped (racial category)	White, American Indian/Alaska Native, Korean
419	White, American Indian/Alaska Native, Vietnamese	Directly mapped (racial category)	White, American Indian/Alaska Native, Vietnamese
420	White, American Indian/Alaska Native, Other Asian	Directly mapped (racial category)	White, American Indian/Alaska Native, Other Asian
421	White, American Indian/Alaska Native, Asian (nonspecific)	Directly mapped (racial category)	White, American Indian/Alaska Native, Asian Nonspecific

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
422	White, Black/African American, Indian (Asian or American unclear)	Directly mapped (racial category)	QD05: White, Black/African American, American Indian/Alaska Native QD05ASIA: White, Black/African American, Asian Indian
423	White, Native Hawaiian, Indian (Asian or American unclear)	Directly mapped (racial category)	QD05: White, Native Hawaiian, American Indian/Alaska Native QD05ASIA: White, Native Hawaiian, Asian Indian
424	White, Other Pacific Islander, Indian (Asian or American unclear)	Directly mapped (racial category)	QD05: White, Other Pacific Islander, American Indian/Alaska Native QD05ASIA: White, Other Pacific Islander, Asian
600	Mexican & Guatemalan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
601	Mexican & El Salvadoran	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
602	Mexican & Honduran	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
603	Mexican & Nicaraguan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
604	Mexican & Costa Rican	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
605	Mexican & Panamanian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
606	Mexican & Colombian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
607	Mexican & Venezuelan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
608	Mexican & Ecuadorian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
609	Mexican & Peruvian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
610	Mexican & Bolivian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
611	Mexican & Chilean	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
612	Mexican & Argentine	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
613	Mexican & Paraguayan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
614	Mexican & Uruguayan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
615	Mexican & Brazilian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
616	Puerto Rican & Guatemalan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
617	Puerto Rican & El Salvadoran	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
618	Puerto Rican & Honduran	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
619	Puerto Rican & Nicaraguan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
620	Puerto Rican & Costa Rican	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
621	Puerto Rican & Panamanian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
622	Puerto Rican & Colombian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
623	Puerto Rican & Venezuelan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
624	Puerto Rican & Ecuadorian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
625	Puerto Rican & Peruvian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
626	Puerto Rican & Bolivian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
627	Puerto Rican & Chilean	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
628	Puerto Rican & Argentine	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
629	Puerto Rican & Paraguayan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
630	Puerto Rican & Uruguayan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
631	Puerto Rican & Brazilian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
632	Cuban & Guatemalan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
633	Cuban & El Salvadoran	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
634	Cuban & Honduran	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
635	Cuban & Nicaraguan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
636	Cuban & Costa Rican	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
637	Cuban & Panamanian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
638	Cuban & Colombian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
639	Cuban & Venezuelan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
640	Cuban & Ecuadorian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
641	Cuban & Peruvian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
642	Cuban & Bolivian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
643	Cuban & Chilean	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
644	Cuban & Argentine	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
645	Cuban & Paraguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
646	Cuban & Uruguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
647	Cuban & Brazilian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
648	Dominican & Guatemalan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
649	Dominican & El Salvadoran	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
650	Dominican & Honduran	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
651	Dominican & Nicaraguan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
652	Dominican & Costa Rican	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
653	Dominican & Panamanian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
654	Dominican & Colombian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
655	Dominican & Venezuelan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
656	Dominican & Ecuadorian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
657	Dominican & Peruvian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
658	Dominican & Bolivian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
659	Dominican & Chilean	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
660	Dominican & Argentine	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
661	Dominican & Paraguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
662	Dominican & Uruguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
663	Dominican & Brazilian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
664	Spanish (from Spain) & Guatemalan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
665	Spanish (from Spain) & El Salvadoran	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
666	Spanish (from Spain) & Honduran	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
667	Spanish (from Spain) & Nicaraguan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
668	Spanish (from Spain) & Costa Rican	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
669	Spanish (from Spain) & Panamanian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
670	Spanish (from Spain) & Colombian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
671	Spanish (from Spain) & Venezuelan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
672	Spanish (from Spain) & Ecuadorian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
673	Spanish (from Spain) & Peruvian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
674	Spanish (from Spain) & Bolivian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
675	Spanish (from Spain) & Chilean	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
676	Spanish (from Spain) & Argentine	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
677	Spanish (from Spain) & Paraguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
678	Spanish (from Spain) & Uruguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
679	Spanish (from Spain) & Brazilian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
680	Colombian & Venezuelan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
681	Colombian & Ecuadorian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
682	Colombian & Peruvian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
683	Colombian & Bolivian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
684	Colombian & Chilean	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
685	Colombian & Argentine	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
686	Colombian & Paraguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
687	Colombian & Uruguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
688	Colombian & Brazilian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
689	Venezuelan & Ecuadoran	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
690	Venezuelan & Peruvian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
691	Venezuelan & Bolivian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
692	Venezuelan & Chilean	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
693	Venezuelan & Argentine	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
694	Venezuelan & Paraguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
695	Venezuelan & Uruguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
696	Venezuelan & Brazilian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
697	Ecuadoran & Peruvian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
698	Ecuadoran & Bolivian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
699	Ecuadoran & Chilean	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
700	Ecuadoran & Argentine	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
701	Ecuadoran & Paraguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
702	Ecuadoran & Uruguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
703	Ecuadoran & Brazilian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
704	Peruvian & Bolivian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
705	Peruvian & Chilean	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
706	Peruvian & Argentine	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
707	Peruvian & Paraguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
708	Peruvian & Uruguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
709	Peruvian & Brazilian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
710	Bolivian & Chilean	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
711	Bolivian & Argentine	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
712	Bolivian & Paraguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
713	Bolivian & Uruguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
714	Bolivian & Brazilian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
715	Chilean & Argentine	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
716	Chilean & Paraguayan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
717	Chilean & Uruguayan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
718	Chilean & Brazilian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
719	Argentine & Paraguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
720	Argentine & Uruguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
721	Argentine & Brazilian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
722	Paraguayan & Uruguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
723	Paraguayan & Brazilian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
724	Uruguayan & Brazilian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
725	Guatemalan & Colombian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
726	El Salvadoran & Colombian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
727	Honduran & Colombian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
728	Nicaraguan & Colombian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
729	Costa Rican & Colombian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
730	Panamanian & Colombian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
731	Guatemalan & Venezuelan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
732	El Salvadoran & Venezuelan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
733	Honduran & Venezuelan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
734	Nicaraguan & Venezuelan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
735	Costa Rican & Venezuelan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
736	Panamanian & Venezuelan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
737	Guatemalan & Ecuadorean	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
738	El Salvadoran & Ecuadorean	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
739	Honduran & Ecuadorean	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
740	Nicaraguan & Ecuadorean	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
741	Costa Rican & Ecuadorean	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
742	Panamanian & Ecuadorean	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
743	Guatemalan & Peruvian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
744	El Salvadoran & Peruvian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
745	Honduran & Peruvian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
746	Nicaraguan & Peruvian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
747	Costa Rican & Peruvian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
748	Panamanian & Peruvian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
749	Guatemalan & Bolivian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
750	El Salvadoran & Bolivian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
751	Honduran & Bolivian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
752	Nicaraguan & Bolivian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
753	Costa Rican & Bolivian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
754	Panamanian & Bolivian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
755	Guatemalan & Chilean	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
756	El Salvadoran & Chilean	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
757	Honduran & Chilean	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
758	Nicaraguan & Chilean	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
759	Costa Rican & Chilean	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
760	Panamanian & Chilean	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
761	Guatemalan & Argentine	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
762	El Salvadoran & Argentine	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
763	Honduran & Argentine	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
764	Nicaraguan & Argentine	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
765	Costa Rican & Argentine	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
766	Panamanian & Argentine	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
767	Guatemalan & Paraguayan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
768	El Salvadoran & Paraguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
769	Honduran & Paraguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
770	Nicaraguan & Paraguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
771	Costa Rican & Paraguayan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
772	Panamanian & Paraguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
773	Guatemalan & Uruguayan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
774	El Salvadoran & Uruguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
775	Honduran & Uruguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
776	Nicaraguan & Uruguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
777	Costa Rican & Uruguayan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
778	Panamanian & Uruguayan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
779	Guatemalan & Brazilian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
780	El Salvadoran & Brazilian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
781	Honduran & Brazilian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
782	Nicaraguan & Brazilian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
783	Costa Rican & Brazilian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
784	Panamanian & Brazilian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
785	Guatemalan & El Salvadoran	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
786	Guatemalan & Honduran	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
787	Guatemalan & Nicaraguan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
788	Guatemalan & Costa Rican	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
789	Guatemalan & Panamanian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
790	El Salvadoran & Honduran	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
791	El Salvadoran & Nicaraguan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
792	El Salvadoran & Costa Rican	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
793	El Salvadoran & Panamanian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
794	Honduran & Nicaraguan	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
795	Honduran & Costa Rican	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
796	Honduran & Panamanian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
797	Nicaraguan & Costa Rican	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
798	Nicaraguan & Panamanian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
799	Costa Rican & Panamanian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
800	Mexican & Jamaican	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
801	Puerto Rican & Jamaican	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
802	Cuban & Jamaican	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
803	Dominican & Jamaican	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
804	Spanish (from Spain) & Jamaican	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
805	Mexican & European (not Spanish)	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
806	Puerto Rican & European (not Spanish)	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
807	Cuban & European (not Spanish)	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
808	Dominican & European (not Spanish)	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
809	Spanish (from Spain) & Other European	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
810	Trinidadian & Mexican	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
811	Trinidadian & Puerto Rican	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
812	Trinidadian & Cuban	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
813	Trinidadian & Dominican	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
814	Trinidadian & Spanish (from Spain)	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
815	Guatemalan & European (not Spanish)	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
816	El Salvador & European (not Spanish)	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
817	Honduran & European (not Spanish)	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
818	Nicaraguan & European (not Spanish)	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
819	Costa Rican & European (not Spanish)	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
820	Panamanian & European (not Spanish)	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
821	Colombian & European (not Spanish)	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
822	Venezuelan & European (not Spanish)	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
823	Ecuadoran & European (not Spanish)	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
824	Peruvian & European (not Spanish)	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
825	Bolivian & European (not Spanish)	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
826	Chilean & European (not Spanish)	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
827	Argentine & European (not Spanish)	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
828	Paraguayan & European (not Spanish)	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
829	Uruguayan & European (not Spanish)	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
830	Brazilian & European (not Spanish)	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
831	(part) Mexican, ½ (part) White	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
832	(part) Mexican, ½ (part) Black/African American	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
833	(part) Mexican, ½ (part) American Indian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
834	(part) Mexican, ½ (part) Hawaiian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
835	(part) Mexican, ½ (part) Other Pacific Islander	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
836	(part) Mexican, ½ (part) Asian Indian	Indirectly mapped (QD05) (double census) ¹	Not a Direct Map
837	(part) Mexican, ½ (part) Chinese	Indirectly mapped (QD05) (double census) ¹	Not a Direct Map
838	(part) Mexican, ½ (part) Filipino	Indirectly mapped (QD05) (double census) ¹	Not a Direct Map
839	(part) Mexican, ½ (part) Japanese	Indirectly mapped (QD05) (double census) ¹	Not a Direct Map
840	(part) Mexican, ½ (part) Korean	Indirectly mapped (QD05) (double census) ¹	Not a Direct Map
841	(part) Mexican, ½ (part) Vietnamese	Indirectly mapped (QD05) (double census) ¹	Not a Direct Map
842	(part) Mexican, ½ (part) Other Asian	Indirectly mapped (QD05) (double census) ¹	Not a Direct Map
843	(part) Puerto Rican, ½ (part) White	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
844	(part) Puerto Rican, ½ (part) Black/African American	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
845	(part) Puerto Rican, ½ (part) American Indian/Alaska Native	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
846	(part) Puerto Rican, ½ (part) Native Hawaiian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
847	(part) Puerto Rican, ½ (part) Other Pacific Islander	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
848	(part) Puerto Rican, ½ (part) Asian Indian	Indirectly mapped (QD05) (double census) ¹	Not a Direct Map
849	(part) Puerto Rican, ½ (part) Chinese	Indirectly mapped (QD05) (double census) ¹	Not a Direct Map
850	(part) Puerto Rican, ½ (part) Filipino	Indirectly mapped (QD05) (double census) ¹	Not a Direct Map
851	(part) Puerto Rican, ½ (part) Japanese	Indirectly mapped (QD05) (double census) ¹	Not a Direct Map
852	(part) Puerto Rican, ½ (part) Korean	Indirectly mapped (QD05) (double census) ¹	Not a Direct Map

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
853	(part) Puerto Rican, ½ (part) Vietnamese	Indirectly mapped (QD05) (double census) ¹	Not a Direct Map
854	(part) Puerto Rican, ½ (part) Other Asian	Indirectly mapped (QD05) (double census) ¹	Not a Direct Map
855*	(part) Hispanic/Latino, ½ (part) White	Directly mapped (racial category)	QD05: White QD05ASIA: O.A.
856*	(part) Hispanic/Latino, ½ (part) Black/African American	Directly mapped (racial category)	QD05: Black/African American QD05ASIA: O.A.
857*	(part) Hispanic/Latino, ½ (part) American Indian/Alaska Native	Directly mapped (racial category)	QD05: American Indian/Alaska Native QD05ASIA: O.A.
858*	(part) Hispanic/Latino, ½ (part) Native Hawaiian	Directly mapped (racial category)	QD05: Native Hawaiian QD05ASIA: O.A.
859*	(part) Hispanic/Latino, ½ (part) Other Pacific Islander	Directly mapped (racial category)	QD05: Other Pacific Islander QD05ASIA: O.A.
860*	(part) Hispanic/Latino, ½ (part) Asian Indian	Directly mapped (racial category)	Asian Indian
861*	(part) Hispanic/Latino, ½ (part) Chinese	Directly mapped (racial category)	Chinese
862*	(part) Hispanic/Latino, ½ (part) Filipino	Directly mapped (racial category)	Filipino
863*	(part) Hispanic/Latino, ½ (part) Japanese	Directly mapped (racial category)	Japanese
864*	(part) Hispanic/Latino, ½ (part) Korean	Directly mapped (racial category)	Korean
865*	(part) Hispanic/Latino, ½ (part) Vietnamese	Directly mapped (racial category)	Vietnamese
866*	(part) Hispanic/Latino, ½ (part) Other Asian	Directly mapped (racial category)	Other Asian

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
869	Haitian & Jamaican	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
870	Haitian & Dominican	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
871	Honduran & Haitian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
872	Guatemalan & Iranian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
873	Panamanian & Jamaican	Indirectly mapped (QD05) (double census) ^I	QD05ASIA: O.A.
874	Cuban & Thai	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
875	Venezuelan & Trinidadian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
876	Puerto Rican & Arab	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
877	Puerto Rican & Virgin Islands	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
878	Mexican & Samoan	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
879	Salvadoran & Egyptian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
880	Costa Rican & Haitian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
881	Mexican & Iranian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
882	Spanish & Barbadian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
883	Peruvian & Other Middle Eastern	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
884	Puerto Rican & African	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
885	Jamaican & Egyptian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
886	Argentine & Turkish	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
887	Mexican & Egyptian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
888	Guatemalan & Canadian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
889	Haitian & European	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
890	Argentine, Cuban, & Spanish	Indirectly mapped (QD05) (triple census)	QD05ASIA: O.A.
891	Mexican, Cuban, & French	Indirectly mapped (QD05) (triple census) ¹	QD05ASIA: O.A.
892	Mexican, Puerto Rican, & European	Indirectly mapped (QD05) (triple census) ¹	QD05ASIA: O.A.
893	Haitian & Trinidadian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
894	Belizean & Honduran	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
895	Trinidadian & European	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
896	Puerto Rican & Haitian	Indirectly mapped (QD05) (double census)	QD05ASIA: O.A.
900*	Definitely Hispanic/Latino (Hispanic, Latino/a, Chicano/a, etc., not Spain or Dominican Republic)	Codes informative for formal imputation procedures	Not a Direct Map
901*	Definitely Hispanic/Latino (Hispanic Spanish, Español, etc.)	Codes informative for formal imputation procedures	Not a Direct Map
902*	Definitely Hispanic/Latino (Hispanic Dominican Republic, Dominicano, etc.)	Indirectly mapped (QD05)	QD05ASIA: O.A.
903	Central/South American (no country)	Codes informative for formal imputation procedures	Not a Direct Map
904	Nonwhite (nonspecific/brown)	Codes informative for formal imputation procedures	Not a Direct Map
905*	Hispanic/Latino nonwhite (including trigueno = "dark," moreno)	Codes informative for formal imputation procedures	Not a Direct Map
906*	Mezclado, Mezclada (Hispanic/Latino mixed)	Codes informative for formal imputation procedures	Not a Direct Map
907	Mixed	Codes informative for formal imputation procedures	Not a Direct Map
908	Olive	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
909	Creole	Indirectly mapped	QD05ASIA: O.A.
910	Arab	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
911	Jewish	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.
912	Kurd	Directly mapped (geographic category)	Other Asian
913	Chaldean/Caldanian/ Assyrian	Directly mapped (geographic category)	Other Asian
914	Romany/Gypsy	Directly mapped (geographic category)	QD05: White QD05ASIA: O.A.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
915	Central/South American & West Indies	Indirectly mapped	QD05ASIA: O.A.
916	Central/South American & Mexican	Codes informative for formal imputation procedures	QD05ASIA: O.A.
917	Central/South American & Puerto Rican	Codes informative for formal imputation procedures	QD05ASIA: O.A.
918	Central/South American & Cuban	Codes informative for formal imputation procedures	QD05ASIA: O.A.
919	Central/South American & Dominican	Codes informative for formal imputation procedures	QD05ASIA: O.A.
920	Central/South American & Spanish	Codes informative for formal imputation procedures	QD05ASIA: O.A.
921	Arab/Asian	QD05: Directly mapped (racial category) QD05ASIA: Directly mapped (geographic category)	QD05: White & Asian QD05ASIA: Other Asian
922	Arab/European	Directly mapped (geographic category)	QD05: White QD05ASIA: White & Other Asian
923	Arab/African	Directly mapped (geographic category)	QD05: White & Black/African American QD05ASIA: Other Asian & Black/African American
924	Arab/Chaldean	Directly mapped (racial category)	QD05: White & Other Asian QD05ASIA: O.A.
925	European & Asian Indian	Directly mapped (racial category)	White & Asian Indian
926	West Indies & Belize	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
927	West Indies & Cape Verde	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
928	Arab & Cape Verde	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
929	Aryan	Directly mapped (geographic category)	QD05: White QD05ASIA: Asian Indian
930	Turkish & Lebanese	Directly mapped (racial category)	White
932	Puerto Rican & Dominican	Codes informative for formal imputation procedures	QD05ASIA: O.A.
933	Puerto Rican & Spanish (from Spain)	Codes informative for formal imputation procedures	QD05ASIA: O.A.

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
934	Dominican & Asian Indian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
935	Cuban & Spanish (from Spain)	Codes informative for formal imputation procedures	QD05ASIA: O.A.
936	Dominican & Spanish (from Spain)	Codes informative for formal imputation procedures	QD05ASIA: O.A.
937	Cuban & Indian (Asian or American unclear)	Directly mapped (racial category)	QD05: American Indian/Alaska Native & Cuban QD05ASIA: Asian Indian & Cuban
938	Cuban & American Indian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
939	Cuban & Asian Indian	Indirectly mapped (QD05) (double census) ¹	Not a Direct Map
940	Thai & Malaysian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
941	Jamaican & American Indian	Indirectly mapped (QD05) (double census) ¹	QD05ASIA: O.A.
942	Jamaican & Asian Indian	Indirectly mapped (QD05) (double census) ¹	Not a Direct Map
951	White and something else	Directly mapped (racial category)	QD05: White (Multiple Race) QD05ASIA: O.A.
952	Black/African American and something else	Directly mapped (racial category)	QD05: Black/African American (Multiple Race) QD05ASIA: O.A.
953	American Indian/Alaska Native and something else	Directly mapped (racial category)	QD05: American Indian/Alaska Native (Multiple Race) QD05ASIA: O.A.
954	Native Hawaiian and something else	Directly mapped (racial category)	QD05: Native Hawaiian (Multiple Race) QD05ASIA: O.A.
955	Other Pacific Islander and something else	Directly mapped (racial category)	QD05: Other Pacific Islander (Multiple Race) QD05ASIA: O.A.
956	Asian Indian and something else	Directly mapped (racial category)	Asian Indian (Multiple Race)
957	Chinese and something else	Directly mapped (racial category)	Chinese (Multiple Race)
958	Filipino and something else	Directly mapped (racial category)	Filipino (Multiple Race)

Table C.1 Descriptions of Race Codes and the Categories to Which They Mapped (continued)

Race Code	Race Name	Type¹	Category to Which Race Code Directly Mapped
959	Japanese and something else	Directly mapped (racial category)	Japanese (Multiple Race)
960	Korean and something else	Directly mapped (racial category)	Korean (Multiple Race)
961	Vietnamese and something else	Directly mapped (racial category)	Vietnamese (Multiple Race)
962	Other Asian and something else	Directly mapped (racial category)	Other Asian (Multiple Race)
963	Asian (nonspecific) and something else	Codes useful for formal imputation procedures	Not a Direct Map
964	Indian (Asian or American unclear) and something else	Directly mapped (racial category)	QD05: American Indian/Alaska Native (Multiple Race) QD05ASIA: Asian Indian (Multiple Race)
965	Brown & White	Directly mapped (racial category)	White & Black/African American
985	Bad data	Noninformative code	Not a Direct Map
994	Unknown/"Don't Know"	Noninformative code	Not a Direct Map
997	Rather Not Say/"Refused" ("American" or "All of Them")	Noninformative code	Not a Direct Map

* These codes caused the Hispanic/Latino indicator to be edited to a "yes" if QD03 was missing or "no." The code that caused the Hispanic/Latino indicator to be edited to a "no" was a Hispanic/Latino code (600) and is listed in [Table C.4](#).

¹ Among the indirectly mapped codes, codes where an imputation was possible based on census information are indicated by the superscript "I." See Section C.2.1.2 for details.

² The abbreviation "O.A." is equivalent to "Other Asian."

Table C.2 Proportional Racial Allocations for Indirectly Mapped Codes

Race Code	Race Name	Probabilities
50	Belize	6.1% Black/African American, 10.6% American Indian/Alaska Native, 24.9% White and Black/African American, 48.7% White and American Indian/Alaska Native, 9.7% Unrestricted Imputation QD05 when in combination with another race: 6.1% Black/African American, 13.8% American Indian/Alaska Native, 28.1% White and Black/African American, 51.9% White and American Indian/Alaska Native
51	Guyana	QD05: 36% Black/African American, 7% American Indian/Alaska Native, 50% Asian Indian, 7% Unrestricted Imputation
52	Suriname	QD05: 1% White, 10% Black/African American, 2% American Indian/Alaska Native, 37% Asian Indian, 2% Chinese, 15% Other Asian, 31% White and Black/African American, 2% Unrestricted Imputation
53	Haiti	95% Black/African American, 5% White and Black/African American
54	Trinidad and Tobago	QD05: 34.2% Black/African American, 35.4% Asian Indian, 7.7% Black/African American and Asian Indian, 15.3% Impute among those with multiple races, 7.4% Unrestricted Imputation QD05 when in combination with another race: 36.7% Black/African American, 37.9% Asian Indian, 25.5% Black/African American and Asian Indian
55	Jamaica	QD05: 92.1% Black/African American, 0.8% Asian Indian, 6.1% White and Black/African American, 1% Unrestricted Imputation QD05 when in combination with another race: 92.4% Black/African American, 1.1% Asian Indian, 6.4% White and Black/African American
56	Virgin Is (St Thomas, St Croix)	QD05: 13.1% White, 76.2% Black/African American, 1.1% Asian Nonspecific, 3.5% White and Black/African American, 6.1% Unrestricted Imputation QD05 when in combination with another race: 15.1% White, 78.2% Black/African American, 1.1% Asian Nonspecific, 5.5% White and Black/African American
57	Bahamas	QD05: 4.7% White, 90.6% Black/African American, 2.1% White and Black/African American, 2.6% Unrestricted Imputation
58	Barbados	2.7% White, 92.4% Black/African American, 3.1% White and Black/African American, 1.3% Asian Indian, 0.5% Unrestricted Imputation QD05 when in combination with another race: 2.9% White, 92.6% Black/African American, 3.3% White and Black/African American, 1.3% Asian Indian
59	Grenada	82% Black/African American, 13% White and Black/African American, 5% Unrestricted Imputation
60	St. Lucia	QD05: 85.3% Black/African American, 2.2% Asian Indian, 5.45% White and Black/African American, 5.45% Black/African American and Asian Indian, 1.7% Unrestricted Imputation

Table C.2 Proportional Racial Allocations for Indirectly Mapped Codes (continued)

Race Code	Race Name	Probabilities
63	Other West Indies	QD05: 80% Black/African American, 14% Asian Nonspecific, 6% Unrestricted Imputation QD05 when in combination with another race: 85.1% Black/African American, 14.9% Asian Nonspecific
64	Brazil	47.7% White, 7.6% Black/African American, 0.5% American Indian/Alaska Native, 43.1% White and Black/African American, 1.1% Asian Nonspecific
65	Canada	QD05: 91.3% White, 4.2% American Indian/Alaska Native, 4.5% Chinese
70	Mexico	QD05 when in combination with another race: 9.3% White, 30.3% American Indian/Alaska Native, 60.3% White and American Indian/Alaska Native
71	Puerto Rico	QD05 when in combination with another race: 75.8% White, 12.4% Black/African American, 2.8% American Indian/Alaska Native, 2.8% Native Hawaiian, 2.8% Other Pacific Islander, 1.1% White and Black/African American, 1.1% White and American Indian/Alaska Native, 1.1% Black/African American and American Indian/Alaska Native
72	Cuba	QD05 when in combination with another race: 64.1% White, 9.3% Black/African American, 26.6% White and American Indian/Alaska Native
73	Dominican Republic	QD05 when in combination with another race: 16% White, 11% Black/African American, 73% White and Black/African American
74	Guatemala	40.5% American Indian/Alaska Native, 59.4% White and American Indian/Alaska Native, 0.1% Unrestricted Imputation QD05 when in combination with another race: 40.5% American Indian/Alaska Native, 59.5% White and American Indian/Alaska Native
75	Honduras	1% White, 2% Black/African American, 7% American Indian/Alaska Native, 90% White and American Indian/Alaska Native
76	El Salvador	12.7% White, 1% American Indian/Alaska Native, 86.3% White and American Indian/Alaska Native
77	Nicaragua	17% White, 9% Black/African American, 5% American Indian/Alaska Native, 69% White and American Indian/Alaska Native
78	Costa Rica	QD05: 1.1% Black/African American, 2.4% American Indian/Alaska Native, 6.7% White and Black/African American, 83.6% Imputed among White or White and American Indian, 6.2% Unrestricted Imputation QD05 when in combination with another race: 43.9% White, 1.1% Black/African American, 2.4% American Indian/Alaska Native, 8.8% White and Black/African American, 43.9% White and American Indian/Alaska Native
79	Panama	10% White, 14% Black/African American, 6% American Indian/Alaska Native, 70% White and American Indian/Alaska Native

Table C.2 Proportional Racial Allocations for Indirectly Mapped Codes (continued)

Race Code	Race Name	Probabilities
80	Colombia	20% White, 4% Black/African American, 1% American Indian/Alaska Native, 14% White and Black/African American, 58% White and American Indian/Alaska Native, 3% Black/African American and American Indian/Alaska Native
81	Venezuela	21% White, 10% Black/African American, 2% American Indian/Alaska Native, 67% White and American Indian/Alaska Native
82	Ecuador	6.1% White, 7.2% Black/African American, 7% American Indian/Alaska Native, 79.3% White and American Indian/Alaska Native, 0.4% Unrestricted Imputation
83	Peru	15% White, 1% Black/African American, 45% American Indian/Alaska Native, 1% Chinese, 1% Japanese, 37% White and American Indian/Alaska Native
84	Bolivia	15% White, 55% American Indian/Alaska Native, 30% White and American Indian/Alaska Native
85	Chile	10.8% American Indian/Alaska Native, 88.9% White and American Indian/Alaska Native, 0.3% Unrestricted Imputation QD05 when in combination with another race: 10.8% American Indian/Alaska Native, 89.2% White and American Indian/Alaska Native
86	Argentina	97% White, 3% White and American Indian/Alaska Native
87	Paraguay	2.5% White, 2.5% American Indian/Alaska Native, 95% White and American Indian/Alaska Native
88	Uruguay	88% White, 4% Black/African American, 8% White and American Indian/Alaska Native
103	United Kingdom	87.2% White, 3% Black/African American, 4.2% Asian Indian, 2% Impute Among Multiple Race, 3.5% Unrestricted Imputation
125	Saudi Arabia	QD05: 90% White, 10% Asian Indian
128	UAE	QD05: 30.5% White, 50% Asian Indian, 11.5% Other Asian, 8% Not American Indian/Alaska Native
129	Qatar	QD05: 40% White, 36% Asian Indian, 10% Other Asian, 14% Not American Indian/Alaska Native
130	Bahrain	QD05: 51.7% White, 1.6% Black/African American, 45.5% Asian Indian, 1.2% Impute among those who are not American Indian/Alaska Native
140	Kazakhstan	26.9% White, 68.6% Other Asian, 4.5% Impute among those who are not American Indian/Alaska Native
141	Uzbekistan	5.5% White, 92% Other Asian, 2.5% Impute among those who are not American Indian/Alaska Native
142	Tadjikistan	1.1% White, 96.3% Other Asian, 2.6% Impute among those who are not American Indian/Alaska Native
143	Kyrgyzstan	13.5% White, 80.8% Other Asian, 5.7% Impute among those who are not American Indian/Alaska Native
144	Turkmenistan	4.0% White, 90.0% Other Asian, 6.0% Impute among those who are not American Indian/Alaska Native

Table C.2 Proportional Racial Allocations for Indirectly Mapped Codes (continued)

Race Code	Race Name	Probabilities
157	Malaysia	6.7% Asian Indian, 22.6% Chinese, 61.9% Other Asian, 8.8% Impute among those who are not American Indian/Alaska Native
159	Singapore	9.2% Asian Indian, 74.2% Chinese, 13.3% Other Asian, 3.3% Impute among those who are not American Indian/Alaska Native
165	Djibouti	2.5% White, 97.5% Black/African American
166	Sudan	39% White, 58% Black/African American, 3% Impute among those who are not American Indian/Alaska Native
168	South Africa	8.9% White, 79.2% Black/African American, 2.5% Asian Indian, 8.9% White and Black/African American, 0.5% Impute among those who are not American Indian/Alaska Native
169	Namibia	6% White, 87.5% Black/African American, 6.5% White and Black/African American
170	Zimbabwe	1% White, 98% Black/African American, .5% Asian Indian, .5% White and Black/African American
171	Zambia	99% Black/African American, 1% Impute among those who are not American Indian/Alaska Native
173	Angola	1% White, 97% Black/African American, 2% White and Black/African American
175	Mauritius	QD05: 2% White, 68% Asian Indian, 3% Chinese, 27% White and Black/African American
177	Cape Verde	1% White, 28% Black/African American, 71% White and Black/African American
179	Mauritania	30% White, 30% Black/African American, 40% White and Black/African American
180	Mali	10% White, 85% Black/African American, 5% Unrestricted Imputation
181	Niger	9.3% White, 89.5% Black/African American, 1.2% Impute among those who are not American Indian/Alaska Native
186	Australia	QD05: 92% White, 7% Asian Nonspecific, 1% Impute among those who are not American Indian/Alaska Native
187	New Zealand	QD05: 67% White, 21.7% Other Pacific Islander, 11.3% Other Asian
902	Definitely Hispanic/Latino (Hispanic Dominican Republic, Dominicano, etc.)	16% White, 11% Black/African American, 73% White and Black/African American
909	Creole	50% White, 50% White and Black/African American
915	Central/South American & West Indies	50% White and Black/African American, 50% Black/African American and American Indian/Alaska Native

Table C.3 Procedures for Restricted Imputation for Codes Informative for Formal Imputation Procedures

Race Code	Race Name	Restriction on Donors in Formal Imputation
33	Asian (nonspecific)	Donors were Asian: impute specific Asian group
54	Trinidad & Tobago (QD05: 15% Multiple Race)	Donors included respondents of more than one race
70	Mexico	Donors were Mexican ¹
71	Puerto Rico	Donors were Puerto Rican
72	Cuba	Donors were Cuban
89	Mexico & Puerto Rico	Donors were Mexican, Puerto Rican, or both
90	Mexico & Cuba	Donors were Mexican, Cuban, or both
91	Mexico & Dominican Republic	Donors were Mexican, Dominican, or both
92	Mexico & Spain	Donors were Mexican, Spanish, or both
93	Puerto Rico & Cuba	Donors were Puerto Rican, Cuban, or both
94	Puerto Rico & Dominican Republic	Donors were Puerto Rican, Dominican, or both
95	Puerto Rico & Spain	Donors were Puerto Rican, Spanish, or both
96	Cuba & Dominican Republic	Donors were Cuban, Dominican, or both
97	Cuba & Spain	Donors were Cuban, Spanish, or both
98	Dominican Republic & Spain	Donors were Dominican, Spanish, or both
103	United Kingdom (QD05: 2% Multiple Race)	Donors included respondents of more than one race
128	UAE (QD05: 8% Not American Indian/Alaska Native)	Donors included respondents of any race or races that did not include American Indian/Alaska Native
129	Qatar (QD05: 14% Not American Indian/Alaska Native)	Donors included respondents of any race or races that did not include American Indian/Alaska Native
140	Kazakhstan (QD05: 4.5% Not American Indian/Alaska Native)	Donors included respondents of any race or races that did not include American Indian/Alaska Native
141	Uzbekistan (QD05: 2.5% Not American Indian/Alaska Native)	Donors included respondents of any race or races that did not include American Indian/Alaska Native
142	Tadjikistan (QD05: 2.6% Not American Indian/Alaska Native)	Donors included respondents of any race or races that did not include American Indian/Alaska Native
143	Kyrgyzstan (QD05: 5.7% Not American Indian/Alaska Native)	Donors included respondents of any race or races that did not include American Indian/Alaska Native
144	Turkmenistan (QD05: 6.0% Not American Indian/Alaska Native)	Donors included respondents of any race or races that did not include American Indian/Alaska Native
157	Malaysia (QD05: 8.8% Not American Indian/Alaska Native)	Donors included respondents of any race or races that did not include American Indian/Alaska Native
159	Singapore (QD05: 3.3% Not American Indian/Alaska Native)	Donors included respondents of any race or races that did not include American Indian/Alaska Native
166	Sudan (QD05: 3% Not American Indian/Alaska Native)	Donors included respondents of any race or races that did not include American Indian/Alaska Native
168	South Africa (QD05: 0.5% Not American Indian/Alaska Native)	Donors included respondents of any race or races that did not include American Indian/Alaska Native

Table C.3 Procedures for Restricted Imputation for Codes Informative for Formal Imputation Procedures (continued)

Race Code	Race Name	Restriction on Donors in Formal Imputation
171	Zambia (QD05: 1% Not American Indian/Alaska Native)	Donors included respondents of any race or races that did not include American Indian/Alaska Native
181	Niger (QD05: 1.2% Not American Indian/Alaska Native)	Donors included respondents of any race or races that did not include American Indian/Alaska Native
186	Australia (QD05: 1% Not American Indian/Alaska Native)	Donors included respondents of any race or races that did not include American Indian/Alaska Native
201	Biracial (nonspecific)	Donors were multiple race: imputed constituent races ²
900	Definitely Hispanic/Latino (Hispanic, Latino/a, Chicano/a, etc., not Spain, D.R.)	Donors were Hispanic/Latino
901	Definitely Hispanic/Latino (Hispanic Spanish, Español, etc.)	Donors were Hispanic/Latino
903	Central/South American (no country)	Donors were Central/South American
904	Nonwhite (nonspecific/brown)	Donors were any race but single-race white
905	Hispanic/Latino nonwhite (including trigueno = "dark," moreno)	Donors were Hispanic/Latino who were any race but single-race white
906	Mezclado, Mezclada (Hispanic/Latino mixed)	Donors were multiple race and Hispanic/Latino: imputed constituent races
907	Mixed	Donors were multiple race: imputed constituent races
916	Central/South American & Mexican	Donors were Central/South American, Mexican, or both
917	Central/South American & Puerto Rican	Donors were Central/South American, Puerto Rican, or both
918	Central/South American & Cuban	Donors were Central/South American, Cuban, or both
919	Central/South American & Dominican	Donors were Central/South American, Dominican, or both
920	Central/South American & Spanish	Donors were Central/South American, Spanish, or both
932	Puerto Rican & Dominican	Donors were Puerto Rican, Dominican, or both
933	Puerto Rican & Spanish (from Spain)	Donors were Puerto Rican, Spanish, or both
935	Cuban & Spanish (from Spain)	Donors were Cuban, Spanish, or both
936	Dominican & Spanish (from Spain)	Donors were Dominican, Spanish, or both

¹ Even though a recipient may not be Hispanic/Latino, he or she may still have indicated "Mexican" in the QD05 "OTHER, Specify" response. Donors in this case included both Hispanic/Latino and (though extremely rare) non-Hispanic/Latino Mexicans.

² Because most multiple-race respondents have only two constituent races, any respondent with this code and nothing else is likely to be assigned a biracial donor. However, for the sake of simplicity, respondents with this code were not treated any differently than respondents with code 907 ("Mixed").

Table C.4 Mapping of Hispanic/Latino Group Codes

Hispanic/Latino Code	Hispanic/Latino Group Name	Category to Which Hispanic/Latino Code Directly Mapped
11	Mexican/Mexican American/Mexicano/Chicano	Mexican
12	Puerto Rican	Puerto Rican
13	Central or South American	Central or South American
14	Cuban/Cuban American	Cuban
15	Dominican (Dominican Republic)	Dominican
16	Spanish (from Spain)	Spanish
17	Caribbean Hispanic/Latino (not specified as Dominican)	Other Hispanic
21	Mexican & Puerto Rican	Mexican, Puerto Rican
22	Mexican & Central or South American	Mexican, Central or South American
23	Mexican & Cuban	Mexican, Cuban
24	Mexican & Dominican	Mexican, Dominican
25	Mexican & Spanish (from Spain)	Mexican, Spanish
26	Puerto Rican & Central or South American	Puerto Rican, Central or South American
27	Puerto Rican & Cuban	Puerto Rican, Cuban
28	Puerto Rican & Dominican	Puerto Rican, Dominican
29	Puerto Rican & Spanish (from Spain)	Puerto Rican, Spanish
30	Central or South American & Cuban	Central or South American, Cuban
31	Central or South American & Dominican	Central or South American, Dominican
32	Central or South American & Spanish (from Spain)	Central or South American, Spanish
33	Cuban & Dominican	Cuban, Dominican
34	Cuban & Spanish (from Spain)	Cuban, Spanish
35	Dominican & Spanish (from Spain)	Dominican, Spanish
36	Mexican, Puerto Rican, & Central or South American	Mexican, Puerto Rican, Central or South American
37	Mexican, Puerto Rican, & Cuban	Mexican, Puerto Rican, Cuban
38	Mexican, Puerto Rican, & Dominican	Mexican, Puerto Rican, Dominican
39	Mexican, Puerto Rican, & Spanish (from Spain)	Mexican, Puerto Rican, Spanish
40	Mexican, Central or South American, & Cuban	Mexican, Central or South American, Cuban
41	Mexican, Central or South American, & Dominican	Mexican, Central or South American, Dominican

Table C.4 Mapping of Hispanic/Latino Group Codes (continued)

Hispanic/Latino Code	Hispanic/Latino Group Name	Category to Which Hispanic/Latino Code Directly Mapped
42	Mexican, Central or South American, & Spanish (from Spain)	Mexican, Central or South American, Spanish
43	Mexican, Cuban, & Dominican	Mexican, Cuban, Dominican
44	Mexican, Cuban, & Spanish (from Spain)	Mexican, Cuban, Spanish
45	Mexican, Dominican, & Spanish (from Spain)	Mexican, Dominican, Spanish
46	Puerto Rican, Central or South American, & Cuban	Puerto Rican, Central or South American, Cuban
47	Puerto Rican, Central or South American, & Dominican	Puerto Rican, Central or South American, Dominican
48	Puerto Rican, Central or South American, & Spanish (from Spain)	Puerto Rican, Central or South American, Spanish
49	Puerto Rican, Cuban, & Dominican	Puerto Rican, Cuban, Dominican
50	Puerto Rican, Cuban, & Spanish (from Spain)	Puerto Rican, Cuban, Spanish
51	Puerto Rican, Dominican, & Spanish (from Spain)	Puerto Rican, Dominican, Spanish
52	Central or South American, Cuban, & Dominican	Central or South American, Cuban, Dominican
53	Central or South American, Cuban, & Spanish (from Spain)	Central or South American, Cuban, Spanish
54	Central or South American, Dominican, & Spanish (from Spain)	Central or South American, Dominican, Spanish
55	Cuban, Dominican, and Spanish (from Spain)	Cuban, Dominican, Spanish
56	Portuguese & Mexican	Mexican, Other Hispanic
57	Portuguese & Puerto Rican	Puerto Rican, Other Hispanic
58	Portuguese & Cuban	Cuban, Other Hispanic
59	Portuguese & Central or South American	Central or South American, Other Hispanic
60	Portuguese & Dominican	Dominican, Other Hispanic
61	Portuguese & Spanish (from Spain)	Spanish, Other Hispanic
100	Brazilian	Central or South American
101	Portuguese	Other Hispanic/Latino
102	Cape Verde	Other Hispanic/Latino
103	Belizean (formerly British Honduras)	Central or South American
104	Guyana	Central or South American

Table C.4 Mapping of Hispanic/Latino Group Codes (continued)

Hispanic/Latino Code	Hispanic/Latino Group Name	Category to Which Hispanic/Latino Code Directly Mapped
105	Jamaican	Caribbean Hispanic
106	Other Caribbean (possibly Hispanic)	Caribbean Hispanic
107	Philippines & Guam	Other Hispanic/Latino
108	Brazilian & Portuguese	Central or South American
109	Cape Verde & Portuguese	Other Hispanic/Latino
110	Haitian/Jamaican	Caribbean Hispanic
200	Mexican/Jamaican	Mexican, Caribbean Hispanic
201	Puerto Rican/Jamaican	Puerto Rican, Caribbean Hispanic
202	Central or South American/Jamaican	Central or South American, Caribbean Hispanic
203	Cuban/Jamaican	Cuban, Caribbean Hispanic
204	Dominican/Jamaican	Dominican, Caribbean Hispanic
205	Spanish (from Spain)/Jamaican	Spanish, Caribbean Hispanic
206	Mexican/West Indies	Mexican, Caribbean Hispanic
207	Puerto Rican/West Indies	Puerto Rican, Caribbean Hispanic
208	Central or South American/West Indies	Central or South American, Caribbean Hispanic
209	Cuban/West Indies	Cuban, Caribbean Hispanic
210	Dominican/West Indies	Dominican, Caribbean Hispanic
211	Spanish (from Spain)/West Indies	Spanish, Caribbean Hispanic
212	Mexican/Haitian	Mexican, Caribbean Hispanic
213	Puerto Rican/Haitian	Puerto Rican, Caribbean Hispanic
214	Central or South American/Haitian	Central or South American, Caribbean Hispanic
215	Cuban/Haitian	Cuban, Caribbean Hispanic
216	Dominican/Haitian	Dominican, Caribbean Hispanic

Table C.4 Mapping of Hispanic/Latino Group Codes (continued)

Hispanic/Latino Code	Hispanic/Latino Group Name	Category to Which Hispanic/Latino Code Directly Mapped
217	Spanish (from Spain)/Haitian	Spanish, Caribbean Hispanic
500	Hispanic/Latino	Hispanic/Latino group imputed
501	Hispanic/Latino Mixed/Mezclada	Hispanic group imputed
502	Hispanic Creole	Other Hispanic
600*	Stated clearly as Not Hispanic/Latino	Hispanic/Latino indicator edited to "no"
800	Non-Hispanic/Latino country	Other Hispanic/Latino
801	Racial category (white, black/African American, etc.)	Hispanic group imputed
802	Combination race and non-Hispanic country	Other Hispanic/Latino
985	Bad Data/"Mixed"	Hispanic/Latino group imputed
994	Unknown/"Don't Know"	Hispanic/Latino group imputed
997	American or "All of Them"	Hispanic/Latino group imputed

* This code caused the Hispanic/Latino indicator to be edited to a "no." Codes that caused the Hispanic/Latino indicator to be edited to a "yes" are listed in [Table C.1](#).

Table C.5 Comparison of Codes for Prescription Drugs in 2016 and 2014

Prescription Drug Category	2016 Code	2016 Description	2014 Code	2014 Description
Sedative	101	Butisol	101	Butisol
Sedative	102	Buticaps	102	Buticaps
Sedative	103	Amytal	103	Amytal
Sedative	104	Mebaral	104	Mebaral
Sedative	105	Placidyl, Ethchlorvynol	105	Placidyl, Ethchlorvynol
Sedative	106	Doriden	106	Doriden
Sedative	107	Noludar	107	Noludar
Sedative	108	Nembutal	108	Nembutal
Sedative	109	Seconal	109	Seconal
Sedative	110	Tuinal	110	Tuinal
Sedative	111	Dalmane [see also 1502]	111	Dalmane, Flurazepam
Sedative	112	Restoril [see also 1503]	112	Restoril, Temazepam
Sedative	113	Halcion [see also 1504]	113	Halcion, Triazolam
Sedative	114	Amobarbital	114	Amobarbital
Sedative	115	Phenobarbital	115	Phenobarbital
Sedative	116	Methaqualone, Sopor, Quaalude [see also 138]	116	Methaqualone, Sopor, Quaalude [see also 138]
Sedative	117	Chloral hydrate, Noctec	117	Chloral hydrate, Noctec
Sedative	118	Pentobarbital, Pentothal, Sodium pentothal, Thiopental	118	Pentobarbital, Pentothal, Sodium pentothal, Thiopental
Sedative	119	Secobarbital	119	Secobarbital
Sedative	120	Downers; OTHER Sedative, NOT OTHERWISE SPECIFIED	120	Downers; OTHER Sedative, NOT OTHERWISE SPECIFIED
Sedative	121	USED SEDATIVE, DON'T KNOW NAME	121	USED SEDATIVE, DON'T KNOW NAME
Sedative	122	Over-the-counter sleep aid; not specified <OTC>	122	Over-the-counter sleep aid; not specified <OTC>
Sedative	123	Unisom <OTC>	123	Unisom <OTC>
Sedative	124	Yellow jackets	124	Yellow jackets
Sedative	126	Sominex <OTC>	126	Sominex <OTC>
Sedative	127	Nytol <OTC>	127	Nytol <OTC>
Sedative	128	Nervine <OTC>	128	Nervine <OTC>
Sedative	129	Reds	129	Reds
Sedative	130	Promethazine	130	Phenergan, Promethazine
Sedative	132	Sleepettes <OTC>	132	Sleepettes <OTC>
Sedative	137	7-14	137	7-14
Sedative	138	Mandrax, Mandrix [see also 116]	138	Mandrax, Mandrix [see also 116]
Sedative	139	Black widows	139	Black widows
Sedative	142	Belladonna	142	Belladonna
Sedative	145	Red devil(s)	145	Red devil(s)
Sedative	146	Bentyl, Dicyclomine	146	Bentyl, Dicyclomine
Sedative	152	Eskabarb	152	Eskabarb
Sedative	153	Luminal	153	Luminal

Table C.5 Comparison of Codes for Prescription Drugs in 2016 and 2014 (continued)

Prescription Drug Category	2016 Code	2016 Description	2014 Code	2014 Description
Sedative	154	Alurate	154	Alurate
Sedative	156	Sodium barbital	156	Sodium barbital
Sedative	157	Barbiturates	157	Barbiturates
Sedative	159	Octanox	159	Octanox
Sedative	163	Blue meanies	163	Blue meanies
Sedative	164	Yellow zombies	164	Yellow zombies
Sedative	171	NyQuil <OTC>	171	NyQuil <OTC>
Sedative	172	Dream Ease <OTC>	172	Dream Ease <OTC>
Sedative	173	Wafers	173	Wafers
Sedative	174	Ambien [see also 1501]	174	Ambien, Stilnox, Zolpidem
Sedative	176	Melatonin <OTC>	176	Melatonin <OTC>
Sedative	177	Bean	177	Bean
Sedative	179	Water tee [sic]	179	Water tee [sic]
Sedative	180	Tylenol P.M. <OTC>	180	Tylenol P.M. <OTC>
Sedative	181	Excedrin P.M. <OTC>	181	Excedrin P.M. <OTC>
Sedative	182	Melzone	182	Melzone
Sedative	184	Red and blues	184	Red and blues
Sedative	186	Butalbital	186	Butalbital
Sedative	187	Belladonna plant/flower/root, deadly nightshade	187	Belladonna plant/flower/root, deadly nightshade
Sedative	188	Prosom	188	Estazolam, Prosom
Sedative	189	Sonata, Zaleplon	189	Sonata, Zaleplon
Sedative	190	Dormin <OTC>	190	Dormin <OTC>
Sedative	191	Blues	191	Blues
Sedative	194	Quiet World <OTC>	194	Quiet World <OTC>
Sedative	195	Lotusate, Talbutal	195	Lotusate, Talbutal
Sedative	196	Simply Sleep <OTC>	196	Simply Sleep <OTC>
Sedative	197	Topamax	197	Topamax, Topiramate
Sedative	198	Seducaps	198	Seducaps
Sedative	199	Ansiovit <OTC>	199	Ansiovit <OTC>
Tranquilizer	201	Valium [see also 2506]	201	Valium, Diazepam
Tranquilizer	202	Librium [see also 2507]	202	Librium, Chlordiazepoxide HCl
Tranquilizer	203	Limbitrol [see also 276]	203	Limbitrol [see also 276]
Tranquilizer	204	Menrium	204	Menrium
Tranquilizer	205	Serax [see also 2508]	205	Serax, Oxazepam
Tranquilizer	206	Tranxene	206	Tranxene, Clorazepate, Tranxilium
Tranquilizer	207	Ativan [see also 2503]	207	Ativan, Lorazepam, Orfidal
Tranquilizer	208	Centrax	208	Centrax
Tranquilizer	209	Xanax [see also 2501, 2502]	209	Xanax, Alprazolam, Tafil, Xanax bars
Tranquilizer	210	Paxipam	210	Paxipam
Tranquilizer	211	BuSpar [see also 2511]	211	BuSpar, Buspirone
Tranquilizer	212	Miltown [see also 220]	212	Miltown
Tranquilizer	213	Equanil [see also 220]	213	Equanil

Table C.5 Comparison of Codes for Prescription Drugs in 2016 and 2014 (continued)

Prescription Drug Category	2016 Code	2016 Description	2014 Code	2014 Description
Tranquilizer	214	Deprol	214	Deprol
Tranquilizer	215	Vistaril [see also 2512]	215	Vistaril
Tranquilizer	216	Atarax [see also 2512]	216	Atarax, Hydroxyzine
Tranquilizer	217	Durrax	217	Durrax
Tranquilizer	219	SK-Lygen	219	SK-Lygen
Tranquilizer	220	Meprobamate	220	Meprobamate
Tranquilizer	221	OTHER Tranquilizer, NOT OTHERWISE SPECIFIED	221	OTHER Tranquilizer, NOT OTHERWISE SPECIFIED
Tranquilizer	222	USED TRANQUILIZER, DON'T KNOW NAME	222	USED TRANQUILIZER, DON'T KNOW NAME
Tranquilizer	223	Methocarbamol	223	Carbacot, Marbaxin, Methocarbamol, Robaxin
Tranquilizer	229	Librax	229	Librax
Tranquilizer	234	Flexeril [see also 2509]	234	Flexeril, Cyclobenzaprine
Tranquilizer	237	Midazolam	237	Dormicum, Hypnovel, Midazolam, Versed
Tranquilizer	242	Equagesic	242	Equagesic
Tranquilizer	246	Klonopin [see also 2504, 2505]	246	Klonopin, Clonazepam, Diocam, Rivotril
Tranquilizer	260	PCB	260	PCB
Tranquilizer	266	Verstran	266	Verstran
Tranquilizer	268	Rohypnol, Flunitrazepam, roche, roofies	268	Rohypnol, Flunitrazepam, roche, roofies
Tranquilizer	276	Limbitrol DS [see also 203]	276	Limbitrol DS [see also 203]
Tranquilizer	278	Benzodiazepines	278	Benzodiazepines
Tranquilizer	281	Blue ringer	281	Blue ringer
Tranquilizer	288	Baclofen	288	Baclofen
Tranquilizer	290	Metaxalone, Skelaxin	290	Metaxalone, Skelaxin
Tranquilizer	296	Novoxapam	296	Novoxapam
Tranquilizer	297	Cafia	297	Cafia
Stimulant	301	Dexedrine	301	Dexedrine
Stimulant	302	Dexamyl	302	Dexamyl
Stimulant	303	Eskatrol	303	Eskatrol
Stimulant	304	Benzedrine	304	Benzedrine
Stimulant	305	Biphetamine	305	Biphetamine
Stimulant	306	Desoxyn	306	Desoxyn
Stimulant	307	Tenuate [see also 3505]	307	Tenuate, Diethylpropion
Stimulant	308	Tepanil	308	Tepanil
Stimulant	309	Didrex [see also 3504]	309	Benzphetamine, Didrex
Stimulant	310	Plegine	310	Plegine
Stimulant	311	Preludin	311	Preludin
Stimulant	312	Ionamin	312	Ionamin
Stimulant	313	Adipex-P [see also 3507]	313	Axion, Adipex-P, Fastin, Phentermine
Stimulant	314	Pondimin, Fenfluramine	314	Pondimin, Fenfluramine

Table C.5 Comparison of Codes for Prescription Drugs in 2016 and 2014 (continued)

Prescription Drug Category	2016 Code	2016 Description	2014 Code	2014 Description
Stimulant	315	Voranyl	315	Voranyl
Stimulant	316	Sanorex	316	Sanorex
Stimulant	317	Mazanor	317	Mazanor
Stimulant	318	Ritalin [see also 3502]	318	Ritalin, Methylphenidate
Stimulant	319	Cylert	319	Cylert
Stimulant	320	Dextroamphetamine	320	Dextroamphetamine
Stimulant	321	Methedrine	321	Methedrine
Stimulant	322	Methamphetamine ("ice" or "crank"), Meth [see also 3126]	322	Methamphetamine ("ice" or "crank"), Meth [see also 3126]
Stimulant	323	Obedrin-L.A.	323	Obedrin-L.A.
Stimulant	324	Uppers; OTHER Stimulant, NOT OTHERWISE SPECIFIED	324	Uppers; OTHER Stimulant, NOT OTHERWISE SPECIFIED
Stimulant	325	USED STIMULANT, DON'T KNOW NAME	325	USED STIMULANT, DON'T KNOW NAME
Stimulant	326	White cross	326	White cross
Stimulant	327	Black beauties, black beauty	327	Black beauties, black beauty
Stimulant	328	Pink hearts	328	Pink hearts
Stimulant	329	Bird's eggs, robin's eggs, speckled eggs	329	Bird's eggs, robin's eggs, speckled eggs
Stimulant	330	Blue birds, blue jays, blue robins, speckled bird	330	Blue birds, blue jays, blue robins, speckled bird
Stimulant	331	Blancs, whites	331	Blancs, whites
Stimulant	332	Ephedrine	332	Ephedrine
Stimulant	333	357 Magnum, Magnum 44 <OTC>	333	357 Magnum, Magnum 44 <OTC>
Stimulant	334	Crossroads, cross tops	334	Crossroads, cross tops
Stimulant	335	Dexatrim <OTC> [see also 347, 378]	335	Dexatrim <OTC> [see also 347, 378]
Stimulant	336	Diet pills	336	Diet pills
Stimulant	337	White molly	337	White molly
Stimulant	339	Amphetamines, otherwise unspecified	339	Amphetamines
Stimulant	341	Pink roses	341	Pink roses
Stimulant	342	Purple hearts	342	Purple hearts
Stimulant	344	Bennies	344	Bennies
Stimulant	345	Black molly	345	Black molly
Stimulant	346	Marax	346	Marax
Stimulant	347	Phenylpropanolamine <OTC> [see also 335, 378]	347	Phenylpropanolamine <OTC> [see also 335, 378]
Stimulant	348	Crystal	348	Crystal
Stimulant	349	Phendimetrazine [see also 3506]	349	Bontril, Melfiat, Phendimetrazine
Stimulant	351	Cathinone, cat, khat, q'at	351	Cathinone, cat, khat, q'at

Table C.5 Comparison of Codes for Prescription Drugs in 2016 and 2014 (continued)

Prescription Drug Category	2016 Code	2016 Description	2014 Code	2014 Description
Stimulant	353	No-Doz <OTC>	353	No-Doz <OTC>
Stimulant	354	Vivarin <OTC>	354	Vivarin <OTC>
Stimulant	355	Crystal meth	355	Crystal meth
Stimulant	356	Dexies	356	Dexies
Stimulant	357	Heart-shaped black pills	357	Heart-shaped black pills
Stimulant	358	20/20	358	20/20
Stimulant	359	Criss cross	359	Criss cross
Stimulant	360	Rainbow speed	360	Rainbow speed
Stimulant	363	Bam(s)	363	Bam(s)
Stimulant	364	Pseudoephedrine <OTC>	364	Pseudoephedrine <OTC>
Stimulant	366	Duromine	366	Duromine
Stimulant	368	Peanut butter crank	368	Peanut butter crank
Stimulant	374	Christmas tree(s)	374	Christmas tree(s)
Stimulant	376	TAC	376	TAC
Stimulant	377	Albuterol	377	Albuterol, Pro-Air, Proventil, Ventolin
Stimulant	378	Control <OTC> [see also 335, 347]	378	Control <OTC> [see also 335, 347]
Stimulant	379	Hearts	379	Hearts
Stimulant	380	Pink ladies	380	Pink ladies
Stimulant	381	P.O.T.C.	381	P.O.T.C.
Stimulant	382	Footballs, pink footballs, purple footballs	382	Footballs, pink footballs, purple footballs
Stimulant	383	Pre-Sate	383	Pre-Sate
Stimulant	384	Little crosses	384	Little crosses
Stimulant	385	Little pink ones	385	Little pink ones
Stimulant	386	Herbal ecstasy	386	Herbal ecstasy
Stimulant	388	Pink lemon	388	Pink lemon
Stimulant	390	Max Alert	390	Max Alert
Stimulant	394	Debs	394	Debs
Stimulant	395	Black cats	395	Black cats
Stimulant	396	Black RJs, RJ8	396	Black RJs, RJ8
Stimulant	398	Pink crosses	398	Pink crosses
Pain Reliever	401	Darvon	401	Darvon
Pain Reliever	402	Dolene	402	Dolene
Pain Reliever	403	SK-65	403	SK-65
Pain Reliever	404	Wygesic	404	Wygesic
Pain Reliever	405	Levo-Dromoran	405	Levo-Dromoran
Pain Reliever	406	Percodan	406	Percodan, Roxiprin
Pain Reliever	407	Demerol [see also 4511]	407	Demerol, Meperidine
Pain Reliever	408	Dilaudid, Hydromorphone	408	Dilaudid, Hydromorphone
Pain Reliever	409	Tylenol with Codeine, Tylenol 3, Tylenol 4	409	Tylenol with Codeine, Tylenol 3, Tylenol 4
Pain Reliever	410	Phenaphen with Codeine, Phenaphen 3	410	Phenaphen with Codeine, Phenaphen 3

Table C.5 Comparison of Codes for Prescription Drugs in 2016 and 2014 (continued)

Prescription Drug Category	2016 Code	2016 Description	2014 Code	2014 Description
Pain Reliever	411	Talwin	411	Talwin
Pain Reliever	412	Talwin NX	412	Talwin NX
Pain Reliever	413	Talacen	413	Talacen
Pain Reliever	414	Propoxyphene	414	Propoxyphene
Pain Reliever	415	Codeine pills [see also 4506]	415	Codeine
Pain Reliever	416	Anileridine	416	Anileridine
Pain Reliever	417	Morphine	417	Morphine, Roxanol
Pain Reliever	418	Methadone [see also 4512]	418	Methadone, Dolophine
Pain Reliever	419	Stadol, Stadol NS	419	Stadol, Stadol NS
Pain Reliever	420	OTHER Analgesic, NOT OTHERWISE SPECIFIED	420	OTHER Analgesic, NOT OTHERWISE SPECIFIED
Pain Reliever	421	USED ANALGESIC, DON'T KNOW NAME	421	USED ANALGESIC, DON'T KNOW NAME
Pain Reliever	422	Motrin	422	Motrin
Pain Reliever	423	Percocet	423	Percocet
Pain Reliever	424	Fioricet	424	Fioricet
Pain Reliever	425	Advil <OTC>	425	Advil <OTC>
Pain Reliever	426	Fiorinal	426	Fiorinal
Pain Reliever	427	Darvocet, Darvocet-N	427	Darvocet, Darvocet-N
Pain Reliever	429	Frost 222	429	Frost 222
Pain Reliever	430	Aspirin, Ecotrin, Empirin <OTC>	430	Aspirin, Ecotrin, Empirin <OTC>
Pain Reliever	431	Tylox	431	Tylox
Pain Reliever	432	Rufen	432	Rufen
Pain Reliever	433	Paregoric	433	Paregoric
Pain Reliever	434	Vicodin [see also 4501, 4502, 4503, 4504]	434	Vicodin, Lortab, Lorcet, Anexsia, Co-Gesic, Lorcet-HD, Lorcet Plus, Maxidone, Norco, Zydone
Pain Reliever	435	Tylenol, Tylenol Extra-Strength <OTC>	435	Tylenol, Tylenol Extra-Strength <OTC>
Pain Reliever	436	Bufferin <OTC>	436	Bufferin <OTC>
Pain Reliever	438	Anaprox, Naprosyn, Naproxen, Naxen	438	Anaprox, Naprosyn, Naproxen, Naxen
Pain Reliever	439	222, Canadian 222	439	222, Canadian 222
Pain Reliever	442	Ibuprofen	442	Ibuprofen
Pain Reliever	443	Diffunisal, Dolobid	443	Diffunisal, Dolobid
Pain Reliever	444	Aspirin with Codeine, Empirin 3, Empirin 4	444	Aspirin with Codeine, Empirin 3, Empirin 4
Pain Reliever	445	Norgesic Forte	445	Dolex, Norgesic Forte
Pain Reliever	447	Synalgos, Synalgos DC	447	Synalgos, Synalgos DC
Pain Reliever	450	Acetaminophen, including non-Tylenol brands <OTC>	450	Acetaminophen, including non-Tylenol brands <OTC>
Pain Reliever	453	Organidin with Codeine	453	Organidin with Codeine
Pain Reliever	454	Novocain, Procaine	454	Novocain, Procaine

Table C.5 Comparison of Codes for Prescription Drugs in 2016 and 2014 (continued)

Prescription Drug Category	2016 Code	2016 Description	2014 Code	2014 Description
Pain Reliever	455	Fiorinal with Codeine	455	Fiorinal with Codeine
Pain Reliever	457	T's and blues, tee's and b's, tops and bottoms	457	T's and blues, tee's and b's, tops and bottoms
Pain Reliever	458	Mepergan, Mepergan Fortis	458	Finalin, Mepergan, Mepergan Fortis, Meprozone
Pain Reliever	462	Norflex	462	Norflex
Pain Reliever	465	Fentanyl [see also 4508, 4509]	465	Actiq, Fentanyl, Duragesic, Sublimaze
Pain Reliever	466	Midol <OTC>	466	Midol <OTC>
Pain Reliever	467	Tolectin	467	Tolectin
Pain Reliever	468	Excedrin <OTC>	468	Excedrin <OTC>
Pain Reliever	469	Leritine	469	Leritine
Pain Reliever	470	Cafergot	470	Cafergot
Pain Reliever	471	Butazolidin	471	Butazolidin
Pain Reliever	472	Indocin, Indomethacin	472	Indocin, Indomethacin
Pain Reliever	473	Pyridium	473	Pyridium
Pain Reliever	476	Duradrin, Isocom, Midrin	476	Duradrin, Isocom, Midrin
Pain Reliever	477	Melumbrina, Neo-Melumbrina	477	Melumbrina, Neo-Melumbrina
Pain Reliever	478	Nogirtronia	478	Nogirtronia
Pain Reliever	480	Robitussin AC with Codeine	480	Robitussin AC with Codeine
Pain Reliever	481	Acular, Dolac, Ketorolac tromethamine, Torradlo	481	Acular, Dolac, Ketorolac tromethamine, Torradlo
Pain Reliever	487	Ariel	487	Ariel
Pain Reliever	490	Nalbuphine, Nubain	490	Nalbuphine, Nubain
Pain Reliever	491	PDR	491	PDR
Pain Reliever	495	Red cross	495	Red cross
Pain Reliever	496	Ultram	496	Ultram
Pain Reliever	498	Mefenamic acid, Ponstan, Ponstel	498	Mefenamic acid, Ponstan, Ponstel
Pain Reliever	499	Aleve <OTC>	499	Aleve <OTC>
Sedative	1001	Benadryl, Diphenhydramine <OTC>	1001	Benadryl, Diphenhydramine <OTC>
Sedative	1003	Donnatal	1003	Donnatal
Sedative	1004	Valerian <OTC>	1004	Valerian <OTC>
Sedative	1005	Imovane, Zopiclone	1005	Imovane, Zopiclone
Sedative	1006	Oxcarbazepine, Trileptal	1006	Oxcarbazepine, Trileptal
Sedative	1007	Robitussin P.M. <OTC>	1007	Robitussin P.M. <OTC>
Sedative	1008	Rohr 714, Rorer 714	1008	Rohr 714, Rorer 714
Sedative	1009	Mogadon	1009	Mogadon, Nitrazepam
Sedative	1010	Diprivan	1010	Diprivan, Propofol
Sedative	1011	Eszopiclone, Lunesta	1011	Lunesta
Sedative	1012	Brevital	1012	Brevital, Methohexital
Sedative	1013	Ambien CR	1013	Ambien CR
Sedative	1014	Tylenol Flu Nighttime <OTC>	1014	Tylenol Flu Nighttime <OTC>
Sedative	1015	Wild lettuce	1015	Wild lettuce

Table C.5 Comparison of Codes for Prescription Drugs in 2016 and 2014 (continued)

Prescription Drug Category	2016 Code	2016 Description	2014 Code	2014 Description
Sedative	1016	Phenazepam	1016	Phenazepam
Sedative	1017	Rozerem	1017	Ramelteon, Rozerem
Sedative	1018	Advil PM <OTC>	1018	Advil PM <OTC>
Sedative	1019	Donormyl <OTC>	1019	Donormyl <OTC>
Sedative	1022	Atropine	1022	Atropine
Sedative	1023	Relaxen <OTC>	1023	Relaxen <OTC>
Sedative	1024	Homemade sedative	1024	Homemade sedative
Sedative	1025	Acetaminophen P.M., including non-Tylenol brands <OTC>	1025	Acetaminophen P.M., including non-Tylenol brands <OTC>
Sedative	1026	Moon drops <OTC>	1026	Moon drops <OTC>
Sedative	1027	ZzzQuil <OTC>	1027	ZzzQuil <OTC>
Sedative	1028	Dormidina	1028	Dormidina
Sedative	1029	Sleep-Eze <OTC>	1029	Sleep-Eze <OTC>
Sedative	1030	Alteril	1030	Alteril
Sedative	1031	Zolfresh	None	None
Sedative	1032	Aleve P.M.<OTC>	None	None
Sedative	1501	Zolpidem [see also 174]	174	Ambien, Stilnox, Zolpidem
Sedative	1502	Flurazepam [see also 111]	111	Dalmane, Flurazepam
Sedative	1503	Temazepam [see also 112]	112	Restoril, Temazepam
Sedative	1504	Triazolam [see also 113]	113	Halcion, Triazolam
Tranquilizer	2001	Soma [see also 2510]	2001	Sanoma, Soma, Carisoprodol
Tranquilizer	2002	Muscle relaxers	2002	Muscle relaxers
Tranquilizer	2003	Zanies, zannies, zans, zany bars	2003	Zanies, zannies, zans, zany bars
Tranquilizer	2004	Bromazepam, Lexillum, Lexomil, Lexotan, Lexotanil	2004	Bromazepam, Lexillum, Lexomil, Lexotan, Lexotanil
Tranquilizer	2006	Chlorzoxazone	2006	Chlorzoxazone
Tranquilizer	2007	R2	2007	R2
Tranquilizer	2008	747	2008	747
Tranquilizer	2009	Inapsine	2009	Droperidol, Inapsine
Tranquilizer	2010	Buscopan, Espacil	2010	Buscopan, Espacil
Tranquilizer	2013	Hyoscyamine, Levsin	2013	Hyoscyamine, Levsin
Tranquilizer	2014	Robaxisal	2014	Methoxisal, Robaxisal
Tranquilizer	2015	Flexodin	2015	Flexodin
Tranquilizer	2016	Xanax XR	2016	Xanax XR
Tranquilizer	2017	Motival	2017	Motival
Tranquilizer	2018	4 Bar, Quad Bar	2018	4 Bar, Quad Bar
Tranquilizer	2019	Soma Compound	2019	Soma Compound
Tranquilizer	2022	Tetrazepam	2022	Tetrazepam
Tranquilizer	2023	Tandrilax	2023	Tandrilax
Tranquilizer	2024	Muscoril, Myoril, Neoflax, Thiocolchicoside	2024	Muscoril, Myoril, Neoflax, Thiocolchicoside
Tranquilizer	2025	Clobazam	2025	Clobazam
Tranquilizer	2026	Sedoxil	2026	Sedoxil
Tranquilizer	2027	Homemade tranquilizers	2027	Homemade tranquilizers
Tranquilizer	2028	Maolate, Chlorphenesin, Musil	2028	Chlorphenesin, Maolate, Musil

Table C.5 Comparison of Codes for Prescription Drugs in 2016 and 2014 (continued)

Prescription Drug Category	2016 Code	2016 Description	2014 Code	2014 Description
Tranquilizer	2029	Etizolam	2029	Etizolam
Tranquilizer	2030	Versed	None	None
Tranquilizer	2501	Alprazolam [see also 209]	209	Xanax, Alprazolam, Tafil, Xanax bars
Tranquilizer	2502	Xanax bars [see also 209]	209	Xanax, Alprazolam, Tafil, Xanax bars
Tranquilizer	2503	Lorazepam [see also 207]	207	Ativan, Lorazepam, Orfidal
Tranquilizer	2504	Clonazepam [see also 246]	246	Klonopin, Clonazepam, Diocam, Rivotril
Tranquilizer	2505	Clonazepam product, otherwise unspecified [see also 246]	246	Klonopin, Clonazepam, Diocam, Rivotril
Tranquilizer	2506	Diazepam [see also 201]	201	Valium, Diazepam
Tranquilizer	2507	Chlordiazepoxide HCl [see also 202]	202	Librium, Chlordiazepoxide HCl
Tranquilizer	2508	Oxazepam [see also 205]	205	Serax, Oxazepam
Tranquilizer	2509	Cyclobenzaprine [see also 234]	234	Flexeril, Cyclobenzaprine
Tranquilizer	2510	Carisoprodol [see also 2001]	2001	Sanoma, Soma, Carisoprodol
Tranquilizer	2511	Buspirone [see also 211]	211	BuSpar, Buspirone
Tranquilizer	2512	Hydroxyzine [see also 215, 216]	216	Atarax, Hydroxyzine
Stimulant	3001	Brown and clear	3001	Brown and clear
Stimulant	3002	Pep ups	3002	Pep ups
Stimulant	3003	Primatene tablets <OTC>	3003	Primatene tablets <OTC>
Stimulant	3004	Black Cadillac	3004	Black Cadillac
Stimulant	3006	Green amps	3006	Green amps
Stimulant	3007	Nasal inhalant	3007	Nasal inhalant
Stimulant	3008	Miniwhites	3008	Miniwhites
Stimulant	3009	Ephedrone, Methcathinone	3009	Ephedrone, Methcathinone
Stimulant	3010	KOOD12	3010	KOOD12
Stimulant	3011	Escalation	3011	Escalation
Stimulant	3012	Stud 100, stud the rock	3012	Stud 100, stud the rock
Stimulant	3013	Tweek	3013	Tweek
Stimulant	3014	Uptime	3014	Uptime
Stimulant	3015	Fen-Phen (fenfluramine and phentermine)	3015	Fen-Phen (fenfluramine and phentermine)
Stimulant	3017	Dayquil <OTC>	3017	Dayquil <OTC>
Stimulant	3018	Up-johns	3018	Up-johns
Stimulant	3019	Peppies	3019	Peppies
Stimulant	3020	White powder	3020	White powder
Stimulant	3021	Blue balls	3021	Blue balls
Stimulant	3023	Adderall	3023	Adderall
Stimulant	3024	Yohimbe	3024	Yohimbe
Stimulant	3025	Moles [sic]	3025	Moles [sic]
Stimulant	3027	Energy pills	3027	Energy pills
Stimulant	3028	Blackbirds	3028	Blackbirds
Stimulant	3029	Soccerballs	3029	Soccerballs

Table C.5 Comparison of Codes for Prescription Drugs in 2016 and 2014 (continued)

Prescription Drug Category	2016 Code	2016 Description	2014 Code	2014 Description
Stimulant	3030	Red hearts	3030	Red hearts
Stimulant	3031	Brethine, Bricanyl, Terbutaline	3031	Brethine, Bricanyl, Terbutaline
Stimulant	3032	Heads up	3032	Heads up
Stimulant	3033	Metabolife <OTC>	3033	Metabolife <OTC>
Stimulant	3034	Speckleberry	3034	Speckleberry
Stimulant	3035	Blacks	3035	Blacks
Stimulant	3036	Ripped fuel <OTC>	3036	Ripped fuel <OTC>
Stimulant	3038	Obenil	3038	Obenil
Stimulant	3039	Theophed	3039	Theophed
Stimulant	3040	Hyper pills	3040	Hyper pills
Stimulant	3041	Desbutal	3041	Desbutal
Stimulant	3042	30-30	3042	30-30
Stimulant	3043	Brown bodies	3043	Brown bodies
Stimulant	3044	Caffeine pills <OTC>	3044	Caffeine pills <OTC>
Stimulant	3045	Stay awakes <OTC>	3045	Stay awakes <OTC>
Stimulant	3047	Sudafed <OTC>	3047	Sudafed <OTC>
Stimulant	3048	Diet Pep <OTC>	3048	Diet Pep <OTC>
Stimulant	3049	Truck driver's speed	3049	Truck driver's speed
Stimulant	3050	Mini thins	3050	Mini thins
Stimulant	3051	Betel nut, beetlenut	3051	Betel nut, beetlenut
Stimulant	3052	Speed; otherwise unspecified	3052	Speed; otherwise unspecified
Stimulant	3053	Maxair, Pirbuterol	3053	Maxair, Pirbuterol
Stimulant	3054	Up you(r) gas <OTC>	3054	Up you(r) gas <OTC>
Stimulant	3055	Clenbuterol	3055	Clenbuterol
Stimulant	3056	Xenadrine <OTC>	3056	Xenadrine <OTC>
Stimulant	3057	Diet Fuel <OTC>	3057	Diet Fuel <OTC>
Stimulant	3058	Metabolift <OTC>	3058	Metabolift <OTC>
Stimulant	3059	Metabolize and Save <OTC>	3059	Metabolize and Save <OTC>
Stimulant	3060	Hydroxycut <OTC>	3060	Hydroxycut <OTC>
Stimulant	3061	Metab-O-Lite <OTC>	3061	Metab-O-Lite <OTC>
Stimulant	3062	Homemade stimulants	3062	Homemade stimulants
Stimulant	3063	Dymetadrine <OTC>	3063	Dymetadrine <OTC>
Stimulant	3064	Ephedra, Ma Huang	3064	Ephedra, Ma Huang
Stimulant	3065	MA-300 <OTC>	3065	MA-300 <OTC>
Stimulant	3066	Alert Blue and Clear <OTC>	3066	Alert Blue and Clear <OTC>
Stimulant	3068	Mini Two-way Action <OTC>	3068	Mini Two-way Action <OTC>
Stimulant	3069	DextroStat	3069	DextroStat
Stimulant	3070	Synatan	3070	Synatan
Stimulant	3071	Adipokinetix <OTC>	3071	Adipokinetix <OTC>
Stimulant	3072	Purple Blast <OTC>	3072	Purple Blast <OTC>
Stimulant	3073	Over-the-counter diet pill; otherwise unspecified <OTC>	3073	Over-the-counter diet pill; otherwise unspecified <OTC>
Stimulant	3074	Stay Trim <OTC>	3074	Stay Trim <OTC>
Stimulant	3075	Inibex	3075	Inibex

Table C.5 Comparison of Codes for Prescription Drugs in 2016 and 2014 (continued)

Prescription Drug Category	2016 Code	2016 Description	2014 Code	2014 Description
Stimulant	3076	Norephedrine <OTC>	3076	Norephedrine <OTC>
Stimulant	3077	Concerta	3077	Concerta
Stimulant	3078	Blue and clears	3078	Blue and clears
Stimulant	3080	Meridia, Sibutramine	3080	Meridia, Sibutramine
Stimulant	3081	P2P, Phenyl-2-propanone	3081	P2P, Phenyl-2-propanone
Stimulant	3082	Fat burner; not otherwise specified	3082	Fat burner; not otherwise specified
Stimulant	3083	Redotex	3083	Redotex
Stimulant	3084	Cartwheels	3084	Cartwheels
Stimulant	3085	Thermo Carb <OTC>	3085	Thermo Carb <OTC>
Stimulant	3086	Adderall XR	3086	Adderall XR
Stimulant	3088	Super Stingers <OTC>	3088	Super Stingers <OTC>
Stimulant	3089	Christina, Tina	3089	Christina, Tina
Stimulant	3090	Provigil [see also 3508]	3090	Modafinil, Provigil
Stimulant	3091	Black bombers	3091	Black bombers
Stimulant	3092	Metadate CD	3092	Metadate CD
Stimulant	3093	Supercaps	3093	Supercaps
Stimulant	3094	Focalin [see also 3503]	3094	Dexmethylphenidate, Focalin
Stimulant	3095	LoCarb Turbo <OTC>	3095	LoCarb Turbo <OTC>
Stimulant	3096	Yaba	3096	Yaba
Stimulant	3097	Bamadex	3097	Bamadex
Stimulant	3098	Zantrex 3 <OTC>	3098	Zantrex 3 <OTC>
Stimulant	3099	Asenlix, Clobenzorex	3099	Asenlix, Clobenzorex
Stimulant	3100	Over-the-counter stimulant; not otherwise specified <OTC>	3100	Over-the-counter stimulant; not otherwise specified <OTC>
Stimulant	3101	Dexitac Stay Alert Stimulant. <OTC>	3101	Dexitac Stay Alert Stimulant. <OTC>
Stimulant	3102	Methylin	3102	Methylin
Stimulant	3103	RedLine. <OTC>	3103	RedLine. <OTC>
Stimulant	3104	Slimquick. <OTC>	3104	Slimquick. <OTC>
Stimulant	3105	Affinato	3105	Affinato
Stimulant	3106	Sedalmerck	3106	Sedalmerck
Stimulant	3107	Guarana. <OTC>	3107	Guarana. <OTC>
Stimulant	3109	Dietrex <OTC>	3109	Dietrex <OTC>
Stimulant	3110	Dexfenfluramine, Redux	3110	Dexfenfluramine, Redux
Stimulant	3111	Vyvanse	3111	LDX, Lisdexamfetamine dimesylate, Vyvanse
Stimulant	3113	Levophed bitartrate, Norepinephrine	3113	Levophed bitartrate, Norepinephrine
Stimulant	3114	Ritalin LA, Ritalin SR	3114	Ritalin LA
Stimulant	3115	Isometheptene	3115	Isometheptene
Stimulant	3116	MCAT, Mephedrone	3116	MCAT, Mephedrone
Stimulant	3117	Methedrone	3117	Methedrone
Stimulant	3119	MDPV, Methylenedioxypyrovalerone	3119	MDPV, Methylenedioxypyrovalerone

Table C.5 Comparison of Codes for Prescription Drugs in 2016 and 2014 (continued)

Prescription Drug Category	2016 Code	2016 Description	2014 Code	2014 Description
Stimulant	3120	DetoxaTrim <OTC>	3120	DetoxaTrim <OTC>
Stimulant	3122	Nuvigil	3122	Armodafinil, Nuvigil
Stimulant	3123	Adrafinil, Olmifon	3123	Adrafinil, Olmifon
Stimulant	3124	Focalin XR	3124	Focalin XR
Stimulant	3125	3, 4-Ethylenedioxy-N-methylamphetamine, EDMA, MDMC	3125	3, 4-Ethylenedioxy-N-methylamphetamine, EDMA, MDMC
Stimulant	3126	Pervitin [see also 322]	3126	Pervitin [see also 322]
Stimulant	3127	4-FMA, 4-Fluoromethamphetamine	3127	4-FMA, 4-Fluoromethamphetamine
Stimulant	3128	Jet-Alert <OTC>	3128	Jet-Alert <OTC>
Stimulant	3129	JavaSLIM green coffee extract <OTC>	3129	JavaSLIM green coffee extract <OTC>
Stimulant	3130	IFA-Norex	3130	IFA-Norex
Stimulant	3131	5-MAPB	None	None
Stimulant	3187	Greenies	3187	Greenies
Stimulant	3188	Anorex, Leptoprin <OTC>	3188	Anorex, Leptoprin <OTC>
Stimulant	3189	BetaLean <OTC>	3189	BetaLean <OTC>
Stimulant	3190	Primatene Mist <OTC>	3190	Primatene Mist <OTC>
Stimulant	3501	Amphetamine/dextroamphetamine salts	None	None
Stimulant	3502	Methylphenidate [see also 318]	318	Ritalin, Methylphenidate
Stimulant	3503	Dexmethylphenidate [see also 3094]	3094	Dexmethylphenidate, Focalin
Stimulant	3504	Benzphetamine [see also 309]	309	Benzphetamine, Didrex
Stimulant	3505	Diethylpropion [see also 307]	307	Tenuate, Diethylpropion
Stimulant	3506	Bontril [see also 349]	349	Bontril, Melfiat, Phendimetrazine
Stimulant	3507	Phentermine [see also 313]	313	Axion, Adipex-P, Fastin, Phentermine
Stimulant	3508	Modafinil [see also 3090]	3090	Modafinil, Provigil
Pain Reliever	4001	Actron, Ketoprofen, Orudis, Oruvail	4001	Actron, Ketoprofen, Orudis, Oruvail
Pain Reliever	4002	Imitrex	4002	Imitrex, Sumatriptan
Pain Reliever	4005	Etodolac, Lodine	4005	Etodolac, Lodine
Pain Reliever	4007	Happy pills	4007	Happy pills
Pain Reliever	4008	Brown gel pills	4008	Brown gel pills
Pain Reliever	4009	Smilies	4009	Smilies
Pain Reliever	4011	Atretol, Carbamazepine, Carbatrol, Epitol, Tegretol	4011	Atretol, Carbamazepine, Carbatrol, Epitol, Tegretol
Pain Reliever	4013	Fenoprofen, Nalfron	4013	Fenoprofen, Nalfron
Pain Reliever	4018	Nuprin <OTC>	4018	Nuprin <OTC>
Pain Reliever	4022	Buprenex	4022	Buprenex
Pain Reliever	4023	Pamprin <OTC>	4023	Pamprin <OTC>
Pain Reliever	4024	Reynol [sic]	4024	Reynol [sic]

Table C.5 Comparison of Codes for Prescription Drugs in 2016 and 2014 (continued)

Prescription Drug Category	2016 Code	2016 Description	2014 Code	2014 Description
Pain Reliever	4025	Viserol [sic]	4025	Viserol [sic]
Pain Reliever	4026	Ventrex [sic]	4026	Ventrex [sic]
Pain Reliever	4027	Rorer	4027	Rorer
Pain Reliever	4028	Arthrotec	4028	Arthrotec
Pain Reliever	4029	Feldine, Piroxicam	4029	Feldine, Piroxicam
Pain Reliever	4030	Triple "S"	4030	Triple "S"
Pain Reliever	4031	Cough syrup/medicine with Codeine	4031	Cough syrup/medicine with Codeine
Pain Reliever	4033	Vicoprofen	4033	Vicoprofen
Pain Reliever	4034	Lilly red	4034	Lilly red
Pain Reliever	4035	Nabumetone, Relafin	4035	Nabumetone, Relafin
Pain Reliever	4036	Stagesic	4036	Stagesic
Pain Reliever	4037	MS Contin [see also 4507]	4037	Kadian, MS Contin
Pain Reliever	4038	Propacet	4038	Propacet
Pain Reliever	4039	Daypro, Oxypirozin	4039	Daypro, Oxypirozin
Pain Reliever	4040	Ridaura	4040	Ridaura
Pain Reliever	4041	Maxalt, Rizatriptan	4041	Maxalt
Pain Reliever	4042	M-33	4042	M-33
Pain Reliever	4043	Tramadol	4043	Tramadol
Pain Reliever	4044	Hydrocodone	4044	Hydrocodone
Pain Reliever	4045	Dolgic, Phrenilin Forte, Sedapap	4045	Dolgic, Phrenilin Forte, Sedapap
Pain Reliever	4046	Oxycodone [see also 4505]	4046	Oxycodone or unspecified oxycodone products
Pain Reliever	4047	Roxicet	4047	Roxicet
Pain Reliever	4049	Headache power <OTC>	4049	Headache power <OTC>
Pain Reliever	4050	Percogesic <OTC>	4050	Percogesic <OTC>
Pain Reliever	4051	Celebrex	4051	Celebrex
Pain Reliever	4052	Endocet	4052	Endocet
Pain Reliever	4053	Codydramol	4053	Codydramol
Pain Reliever	4054	Zomax	4054	Zomax
Pain Reliever	4055	Vioxx	4055	Vioxx
Pain Reliever	4056	Cataflam, Diclofenac, Voltaren	4056	Cataflam, Diclofenac, Voltaren
Pain Reliever	4057	Mejoral	4057	Mejoral
Pain Reliever	4058	Totem poles	4058	Totem poles
Pain Reliever	4059	Zomig, Zolmitriptan	4059	Zomig, Zolmitriptan
Pain Reliever	4060	Ansaid, Flurbiprophen, Majezik	4060	Ansaid, Flurbiprophen, Majezik
Pain Reliever	4061	Oramorph SR	4061	Oramorph SR
Pain Reliever	4062	Iophen-C Liquid	4062	Iophen-C Liquid
Pain Reliever	4063	Naprelan	4063	Naprelan
Pain Reliever	4064	Epidrin	4064	Epidrin
Pain Reliever	4065	Roxicodone	4065	Roxicodone
Pain Reliever	4067	OxyContin	4067	OxyContin
Pain Reliever	4068	Roxilox	4068	Roxilox
Pain Reliever	4069	Benflogin	4069	Benflogin
Pain Reliever	4070	Flexagen <OTC>	4070	Flexagen <OTC>

Table C.5 Comparison of Codes for Prescription Drugs in 2016 and 2014 (continued)

Prescription Drug Category	2016 Code	2016 Description	2014 Code	2014 Description
Pain Reliever	4071	Syncol	4071	Syncol
Pain Reliever	4072	Oxycet	4072	Oxycet
Pain Reliever	4073	Temgesic	4073	Temgesic
Pain Reliever	4074	Provil <OTC>	4074	Provil <OTC>
Pain Reliever	4075	Disalcid, Salflex, Salsalate	4075	Disalcid, Salflex, Salsalate
Pain Reliever	4076	Over-the-counter pain reliever; not otherwise specified <OTC>	4076	Over-the-counter pain reliever; not otherwise specified <OTC>
Pain Reliever	4077	Tee's	4077	Tee's
Pain Reliever	4078	Spasmoproxyn	4078	Spasmoproxyn
Pain Reliever	4079	Anbesol <OTC>	4079	Anbesol <OTC>
Pain Reliever	4080	Morphidex	4080	Morphidex
Pain Reliever	4081	Avamigran	4081	Avamigran
Pain Reliever	4082	Analgin, Dipyrone, Metamizol, Metamizole, Novalgin, Prodolin	4082	Analgin, Dipyrone, Metamizol, Metamizole, Novalgin, Prodolin
Pain Reliever	4083	Ultracet	4083	Ultracet
Pain Reliever	4084	Aktren	4084	Aktren
Pain Reliever	4085	Advil Migraine <OTC>	4085	Advil Migraine <OTC>
Pain Reliever	4086	Parafon Forte	4086	Parafon Forte
Pain Reliever	4087	Torbugesic/Torbutrol (veterinary butorphanol) [see also 419]	4087	Torbugesic/Torbutrol (veterinary butorphanol) [see also 419]
Pain Reliever	4088	Synthetic heroin	4088	Synthetic heroin
Pain Reliever	4089	Brompton's solution/mixture	4089	Brompton's solution/mixture
Pain Reliever	4090	Dorixina	4090	Dorixina
Pain Reliever	4091	K4	4091	K4
Pain Reliever	4092	Pain reliever lotion, rub; not specified	4092	Pain reliever lotion, rub; not specified
Pain Reliever	4093	Tylox, Solpadeine	4093	Tylox, Solpadeine
Pain Reliever	4094	Tombstones	4094	Tombstones
Pain Reliever	4095	Dihydrocodeine	4095	Dihydrocodeine
Pain Reliever	4096	Bextra, Valdecoxib	4096	Bextra, Valdecoxib
Pain Reliever	4097	Tussionex	4097	Tussionex
Pain Reliever	4098	Banamine	4098	Banamine
Pain Reliever	4099	Anacin <OTC>	4099	Anacin <OTC>
Pain Reliever	4100	Excedrin Migraine <OTC>	4100	Excedrin Migraine <OTC>
Pain Reliever	4101	Myprodol	4101	Myprodol
Pain Reliever	4102	Meloxicam, Mobic	4102	Meloxicam, Mobic
Pain Reliever	4103	Terpin Hydrate with Codeine	4103	Terpin Hydrate with Codeine
Pain Reliever	4104	Dolo-Neurobion	4104	Dolo-Neurobion
Pain Reliever	4105	Hydrocodone with APAP; otherwise unspecified	4105	Hydrocodone with APAP; otherwise unspecified
Pain Reliever	4106	Atroveran	4106	Atroveran
Pain Reliever	4107	Panlor	4107	Panlor
Pain Reliever	4108	Homemade analgesics	4108	Homemade analgesics
Pain Reliever	4109	Suboxone	4109	Suboxone

Table C.5 Comparison of Codes for Prescription Drugs in 2016 and 2014 (continued)

Prescription Drug Category	2016 Code	2016 Description	2014 Code	2014 Description
Pain Reliever	4110	Avinza	4110	Avinza
Pain Reliever	4111	Clinoril, Sulinac	4111	Clinoril, Sulinac
Pain Reliever	4112	Saridone	4112	Saridone
Pain Reliever	4113	Cosanyl	4113	Cosanyl
Pain Reliever	4114	Nimesulide	4114	Nimesulide
Pain Reliever	4115	Pantopon	4115	Pantopon
Pain Reliever	4116	Phenergan with Codeine, promethazine with Codeine	4116	Phenergan with Codeine, promethazine with Codeine
Pain Reliever	4117	Oxy-IR	4117	Oxy-IR
Pain Reliever	4118	Sufenta, Sufentanil	4118	Sufenta, Sufentanil
Pain Reliever	4119	Buprenorphine	4119	Buprenorphine, Subutex
Pain Reliever	4120	Opana [see also 4510]	4120	Numorphan, Opana, Oxymorphone
Pain Reliever	4121	Dorflex	4121	Dorflex
Pain Reliever	4122	Combiflam	4122	Combiflam
Pain Reliever	4123	Dolorex	4123	Dolorex
Pain Reliever	4124	Relpax	4124	Relpax
Pain Reliever	4125	Hycodan, Hydromet, Tussigon	4125	Hycodan, Hydromet, Tussigon
Pain Reliever	4126	Panafen Plus	4126	Panafen Plus
Pain Reliever	4127	Panadeine	4127	Panadeine
Pain Reliever	4128	Ibuprofen 200 mg/Children's ibuprofen <OTC>	4128	Ibuprofen 200 mg/Children's ibuprofen <OTC>
Pain Reliever	4129	Naproxen Sodium 220 mg <OTC>	4129	Naproxen Sodium 220 mg <OTC>
Pain Reliever	4130	Dalgan, Dezocine	4130	Dalgan, Dezocine
Pain Reliever	4131	Diconal, Dipipanone	4131	Diconal, Dipipanone
Pain Reliever	4132	Embeda	4132	Embeda
Pain Reliever	4133	Tramadol extended release	4133	Tramadol extended release, otherwise specified
Pain Reliever	4134	Ultram ER	4134	Ultram ER
Pain Reliever	4135	Arcoxia, Etoricoxib	4135	Arcoxia, Etoricoxib
Pain Reliever	4136	Sedalmex <OTC>	4136	Sedalmex <OTC>
Pain Reliever	4137	Nucynta	4137	Nucynta
Pain Reliever	4138	Sprix	4138	Sprix
Pain Reliever	4139	Bute tabs, Phenylbutazone	4139	Bute tabs, Phenylbutazone
Pain Reliever	4140	Opana ER (extended release)	4140	Opana ER (extended release)
Pain Reliever	4141	Alfenta, Alfentanil	4141	Alfenta, Alfentanil
Pain Reliever	4142	Remifentanil	4142	Remifentanil
Pain Reliever	4143	Vicotuss	4143	Vicotuss
Pain Reliever	4144	Bancap HC	4144	Bancap HC
Pain Reliever	4145	Daprisal	4145	Daprisal
Pain Reliever	4146	Traumeel	4146	Traumeel
Pain Reliever	4147	Bio Electro pain reliever <OTC>	None	None
Pain Reliever	4148	Butrans transdermal system/patch	None	None
Pain Reliever	4149	Pain patch, otherwise unspecified	None	None

Table C.5 Comparison of Codes for Prescription Drugs in 2016 and 2014 (continued)

Prescription Drug Category	2016 Code	2016 Description	2014 Code	2014 Description
Pain Reliever	4150	Subutex	4119	Buprenorphine, Subutex
Pain Reliever	4152	Flanid	None	None
Pain Reliever	4153	Morphine extended release	None	None
Pain Reliever	4154	Musflex	None	None
Pain Reliever	4155	Numorphone	None	None
Pain Reliever	4156	Apronax	None	None
Pain Reliever	4501	Lortab [see also 434]	434	Vicodin, Lortab, Lorcet, Anexsia, Co-Gesic, Lorcet-HD, Lorcet Plus, Maxidone, Norco, Zydone
Pain Reliever	4502	Norco [see also 434]	434	Vicodin, Lortab, Lorcet, Anexsia, Co-Gesic, Lorcet-HD, Lorcet Plus, Maxidone, Norco, Zydone
Pain Reliever	4503	Lorcet [see also 434]	434	Vicodin, Lortab, Lorcet, Anexsia, Co-Gesic, Lorcet-HD, Lorcet Plus, Maxidone, Norco, Zydone
Pain Reliever	4504	Anexsia [see also 434]	434	Vicodin, Lortab, Lorcet, Anexsia, Co-Gesic, Lorcet-HD, Lorcet Plus, Maxidone, Norco, Zydone
Pain Reliever	4505	Oxycodone product, otherwise unspecified [see also 4046]	4046	Oxycodone or unspecified oxycodone products
Pain Reliever	4506	Codeine, not in pill form [see also 415]	415	Codeine
Pain Reliever	4507	Kadian [see also 4037]	4037	Kadian, MS Contin
Pain Reliever	4508	Duragesic [see also 465]	465	Actiq, Fentanyl, Duragesic, Sublimaze
Pain Reliever	4509	Sublimaze [see also 465]	465	Actiq, Fentanyl, Duragesic, Sublimaze
Pain Reliever	4510	Oxymorphone [see also 4120]	4120	Numorphan, Opana, Oxymorphone
Pain Reliever	4511	Meperidine [see also 407]	407	Demerol, Meperidine
Pain Reliever	4512	Dolophine [see also 418]	418	Methadone, Dolophine
Hallucinogen	6501	Ketaset [code 280 prior to 2015]	280	Ketaset
Other	8501	Haldol, Haloperidol [code 113 prior to 2015]	131	Haldol, Haloperidol
Other	8502	Dimetapp <OTC> [code 136 prior to 2015]	136	Dimetapp <OTC>
Other	8503	Dilantin [code 141 prior to 2015]	141	Dilantin
Other	8504	Actifed <OTC> [code 148 prior to 2015]	148	Actifed <OTC>
Other	8505	Quetiapine, Seroquel [code 178 prior to 2015]	178	Quetiapine, Seroquel
Other	8506	Mysoline [code 183 prior to 2015]	183	Mysoline

Table C.5 Comparison of Codes for Prescription Drugs in 2016 and 2014 (continued)

Prescription Drug Category	2016 Code	2016 Description	2014 Code	2014 Description
Other	8507	Promazine [code 192 prior to 2015]	192	Promazine
Other	8508	Meclizine [code 193 prior to 2015]	193	Meclizine
Other	8509	Triavil [code 224 prior to 2015]	224	Triavil
Other	8510	Desyrel, Trazodone hydrochloride [code 225 prior to 2015]	225	Desyrel, Trazodone hydrochloride
Other	8511	Stelazine, Trifluoperazine HCl [code 226 prior to 2015]	226	Stelazine, Trifluoperazine HCl
Other	8512	Thorazine [code 227 prior to 2015]	227	Thorazine
Other	8513	Amitriptyline, Elavil [code 228 prior to 2015]	228	Amitriptyline, Elavil
Other	8514	Compazine, Prochlorperazine [code 230 prior to 2015]	230	Compazine, Prochlorperazine
Other	8515	Mellaril [code 232 prior to 2015]	232	Mellaril
Other	8516	Trilafon, Perphenazine [code 233 prior to 2015]	233	Trilafon, Perphenazine
Other	8517	Imipramine [code 235 prior to 2015]	235	Imipramine
Other	8518	Lithium carbonate, Lithonate [code 240 prior to 2015]	240	Lithium carbonate, Lithonate
Other	8519	Prozac, Fluoxetine [code 241 prior to 2015]	241	Prozac, Fluoxetine
Other	8520	Etrafon [code 243 prior to 2015]	243	Etrafon
Other	8521	Tryptophan [code 245 prior to 2015]	245	Tryptophan
Other	8522	Adapin, Doxepin, Sinequan [code 247 prior to 2015]	247	Adapin, Doxepin, Sinequan
Other	8523	Asendin [code 249 prior to 2015]	249	Asendin
Other	8524	Aventyl, Nortab, Nortriptyline, Pamelor [code 250 prior to 2015]	250	Aventyl, Nortab, Nortriptyline, Pamelor
Other	8525	Zestril [code 257 prior to 2015]	257	Zestril
Other	8526	Desipramine, Norpramin [code 259 prior to 2015]	259	Desipramine, Norpramin
Other	8527	Prolixin [code 262 prior to 2015]	262	Prolixin
Other	8528	Vivactil [code 263 prior to 2015]	263	Vivactil
Other	8529	Marezine [code 267 prior to 2015]	267	Marezine
Other	8530	Sertraline hydrochloride, Zoloft [code 271 prior to 2015]	271	Sertraline hydrochloride, Zoloft
Other	8531	Loxitane [code 275 prior to 2015]	275	Loxitane
Other	8532	Clozapine, Clozaril [code 282 prior to 2015]	282	Clozapine, Clozaril
Other	8533	Paroxetine, Paxil [code 283 prior to 2015]	283	Paroxetine, Paxil
Other	8534	Tofranil [code 284 prior to 2015]	284	Tofranil

Table C.5 Comparison of Codes for Prescription Drugs in 2016 and 2014 (continued)

Prescription Drug Category	2016 Code	2016 Description	2014 Code	2014 Description
Other	8535	Risperdal, Risperdol, Risperidone [code 285 prior to 2015]	285	Risperdal, Risperdol, Risperidone
Other	8536	Olanzapine, Zyprexa [code 289 prior to 2015]	289	Olanzapine, Zyprexa
Other	8537	Celexa, Citalopram [code 291 prior to 2015]	291	Celexa, Citalopram
Other	8538	Serzone [code 292 prior to 2015]	292	Serzone
Other	8539	Anafranil, Clomipramine [code 293 prior to 2015]	293	Anafranil, Clomipramine
Other	8540	Fluvoxamine, Luvox [code 294 prior to 2015]	294	Fluvoxamine, Luvox
Other	8541	Mirtazapine, Remeron [code 295 prior to 2015]	295	Mirtazapine, Remeron
Other	8542	Acepromazine, PromAce [code 298 prior to 2015]	298	Acepromazine, PromAce
Other	8543	Nitroglycerine [code 343 prior to 2015]	343	Nitroglycerine
Other	8544	Cogentin [code 371 prior to 2015]	371	Cogentin
Other	8545	Dimenhydrinate, Dramamine, sea/motion sickness pill <OTC> [code 1002 prior to 2015]	1002	Dimenhydrinate, Dramamine, sea/motion sickness pill <OTC>
Other	8546	Scopolamine [code 1020 prior to 2015]	1020	Hyoscine, Levo-duboisine, Scopolamine
Other	8547	Seroquel XR [code 1021 prior to 2015]	1021	Seroquel XR
Other	8548	Insidon, Opipramol [code 2005 prior to 2015]	2005	Insidon, Opipramol
Other	8549	Lexapro [code 2011 prior to 2015]	2011	Lexapro
Other	8550	Geodon, Ziprasidone [code 2012 prior to 2015]	2012	Geodon, Ziprasidone
Other	8551	Invega [code 2020 prior to 2015]	2020	Invega
Other	8552	Lidone, Moban, Molindone [code 2021 prior to 2015]	2021	Lidone, Moban, Molindone
Other	8553	Bupropion, Wellbutrin, Zyban [code 3026 prior to 2015]	3026	Bupropion, Wellbutrin, Zyban
Other	8554	Effexor [code 3037 prior to 2015]	3037	Effexor
Other	8555	Spancap [code 3067 prior to 2015]	3067	Spancap
Other	8556	Orlistat, Xenical [code 3079 prior to 2015]	3079	Orlistat, Xenical
Other	8557	Strattera [code 3087 prior to 2015]	3087	Strattera
Other	8558	Wellbutrin XL [code 3108 prior to 2015]	3108	Wellbutrin XL
Other	8559	alli weight loss pills <OTC> [code 3118 prior to 2015]	3118	alli weight loss pills <OTC>
Other	8560	Obitol [code 3121 prior to 2015]	3121	Obitol

Table C.5 Comparison of Codes for Prescription Drugs in 2016 and 2014 (continued)

Prescription Drug Category	2016 Code	2016 Description	2014 Code	2014 Description
Other	8561	Akineton, Biperiden [code 4014 prior to 2015]	4014	Akineton, Biperiden
Other	8562	Flarex [code 4032 [prior to 2015]	4032	Flarex
Other	8563	Suprax [code 4066 prior to 2015]	4066	Suprax
Not Used	None	None	140	Seldane
Not Used	None	None	252	Strawberry whip cream
Not Used	None	None	270	Torzien
Not Used	None	None	393	OP 33
Not Used	None	None	397	F-40

Appendix D: Model Summaries

This page intentionally left blank

Appendix D: Model Summaries

D.1 Introduction

The tables in this appendix list the covariates used in all the imputation models that were run in the 2016 National Survey on Drug Use and Health (NSDUH). For each variable or set of variables to which the predictive mean neighborhood (PMN) imputation method or the modified PMN (modPMN) imputation method was applied, two models were run: one to adjust the weights for item nonresponse (response propensity models) and a second to calculate predicted means. Imputation was usually done separately among age groups. Therefore, most of the tables within this appendix, with the exception of those in Section D.4.2, display only one age group at a time.

The models for the demographic variables are presented in Section D.3 and the models for the drug variables are presented in Section D.4. With the exception of the lifetime usage models, separate tables are provided in Section D.4.1 for each drug age group combination. Within Section D.4.2, separate tables are provided for each measure. Dependence and abuse measures are further broken down into four imputation sets. Tables that present the models for each age group for the household composition variables, which are derived from the questionnaire roster items, are provided in Section D.5. Section D.6 presents the models for the income variables, and Section D.7 presents the models for the health insurance variables. Section D.8 presents the models for the roster pair variables.

The definition of terms and variables in the models of the various sections can be found in the chapters that correspond to those sections as follows: Section D.3 (Chapters 4 & 5), Section D.4 (Chapters 6 & 7), Section D.5 (Chapter 11), Section D.6 (Chapter 12), Section D.7 (Chapter 13), and Section D.8 (Chapter 14).

In the tables, the variable "age" is the mean-centered age, where the age was "centered" by subtracting the mean age and where the mean was calculated across all respondents within the age group who were used to build the given model. The variables "age squared" and "age cubed" represent the square and cube, respectively, of this mean-centered age variable. The variable "Orthogonal Age" is the orthogonal polynomial term for age calculated across all respondents within the domain. The variables "Orthogonal Age Squared" and "Orthogonal Age Cubed" are also orthogonal polynomial terms. Using orthogonal polynomials reduces the severity of multicollinearity in the regression models. Also in the tables, when an asterisk "*" is given, it represents an interaction between two variables.

D.2 Screener, Block-Level, and Segment-Level Variables

In the PMN and modPMN procedures, statistical modeling was performed to adjust weights for item nonresponse and also to calculate predicted means in the imputation models. Descriptions of questionnaire-derived variables are described in the main body of this report, with the exception of screener, block-level, and segment-level variables, which are described below. These variables were often used as covariates in both types of models for the PMN and modPMN procedures.

D.2.1 Household Type

Household type was a three-level race/ethnicity variable based on screener data. It was created by recoding the race/ethnicity of the screening head of the household to one of three levels: Hispanic/Latino, non-Hispanic/Latino black/African American, or non-Hispanic/Latino non-black/African American.

D.2.2 Census Region

Census region was a four-level geographic variable recoded from the respondent's state of residence. The four levels were Northeast, Midwest, South, and West. The Northeast includes Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The Midwest consists of Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin. Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia compose the South. Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming compose the West.

D.2.3 Population Density

The population density variable classifies respondents according to their living situation, whether it be in a rural or urban area and, if urban, the size of the urban area. It was used to categorize segments where the respondents lived according to modified 2010 census data, which were adjusted to reflect population increases between census years by intercensal projections by Claritas, Inc.¹ This variable had five levels: segment in core-based statistical area (CBSA)² with 1 million or more people; segment in CBSA with 250,000 to 999,999 people; segment in CBSA with fewer than 250,000 people; segment not in CBSA and not in rural area; and segment not in CBSA and in rural area.

D.2.4 Percentage Hispanic/Latino in Block

The "Percentage Hispanic/Latino in Block" variable was used to categorize blocks according to the concentration of Hispanics/Latinos in the blocks in which the respondents lived, using the adjusted 2010 census data. It had three levels: less than 20 percent, 20 to less than 71 percent, and 71 percent or more.

D.2.5 Percentage Owner Occupied in Block

The "Percentage Owner Occupied in Block" variable was used to categorize blocks according to the concentration of owner-occupied households in the blocks in which the respondents lived, using the adjusted 2010 census data. It was used as a surrogate for income because wealthy blocks tend to have many homeowners, while poor blocks tend to have many

¹ Claritas, Inc. is a market research firm headquartered in San Diego, California.

² CBSAs, developed in response to standards put forth by the Office of Management and Budget (OMB), are metropolitan and micropolitan areas that were designated using data from the 2000 census. More information about CBSAs can be retrieved from <https://www.census.gov/hhes/www/housing/resscg/cbsa.html>.

renters. It had three levels: less than 10 percent, 10 to less than 50 percent, and 50 percent or more.

D.2.6 Percentage Black/African American in Block

The "Percentage Black/African American in Block" variable was used to categorize blocks according to the concentration of black/African-American households in the blocks in which the respondents lived, using the adjusted 2010 census data. It had three levels: less than 10 percent, 10 to less than 40 percent, and 40 percent or more.

D.2.7 Percentage Asian/Other Pacific Islander in Block

The "Percentage Asian/Other Pacific Islander in Block" variable was used to categorize blocks according to the concentration of Asian/Other Pacific Islander households in the blocks in which the respondents lived, using the adjusted 2010 census data. It had three levels: less than 5 percent, 5 to less than 10 percent, and 10 percent or more.

D.2.8 Percentage American Indian/Alaska Native in Block

The "Percentage American Indian/Alaska Native in Block" variable was used to categorize blocks according to the concentration of American Indian/Alaska Native households in the blocks in which the respondents lived, using the adjusted 2010 census data. It had three levels: less than 1 percent, 1 to less than 3 percent, and 3 percent or more.

D.2.9 Percentage Non-Hispanic/Latino Black/African American in Block

The "Percentage Non-Hispanic/Latino Black/African American in Block" variable was used to categorize blocks according to the concentration of non-Hispanic/Latino black/African-American households in the blocks in which the respondents lived, using the adjusted 2010 census data. It had three levels: less than 10 percent, 10 percent to less than 50 percent, and 50 percent or more.

D.3 Demographic Variables

For justifications of the aggregation of age groups for certain imputation steps, see Chapters 4 and 5.

Exhibit D.1 Definitions of Levels for Categorical Variables

Gender

M: Male, F: Female¹

Age Category

Y: 12-17, T: 18-25, A1: 26-34, A2: 35-49, A3: 50+¹

Race/Hispanic Recode (4 Levels)

W: Not Hispanic/Latino and White Only,¹ B: Not Hispanic/Latino and Black/African American Only, O: Not Hispanic/Latino and Other, H: Hispanic/Latino

Race/Hispanic Recode (5 Levels)

W: White Only,¹ B: Black/African American Only, AI: American Indian/Alaska Native Only, A: Asian/Other Pacific Islander Only, MR: Multiple Race

Exhibit D.1 Definitions of Levels for Categorical Variables (continued)

Race/Hispanic Recode (8 Levels)

W: Not Hispanic/Latino and White,¹ B: Not Hispanic/Latino and Black/African American, AI: Not Hispanic/Latino and American Indian/Alaska Native, A: Not Hispanic/Latino and Asian/Other Pacific Islander, MR: Not Hispanic/Latino and Multiple Race, PR: Hispanic/Latino and Puerto Rican, M: Hispanic/Latino and Mexican, O: Not Hispanic/Latino and Other

Education Level (4 Levels)

L: Less than High School, HS: High School Graduate, SC: Some College, C: College Graduate¹

Immigrant Education Level (4 Levels)

L: Less than High School, HS: High School Graduate, SC: Some College, C: College Graduate¹

Employment Status

FT: Full Time, PT: Part Time, UN: Unemployed, OE: Other Employment¹

Immigrant Employment Status

FT: Full Time, PT: Part Time, UN: Unemployed, OE: Other Employment¹

Marital Status (3 Levels)

M: Married, WM: Was Married, NM: Never Been Married¹

Immigrant Marital Status (3 Levels)

M: Married, WM: Was Married, NM: Never Been Married¹

Marital Status (4 Levels)

M: Married, W: Widowed, DS: Divorced or Separated, NM: Never Been Married¹

Population Density (3 Levels)

R: Segment Not in CBSA, SM: Segment in CBSA with Fewer than 1 Million People, L: Segment in CBSA with 1 Million or More People¹

Population Density (5 Levels)

LC: Segment in CBSA with 1 Million or More People, MC: Segment in CBSA from 250,000 to 999,999 People, SC: Segment in CBSA with Fewer than 250,000 People, NC: Segment Not in CBSA and Not in a Rural Area, R: Segment Not in CBSA and in a Rural Area¹

Census Region

N: Northeast, M: Midwest, S: South, W: West¹

Race/Ethnicity of Householder

B: Black/African American, H: Hispanic/Latino, W: White¹

Percentage Asian/Other Pacific Islander in Block

H: $\geq 10\%$, M: $[5,10)\%$, L: $< 5\%$ ¹

Percentage Black/African American in Block

H: $\geq 40\%$, M: $[10,40)\%$, L: $< 10\%$ ¹

Percentage Hispanic/Latino in Block

H: $\geq 71\%$, M: $[20,71)\%$, L: $< 20\%$ ¹

Percentage American Indian/Alaska Native in Block

H: $\geq 3\%$, M: $[1,3)\%$, L: $< 1\%$ ¹

Percentage Owner Occupied in Block

H: $\geq 50\%$, M: $[10,50)\%$, L: $< 10\%$ ¹

Age Category * Gender

MY: Male 12-17, MT: Male 18-25, MA1: Male 26-34, MA2: Male 35-49, MA3: Male 50+,¹ FY: Female 12-17,¹ FT: Female 18-25,¹ FA1: Female 26-34,¹ FA2: Female 35-49,¹ FA3: Female 50+¹

Gender * Race/Hispanic Recode (8 Levels)

MW: Male Not Hispanic/Latino and White,¹ MB: Male Not Hispanic/Latino and Black/African American, MA: Male Not Hispanic/Latino and Asian/Other Pacific Islander, MAI: Male Not Hispanic/Latino and American Indian/Alaska Native, MMR: Male Not Hispanic/Latino and Multiple Race, MPR: Male Hispanic/Latino and Puerto Rican, MM: Male Hispanic/Latino and Mexican, MO: Male Not Hispanic/Latino and Other, FW: Female Not Hispanic/Latino and White,¹ FB: Female Not Hispanic/Latino and Black/African American,¹ FA: Female Not Hispanic/Latino and Asian/Other Pacific Islander,¹ FAI: Female Not Hispanic/Latino and American Indian/Alaska Native,¹ FMR: Female Not Hispanic/Latino and Multiple Race,¹ FPR: Female Hispanic/Latino and Puerto Rican,¹ FM: Female Hispanic/Latino and Mexican,¹ FO: Female Not Hispanic/Latino and Other¹

NOTE: An asterisk "*" represents an interaction between two variables.

CBSA = core-based statistical area.

¹ This is the reference level for this variable, against which effects of other factor levels are measured.

Table D.1 Summaries for Response Propensity Models

Imputation Step	Variables Included in Response Propensity Model
Marital Status	Gender (M); Age Category (Y, T, A1, A2); Population Density (5 Levels) (LC, MC, SC, NC); Census Region (N, M, S); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M); Age Category * Gender (MY, MT, MA1, MA2)
Race 12-17	Census Region (N, M, S); Race/Ethnicity of Householder (B, H); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M)
Race 18-25	Marital Status (3 Levels) (M, WM); Census Region (N, M, S); Race/Ethnicity of Householder (B, H); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M)
Race 26+	Age Category (A1, A2); Marital Status (4 Levels) (M, W, DS); Census Region (N, M, S); Race/Ethnicity of Householder (B, H); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M)
Hispanic/Latino Origin 12-17	Race/Hispanic Recode (5 Levels) (B, AI, A, MR); Census Region (N, M, S); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M)
Hispanic/Latino Origin 18-25	Race/Hispanic Recode (5 Levels) (B, AI, A, MR); Census Region (N, M, S); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M)
Hispanic/Latino Origin 26+	Age Category (A1, A2); Race/Hispanic Recode (5 Levels) (B, AI, A, MR); Census Region (N, M, S); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M)
Hispanic/Latino Group	Gender (M); Age Category (Y, T, A1, A2); Race/Hispanic Recode (5 Levels) (B, AI, A, MR); Census Region (N, M, S); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M); Age Category * Gender (MY, MT, MA1, MA2)
Education Level 12-17	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Census Region (N, M, S); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M)
Education Level 18+	Gender (M); Age Category (T, A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Census Region (N, M, S); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M); Age Category * Gender (MT, MA1, MA2)

Table D.1 Summaries for Response Propensity Models (continued)

Imputation Step	Variables Included in Response Propensity Model
Employment Status 12-25	Gender (M); Age Category (Y); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Census Region (N, M, S); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M); Age Category * Gender (MY)
Employment Status 26+	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Census Region (N, M, S); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M); Age Category * Gender (MA1, MA2)
Born in US 12-17	Gender (M); Race/Hispanic Recode (8 Levels) (B, AI, A, MR, PR, M, O); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Owner Occupied in Block (H, M)
Born in US 18-25	Gender (M); Race/Hispanic Recode (8 Levels) (B, AI, A, MR, PR, M, O); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Owner Occupied in Block (H, M)
Born in US 26+	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (8 Levels) (B, AI, A, MR, PR, M, O); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Owner Occupied in Block (H, M); Age Category * Gender (MA1, MA2)
Age of Entry	Gender (M); Age Category (Y, T, A1, A2); Race/Hispanic Recode (8 Levels) (B, AI, A, MR, PR, M, O); Immigrant Education Level (4 Levels) (L, HS, SC); Immigrant Employment Status (FT, PT, UN); Immigrant Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Owner Occupied in Block (H, M); Age Category * Gender (MY, MT, MA1, MA2); Gender * Race/Hispanic Recode (8 Levels) (MB, MA, MAI, MMR, MPR, MM, MO)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.1](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.2 Summaries for Predictive Mean Models

Imputation Step	Variables Included in Predictive Mean Model
Marital Status	Gender (M); Age; Age Squared ; Age Cubed ; Population Density (5 Levels) (LC, MC, SC, NC); Census Region (N, M, S); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender
Race 12-17	Age; Age Squared; Census Region (N, M, S); Race/Ethnicity of Householder (B, H) ; Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M)
Race 18-25	Age; Marital Status (3 Levels) (M, WM); Census Region (N, M, S); Race/Ethnicity of Householder (B, H) ; Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M)
Race 26+	Age; Marital Status (4 Levels) (M, W, DS); Census Region (N, M, S); Race/Ethnicity of Householder (B, H) ; Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M)
Hispanic/Latino Origin 12-17	Age; Age Squared; Race/Hispanic Recode (5 Levels) (B, AI, A, MR); Census Region (N, M, S); Race/Ethnicity of Householder (B, H); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M)
Hispanic/Latino Origin 18-25	Age; Age Squared; Age Cubed; Race/Hispanic Recode (5 Levels) (B, AI, A, MR); Marital Status (3 Levels) (M, WM); Census Region (N, M, S); Race/Ethnicity of Householder (B, H); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M)
Hispanic/Latino Origin 26+	Age; Age Squared; Age Cubed; Race/Hispanic Recode (5 Levels) (B, AI, A, MR); Marital Status (4 Levels) (M, W, DS); Census Region (N, M, S); Race/Ethnicity of Householder (B, H); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M)
Hispanic/Latino Group	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (5 Levels) (B, AI, A, MR); Census Region (N, M, S); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender
Education Level 12-17	Gender (M); Age ; Age Squared ; Race/Hispanic Recode (4 Levels) (B, O, H); Census Region (N, M, S); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M)

Table D.2 Summaries for Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model
Education Level 18+	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Marital Status (4 Levels) (M, W, DS); Census Region (N, M, S); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender
Employment Status 12-25	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Census Region (N, M, S); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender
Employment Status 26+	Gender (M); Age; Age Squared; Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Marital Status (4 Levels) (M, W, DS); Census Region (N, M, S); Percentage Asian/Other Pacific Islander in Block (H, M); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage American Indian/Alaska Native in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender
Born in US 12-17	Gender (M); Age; Age Squared; Race/Hispanic Recode (8 Levels) (B, AI, A, MR, PR, M, O); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender
Born in US 18-25	Gender (M); Age; Age Squared; Race/Hispanic Recode (8 Levels) (B, AI, A, MR, PR, M, O); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender
Born in US 26+	Gender (M); Age; Age Squared; Race/Hispanic Recode (8 Levels) (B, AI, A, MR, PR, M, O); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender
Age of Entry	Gender (M); Age; Age Squared; Race/Hispanic Recode (8 Levels) (B, AI, A, MR, PR, M, O); Immigrant Education Level (4 Levels) (L, HS, SC); Immigrant Employment Status (FT, PT, UN); Immigrant Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (8 Levels) (MB, MA, MAI, MMR, MPR, MM, MO)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.1](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

D.4 Drug Variables

Exhibit D.2 Definitions of Levels for Categorical Variables

Gender

M: Male, F: Female¹

Age Category²

Y: 12-17,¹ T: 18-25, A1: 26-34, A2: 35-49, A3: 50+¹

Race/Hispanic Recode (2 Levels)

W: Not Hispanic/Latino and White Only, O: Hispanic/Latino or Other than White Only¹

Race/Hispanic Recode (4 Levels)

W: Not Hispanic/Latino and White Only,¹ B: Not Hispanic/Latino and Black/African American Only, O: Not Hispanic/Latino and Other, H: Hispanic/Latino

Education Level (4 Levels)

L: Less than High School, HS: High School Graduate, SC: Some College, C: College Graduate¹

Employment Status

FT: Full Time, PT: Part Time, UN: Unemployed, OE: Other Employment¹

Marital Status (3 Levels)

M: Married, WM: Was Married, NM: Never Been Married¹

Marital Status (4 Levels)

M: Married, W: Widowed, DS: Divorced/Separated, NM: Never Been Married¹

Population Density (3 Levels)

R: Segment Not in CBSA, SM: Segment in CBSA with Fewer than 1 Million People, L: Segment in CBSA with 1 Million or More People¹

Census Region

N: Northeast, M: Midwest, S: South, W: West¹

Gender * Race/Hispanic Recode (4 Levels)

MW: Male Not Hispanic/Latino and White Only,¹ MB: Male Not Hispanic/Latino and Black/African American Only, MO: Male Not Hispanic/Latino and Other, MH: Male Hispanic/Latino, FW: Female Not Hispanic/Latino and White Only,¹ FB: Female Not Hispanic/Latino and Black/African American Only,¹ FO: Female Not Hispanic/Latino and Other,¹ FH: Female Hispanic/Latino¹

Age * Race/Hispanic Recode (4 Levels)

W: Age * Not Hispanic/Latino and White Only,¹ B: Age * Not Hispanic/Latino and Black/African American Only, O: Age * Not Hispanic/Latino and Other, H: Age * Hispanic/Latino

Age Squared * Race/Hispanic Recode (4 Levels)

W: Age Squared * Not Hispanic/Latino and White Only,¹ B: Age Squared * Not Hispanic/Latino and Black/African American Only, O: Age Squared * Not Hispanic/Latino and Other, H: Age Squared * Hispanic/Latino

Orthogonal Age * Race/Hispanic Recode (4 Levels)

W: Orthogonal Age * Not Hispanic/Latino and White Only,¹ B: Orthogonal Age * Not Hispanic/Latino and Black/African American Only, O: Orthogonal Age * Not Hispanic/Latino and Other, H: Orthogonal Age * Hispanic/Latino

Orthogonal Age Squared * Race/Hispanic Recode (4 Levels)

W: Orthogonal Age Squared * Not Hispanic/Latino and White Only,¹ B: Orthogonal Age Squared * Not Hispanic/Latino and Black/African American, O: Orthogonal Age Squared * Not Hispanic/Latino and Other, H: Orthogonal Age Squared * Hispanic/Latino

Cigarette Lifetime Indicator

Y: Yes, N: No¹

Intermediate Drug-Specific Lifetime Indicator

Y: Yes, N: No¹

Drug-Specific Lifetime Indicator

Y: Yes, N: No¹

Drug-Specific Past Year Indicator

Y: Yes, N: No¹

Intermediate Drug-Specific Past Month Indicator

Y: Yes, N: No¹

Exhibit D.2 Definitions of Levels for Categorical Variables (continued)

Pipe Recency

PM: Past Month, LF: Lifetime but Not Past Month, NU: Lifetime Nonuser¹

Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels)

PM: Past Month, PY: Past Year but Not Past Month, LF: Lifetime but Not Past Year, NU: Lifetime Nonuser¹

Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels)

PM: Past Month, PY: Past Year but Not Past Month, LF: Lifetime but Not Past Year¹

Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels)

PM: Past Month, PY: Past Year but Not Past Month, P3Y: Past 3 Years but Not Past Year, LF: Lifetime but Not Past 3 Years, NU: Lifetime Nonuser¹

Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels)

PM: Past Month, PY: Past Year but Not Past Month, P3Y: Past 3 Years but Not Past Year, LF: Lifetime but Not Past 3 Years¹

Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels)

PY: Past Year Any Use, LF: Lifetime but Not Past Year Any Use, NU: Never Used¹

Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels)

PMM: Past Month Misuse, PYM: Past Year Not Past Month Misuse, LFM: Lifetime Not Past Year Misuse, LNM: Lifetime Any Use Not Misuse, NU: Never Used¹

Drug-Specific Past Year Initiation

Y: Yes, N: No¹

Drug-Specific Past Year Dependence Indicator

Y: Yes, N: No¹

Drug-Specific Past Year Abuse Indicator

Y: Yes, N: No¹

Lifetime State Rank

L: Low State Rank (low tertile), M: Middle State Rank (middle tertile), H: High State Rank (high tertile)¹

Hallucinogen Lifetime State Rank

L: Low State Rank (low tertile), M: Middle State Rank (middle tertile), H: High State Rank (high tertile)¹

Any Use Recency State Rank

L: Low State Rank (low tertile), M: Middle State Rank (middle tertile), H: High State Rank (high tertile)¹

Recency State Rank

L: Low State Rank (low tertile), M: Middle State Rank (middle tertile), H: High State Rank (high tertile)¹

Initiation State Rank

L: Low State Rank (low tertile), M: Middle State Rank (middle tertile), H: High State Rank (high tertile)¹

NOTE: An asterisk "*" represents an interaction between two variables.

CBSA = core-based statistical area.

¹ This is the reference level for this variable, against which effects of other factor levels are measured.

² For models fit separately by age group, the age category recode was used only for the adult models and the reference cell was "50+." For models fit among all respondents in the domain, the reference cell was "12-17."

D.4.1 PMN Imputation Method Tables

The tables in this section present models for drug variables that were imputed using PMN methodology. The drugs included in this section are cigarettes, smokeless tobacco, cigars, pipes, alcohol, marijuana, cocaine, crack, and heroin.

Table D.3 Lifetime Response Propensity Models

Age Group	Variables Included in Response Propensity Model
12-17	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Cigarette Lifetime Indicator (Y)
18-25	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Cigarette Lifetime Indicator (Y)
26+	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Cigarette Lifetime Indicator (Y)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.4 Cigarettes: People 12 to 17 Years Old

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	N/A	N/A
Recency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Recency State Rank (L, M)
12-Month Frequency	N/A	N/A
30-Day Frequency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y)	Daily Users: Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Recency State Rank (L, M) Nondaily Users: Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Recency State Rank (L, M)

Table D.4 Cigarettes: People 12 to 17 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Age at First Use	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y); Drug-Specific 30-Day Frequency for Cigarettes; Recency State Rank (L, M)
Ever Daily Used	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y); Drug-Specific 30-Day Frequency for Cigarettes; Age at First Use for Cigarettes; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

N/A = not applicable.

Table D.5 Cigarettes: People 18 to 25 Years Old

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	N/A	N/A
Recency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Recency State Rank (L, M)
12-Month Frequency	N/A	N/A
30-Day Frequency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y)	Daily Users: Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Recency State Rank (L, M)
		Nondaily Users: Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Recency State Rank (L, M)

Table D.5 Cigarettes: People 18 to 25 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Age at First Use	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y); Drug-Specific 30-Day Frequency for Cigarettes; Recency State Rank (L, M)
Ever Daily Used	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y); Drug-Specific 30-Day Frequency for Cigarettes; Age at First Use for Cigarettes; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

N/A = not applicable.

Table D.6 Cigarettes: People 26 Years or Older

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	N/A	N/A
Recency	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Recency State Rank (L, M)
12-Month Frequency	N/A	N/A
30-Day Frequency	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y)	Daily Users: Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Recency State Rank (L, M)
		Nondaily Users: Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Recency State Rank (L, M)

Table D.6 Cigarettes: People 26 Years or Older (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Age at First Use	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y); Drug-Specific 30-Day Frequency for Cigarettes; Recency State Rank (L, M)
Ever Daily Used	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y); Drug-Specific 30-Day Frequency for Cigarettes; Age at First Use for Cigarettes; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

N/A = not applicable.

Table D.7 Smokeless Tobacco: People 12 to 17 Years Old

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Lifetime State Rank (L, M)
Recency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cigars (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cigars (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)
12-Month Frequency	N/A	N/A
30-Day Frequency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cigars (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cigars (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)

Table D.7 Smokeless Tobacco: People 12 to 17 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Age at First Use	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cigars (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Smokeless Tobacco (PM, PY, P3Y)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Daily Cigarette Use (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Age at First Use for Daily Cigarette Use, and Cigarettes; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

N/A = not applicable.

Table D.8 Smokeless Tobacco: People 18 to 25 Years Old

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Lifetime State Rank (L, M)
Recency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cigars (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cigars (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)
12-Month Frequency	N/A	N/A
30-Day Frequency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S) ; Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cigars (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cigars (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)

Table D.8 Smokeless Tobacco: People 18 to 25 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Age at First Use	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cigars (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Smokeless Tobacco (PM, PY, P3Y)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Daily Cigarette Use (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Age at First Use for Daily Cigarette Use, and Cigarettes; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

N/A = not applicable.

Table D.9 Smokeless Tobacco: People 26 Years or Older

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Lifetime State Rank (L, M)
Recency	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cigars (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cigars (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)
12-Month Frequency	N/A	N/A
30-Day Frequency	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cigars (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cigars (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)

Table D.9 Smokeless Tobacco: People 26 Years or Older (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Age at First Use	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cigars (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Smokeless Tobacco (PM, PY, P3Y)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Daily Cigarette Use (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Age at First Use for Daily Cigarette Use, and Cigarettes; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

N/A = not applicable.

Table D.10 Cigars: People 12 to 17 Years Old

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Smokeless Tobacco (Y); Lifetime State Rank (L, M)
Recency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Alcohol (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)
12-Month Frequency	N/A	N/A
30-Day Frequency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Alcohol (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF), and Smokeless Tobacco (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF), and Smokeless Tobacco (PM, PY, P3Y, LF); Recency State Rank (L, M)

Table D.10 Cigars: People 12 to 17 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Age at First Use	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Alcohol (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF), and Smokeless Tobacco (PM, PY, P3Y, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Daily Cigarette Use (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF), and Smokeless Tobacco (PM, PY, P3Y, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y); Drug-Specific 30-Day Frequency for Cigars; Drug-Specific Age at First Use for Daily Cigarette Use, Smokeless Tobacco, and Cigarettes; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

N/A = not applicable.

Table D.11 Cigars: People 18 to 25 Years Old

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Smokeless Tobacco (Y); Lifetime State Rank (L, M)
Recency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Alcohol (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)
12-Month Frequency	N/A	N/A
30-Day Frequency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Alcohol (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF), and Smokeless Tobacco (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF), and Smokeless Tobacco (PM, PY, P3Y, LF); Recency State Rank (L, M)

Table D.11 Cigars: People 18 to 25 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Age at First Use	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Alcohol (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF), and Smokeless Tobacco (PM, PY, P3Y, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Daily Cigarette Use (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF), and Smokeless Tobacco (PM, PY, P3Y, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y); Drug-Specific 30-Day Frequency for Cigars; Drug-Specific Age at First Use for Daily Cigarette Use, Smokeless Tobacco, and Cigarettes; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

N/A = not applicable.

Table D.12 Cigars: People 26 Years or Older

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Smokeless Tobacco (Y); Lifetime State Rank (L, M)
Recency	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Alcohol (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)
12-Month Frequency	N/A	N/A
30-Day Frequency	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Alcohol (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF), and Smokeless Tobacco (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF), and Smokeless Tobacco (PM, PY, P3Y, LF); Recency State Rank (L, M)

Table D.12 Cigars: People 26 Years or Older (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Age at First Use	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Alcohol (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF), and Smokeless Tobacco (PM, PY, P3Y, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), and Daily Cigarette Use (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF), and Smokeless Tobacco (PM, PY, P3Y, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y); Drug-Specific 30-Day Frequency for Cigars; Drug-Specific Age at First Use for Daily Cigarette Use, Smokeless Tobacco, and Cigarettes; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

N/A = not applicable.

Table D.13 Pipes: People 12 to 17 Years Old

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Cigars (Y), and Smokeless Tobacco (Y); Lifetime State Rank (L, M)
Recency	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
12-Month Frequency	N/A	N/A
30-Day Frequency	N/A	N/A
Age at First Use	N/A	N/A

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

N/A = not applicable.

Table D.14 Pipes: People 18 to 25 Years Old

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Cigars (Y), and Smokeless Tobacco (Y); Lifetime State Rank (L, M)
Recency	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
12-Month Frequency	N/A	N/A
30-Day Frequency	N/A	N/A
Age at First Use	N/A	N/A

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

N/A = not applicable.

Table D.15 Pipes: People 26 Years or Older

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Cigars (Y), and Smokeless Tobacco (Y); Lifetime State Rank (L, M)
Recency	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)
12-Month Frequency	N/A	N/A
30-Day Frequency	N/A	N/A
Age at First Use	N/A	N/A

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

N/A = not applicable.

Table D.16 Alcohol: People 12 to 17 Years Old

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Cigars (Y), Smokeless Tobacco (Y), and Pipes (Y); Lifetime State Rank (L, M)
Recency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)
12-Month Frequency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Intermediate Drug-Specific Past Month Indicator for Alcohol (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Intermediate Drug-Specific Past Month Indicator for Alcohol (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)

Table D.16 Alcohol: People 12 to 17 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
30-Day Frequency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Alcohol	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Alcohol; Recency State Rank (L, M)
Age at First Use	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y), Crack (Y), Heroin (Y), Marijuana (Y), and Cocaine (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Alcohol; Drug-Specific 30-Day Frequency for Alcohol; Drug-Specific Age at First Use for Daily Cigarette Use, Cigarettes, Smokeless Tobacco, and Cigars; Recency State Rank (L, M)

Table D.16 Alcohol: People 12 to 17 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
30-Day Binge Drinking Frequency	Males: Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Alcohol Females: Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Alcohol	Males: Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Alcohol; Drug-Specific 30-Day Frequency for Alcohol; Recency State Rank (L, M) Females: Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Alcohol; Drug-Specific 30-Day Frequency for Alcohol; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.17 Alcohol: People 18 to 25 Years Old

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Cigars (Y), Smokeless Tobacco (Y), and Pipes (Y); Lifetime State Rank (L, M)
Recency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)
12-Month Frequency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Intermediate Drug-Specific Past Month Indicator for Alcohol (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Intermediate Drug-Specific Past Month Indicator for Alcohol (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)

Table D.17 Alcohol: People 18 to 25 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
30-Day Frequency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Alcohol	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Alcohol; Recency State Rank (L, M)
Age at First Use	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y), Crack (Y), Heroin (Y), Marijuana (Y), and Cocaine (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Alcohol; Drug-Specific 30-Day Frequency for Alcohol; Drug-Specific Age at First Use for Daily Cigarette Use, Cigarettes, Smokeless Tobacco, and Cigars; Recency State Rank (L, M)

Table D.17 Alcohol: People 18 to 25 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
30-Day Binge Drinking Frequency	Males: Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Alcohol Females: Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Alcohol	Males: Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Alcohol; Drug-Specific 30-Day Frequency for Alcohol; Recency State Rank (L, M) Females: Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Alcohol; Drug-Specific 30-Day Frequency for Alcohol; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.18 Alcohol: People 26 Years or Older

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Cigars (Y), Smokeless Tobacco (Y), and Pipes (Y); Lifetime State Rank (L, M)
Recency	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)
12-Month Frequency	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Intermediate Drug-Specific Past Month Indicator for Alcohol (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Intermediate Drug-Specific Past Month Indicator for Alcohol (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)

Table D.18 Alcohol: People 26 Years or Older (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
30-Day Frequency	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Alcohol	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Alcohol; Recency State Rank (L, M)
Age at First Use	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y), Crack (Y), Heroin (Y), Marijuana (Y), and Cocaine (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Alcohol; Drug-Specific 30-Day Frequency for Alcohol; Drug-Specific Age at First Use for Daily Cigarette Use, Cigarettes, Smokeless Tobacco, and Cigars; Recency State Rank (L, M)

Table D.18 Alcohol: People 26 Years or Older (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
30-Day Binge Drinking Frequency	Males: Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Alcohol Females: Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Alcohol	Males: Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Alcohol; Drug-Specific 30-Day Frequency for Alcohol; Recency State Rank (L, M) Females: Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), Marijuana (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Alcohol; Drug-Specific 30-Day Frequency for Alcohol; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.19 Marijuana: People 12 to 17 Years Old

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Alcohol (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Lifetime State Rank (L, M)
Recency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)
12-Month Frequency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), and Crack (Y); Intermediate Drug-Specific Past Month Indicator for Marijuana (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), and Crack (Y); Intermediate Drug-Specific Past Month Indicator for Marijuana (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)

Table D.19 Marijuana: People 12 to 17 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
30-Day Frequency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Marijuana	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Marijuana; Recency State Rank (L, M)
Age at First Use	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Marijuana (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y), Crack (Y), Heroin (Y), and Cocaine (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Marijuana (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Marijuana; Drug-Specific 30-Day Frequency for Marijuana; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.20 Marijuana: People 18 to 25 Years Old

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Alcohol (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Lifetime State Rank (L, M)
Recency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)

Table D.20 Marijuana: People 18 to 25 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
12-Month Frequency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), and Crack (Y); Intermediate Drug-Specific Past Month Indicator for Marijuana (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), and Crack (Y); Intermediate Drug-Specific Past Month Indicator for Marijuana (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)
30-Day Frequency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Marijuana	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Marijuana; Recency State Rank (L, M)

Table D.20 Marijuana: People 18 to 25 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Age at First Use	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Marijuana (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y), Crack (Y), Heroin (Y), and Cocaine (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Marijuana (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Marijuana; Drug-Specific 30-Day Frequency for Marijuana; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.21 Marijuana: People 26 Years or Older

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Gender (M); Age; Age Squared; Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Alcohol (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Lifetime State Rank (L, M)
Recency	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)

Table D.21 Marijuana: People 26 Years or Older (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
12-Month Frequency	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), and Crack (Y); Intermediate Drug-Specific Past Month Indicator for Marijuana (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), and Crack (Y); Intermediate Drug-Specific Past Month Indicator for Marijuana (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)
30-Day Frequency	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Marijuana	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Marijuana; Recency State Rank (L, M)

Table D.21 Marijuana: People 26 Years or Older (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Age at First Use	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cocaine (Y), Heroin (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Marijuana (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y), Crack (Y), Heroin (Y), and Cocaine (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Marijuana (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Marijuana; Drug-Specific 30-Day Frequency for Marijuana; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.22 Cocaine (and Crack): People 12 to 17 Years Old

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Cocaine: Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Alcohol (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Lifetime State Rank (L, M) Crack: Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Alcohol (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Lifetime State Rank (L, M)
Recency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Crack (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Crack (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)

Table D.22 Cocaine (and Crack): People 12 to 17 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
12-Month Frequency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Crack (Y), and Heroin (Y); Intermediate Drug-Specific Past Month Indicator for Cocaine (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Crack (Y), and Heroin (Y); Intermediate Drug-Specific Past Month Indicator for Cocaine (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)
30-Day Frequency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Crack (Y), and Heroin (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Cocaine	Age; Census Region (N, M, S); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Cocaine

Table D.22 Cocaine (and Crack): People 12 to 17 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Age at First Use	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Crack (Y), and Heroin (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Cocaine (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y), Heroin (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Cocaine (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Cocaine; Drug-Specific 30-Day Frequency for Cocaine; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Marijuana, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.23 Cocaine (and Crack): People 18 to 25 Years Old

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Cocaine: Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Alcohol (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Lifetime State Rank (L, M) Crack: Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Alcohol (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Lifetime State Rank (L, M)
Recency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Crack (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Crack (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)

Table D.23 Cocaine (and Crack): People 18 to 25 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
12-Month Frequency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Crack (Y), and Heroin (Y); Intermediate Drug-Specific Past Month Indicator for Cocaine (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Crack (Y), and Heroin (Y); Intermediate Drug-Specific Past Month Indicator for Cocaine (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)
30-Day Frequency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Crack (Y), and Heroin (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Cocaine	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Crack (Y), and Heroin (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Cocaine; Recency State Rank (L, M)

Table D.23 Cocaine (and Crack): People 18 to 25 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Age at First Use	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Crack (Y), and Heroin (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Cocaine (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y), Heroin (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Cocaine (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Cocaine; Drug-Specific 30-Day Frequency for Cocaine; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Marijuana, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.24 Cocaine (and Crack): People 26 Years or Older

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Cocaine: Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Alcohol (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Lifetime State Rank (L, M) Crack: Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Alcohol (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Lifetime State Rank (L, M)
Recency	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Crack (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Crack (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)

Table D.24 Cocaine (and Crack): People 26 Years or Older (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
12-Month Frequency	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Crack (Y), and Heroin (Y); Intermediate Drug-Specific Past Month Indicator for Cocaine (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Crack (Y), and Heroin (Y); Intermediate Drug-Specific Past Month Indicator for Cocaine (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)
30-Day Frequency	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Crack (Y), and Heroin (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Cocaine	Gender (M); Age; Race/Hispanic Recode (4 Levels) (B, O, H); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Crack (Y); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Cocaine

Table D.24 Cocaine (and Crack): People 26 Years or Older (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Age at First Use	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Crack (Y), and Heroin (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Cocaine (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y), Heroin (Y), and Crack (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Cocaine (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Cocaine; Drug-Specific 30-Day Frequency for Cocaine; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Marijuana, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.25 Heroin: People 12 to 17 Years Old

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Gender (M); Age; Age Squared ; Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y) ; Lifetime State Rank (L, M)
Recency: Past Year vs. Not Past Year	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Recency State Rank (L, M)
Recency: Past Month vs. Past Year Not Past Month	No Model Used; No Nonrespondents	Gender (M) ; Age; Age Squared ; Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Recency State Rank (L, M)

Table D.25 Heroin: People 12 to 17 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
12-Month Frequency	No Model Used; No Nonrespondents	Gender (M); Age; Age Squared ; Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S) ; Age * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Intermediate Drug-Specific Past Month Indicator for Heroin (Y); Pipe Recency (PM, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)
30-Day Frequency	No Model Used; No Nonrespondents	Age; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Crack (PM), and Marijuana (PM)

Table D.25 Heroin: People 12 to 17 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Age at First Use	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Heroin (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Heroin (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Heroin; Drug-Specific 30-Day Frequency for Heroin; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Marijuana, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.26 Heroin: People 18 to 25 Years Old

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Lifetime State Rank (L, M)
Recency: Past Year vs. Not Past Year	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)

Table D.26 Heroin: People 18 to 25 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Recency: Past Month vs. Past Year Not Past Month	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)
12-Month Frequency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Intermediate Drug-Specific Past Month Indicator for Heroin (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Intermediate Drug-Specific Past Month Indicator for Heroin (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)

Table D.26 Heroin: People 18 to 25 Years Old (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
30-Day Frequency	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S) ; Pipe Recency (PM, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF) ; Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Heroin	Race/Hispanic Recode (2 Levels) (W); Population Density (3 Levels) (R); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Crack (PM); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Heroin
Age at First Use	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S) ; Pipe Recency (PM, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF) ; Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Heroin (PM, PY) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared ; Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN) ; Marital Status (3 Levels) (M, WM) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S) ; Age * Gender; Age Squared * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y) ; Pipe Recency (PM, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF) ; Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Heroin (PM, PY) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Heroin ; Drug-Specific 30-Day Frequency for Heroin ; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Marijuana, Smokeless Tobacco, and Daily Cigarette Use ; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.27 Heroin: People 26 Years or Older

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Lifetime	See Table D.3 .	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Lifetime State Rank (L, M)
Recency: Past Year vs. Not Past Year	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Recency State Rank (L, M)

Table D.27 Heroin: People 26 Years or Older (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Recency: Past Month vs. Past Year Not Past Month	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S) ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Drug-Specific Lifetime Indicator for Cigars (Y) , Crack (Y) , Pipes (Y) , Smokeless Tobacco (Y) , and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF) , and Marijuana (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Recency State Rank (L, M)	Gender (M); Age; Age Squared ; Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y) , Crack (Y) , Pipes (Y) , Smokeless Tobacco (Y) , and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF) , and Marijuana (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Recency State Rank (L, M)
12-Month Frequency	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S); Intermediate Drug-Specific Past Month Indicator for Heroin (Y); Pipe Recency (PM, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF) , Crack (PM, PY, LF) , Marijuana (PM, PY, LF) , and Cocaine (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF) , Smokeless Tobacco (PM, PY, P3Y, LF) , and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Age * Race/Hispanic Recode (4 Levels) (B, O, H); Intermediate Drug-Specific Past Month Indicator for Heroin (Y) ; Pipe Recency (PM, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF) , Crack (PM, PY, LF) , Marijuana (PM, PY, LF) , and Cocaine (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF) , Smokeless Tobacco (PM, PY, P3Y, LF) , and Cigarettes (PM, PY, P3Y, LF) ; Recency State Rank (L, M)

Table D.27 Heroin: People 26 Years or Older (continued)

Imputation Step	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
30-Day Frequency	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Heroin	Gender (M); Employment Status (FT); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Cocaine (PM); Intermediate Drug-Specific 12-Month Frequency, Domain Restricted to Past Month Users for Heroin
Age at First Use	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Heroin (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF)	Gender (M); Age; Age Squared; Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Heroin (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Heroin; Drug-Specific 30-Day Frequency for Heroin; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Marijuana, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

D.4.2 Modified PMN Imputation Method Tables

Under the modified PMN (modPMN) imputation methodology, the response propensity step is used to segregate respondents and nonrespondents into five bins of roughly equal size based on their estimated response propensity. The prediction step is then performed separately within each bin that includes at least one item nonrespondent. This is in contrast to PMN, which, generally, performs the response propensity and prediction steps separately by age group. In addition, imputation sets are generally defined across drugs within a measure (e.g., 30-day frequency of use) under the modPMN methodology and a cycling approach is implemented for variables within an imputation set. As a result, the tables within this section are organized by the drug measure used to define an imputation set (e.g., lifetime use). For tables presenting the response propensity models, the models from both cycles are presented. Similarly, tables containing the predictive mean models will show the models from both cycles for each of the five bins. Within each table, drugs are listed in the order in which they are imputed. This ordering is dynamic and can change from year to year depending on the item response rates of variables within the set. An exception to this is the imputation set containing lifetime hallucinogen child drug variables, which has a static order. ModPMN was applied to inhalants, hallucinogens, methamphetamine, pain relievers, tranquilizers, stimulants, and sedatives. Although lifetime hallucinogen child drug variables were imputed using PMN methodology, these tables have been included in the modPMN section for proximity to the parent drug tables and other measures of hallucinogen child drug use (e.g., recency of use and age at first use), which are imputed using the modPMN methodology. For sets with a single drug, cycling was not necessary.

Table D.28 Lifetime Use Response Propensity Models

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Methamphetamine	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Cigarette Lifetime Indicator (Y); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Hallucinogens (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Inhalants (Y)
Inhalants	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Methamphetamine (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Alcohol (Y)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Hallucinogens (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Alcohol (Y)
Hallucinogens	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Inhalants (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Pipes (Y), Smokeless Tobacco (Y), and Methamphetamine (Y)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Cigarette Lifetime Indicator (Y); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.29 Lifetime Use Predictive Mean Models

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Methamphetamine		
Bin 1	No Model Used; No Nonrespondents	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Hallucinogens (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Pipes (Y), Smokeless Tobacco (Y), and Inhalants (Y); Lifetime State Rank (L, M)
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM) ; Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y) ; Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Hallucinogens (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Pipes (Y), Smokeless Tobacco (Y), and Inhalants (Y) ; Lifetime State Rank (L, M)

Table D.29 Lifetime Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y) , Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Hallucinogens (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y) , Pipes (Y), Smokeless Tobacco (Y), and Inhalants (Y); Lifetime State Rank (L, M)
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y) , Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Hallucinogens (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y) , Pipes (Y), Smokeless Tobacco (Y), and Inhalants (Y); Lifetime State Rank (L, M)
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y) , Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Hallucinogens (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y) , Pipes (Y), Smokeless Tobacco (Y), and Inhalants (Y); Lifetime State Rank (L, M)

Table D.29 Lifetime Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Inhalants		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Methamphetamine (Y), Cigars (Y), Cocaine (Y), Crack (Y) , Pipes (Y), Smokeless Tobacco (Y), and Alcohol (Y); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Hallucinogens (Y), Cigars (Y), Cocaine (Y), Crack (Y) , Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Alcohol (Y); Lifetime State Rank (L, M)
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Methamphetamine (Y), Cigars (Y), Cocaine (Y), Crack (Y) , Pipes (Y), Smokeless Tobacco (Y), and Alcohol (Y); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Hallucinogens (Y), Cigars (Y), Cocaine (Y), Crack (Y) , Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Alcohol (Y); Lifetime State Rank (L, M)

Table D.29 Lifetime Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Methamphetamine (Y), Cigars (Y), Cocaine (Y), Crack (Y) , Pipes (Y), Smokeless Tobacco (Y), and Alcohol (Y); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Hallucinogens (Y), Cigars (Y), Cocaine (Y), Crack (Y) , Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Alcohol (Y); Lifetime State Rank (L, M)
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM) ; Census Region (N, M, S); Orthogonal Age * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Methamphetamine (Y), Cigars (Y), Cocaine (Y), Crack (Y) , Pipes (Y), Smokeless Tobacco (Y), and Alcohol (Y); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM) ; Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Hallucinogens (Y), Cigars (Y), Cocaine (Y), Crack (Y) , Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Alcohol (Y); Lifetime State Rank (L, M)

Table D.29 Lifetime Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Hallucinogens		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Inhalants (Y) , Alcohol (Y), Cigars (Y), Cocaine (Y) , Crack (Y), Pipes (Y), Smokeless Tobacco (Y), and Methamphetamine (Y); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Drug-Specific Lifetime Indicator for Alcohol (Y) , Cocaine (Y) , Crack (Y), Inhalants (Y) , Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Lifetime State Rank (L, M)
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Inhalants (Y) , Alcohol (Y), Cigars (Y), Cocaine (Y) , Crack (Y) , Pipes (Y), Smokeless Tobacco (Y), and Methamphetamine (Y) ; Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Drug-Specific Lifetime Indicator for Alcohol (Y) , Cocaine (Y), Crack (Y) , Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Lifetime State Rank (L, M)

Table D.29 Lifetime Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Inhalants (Y) , Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y) , Pipes (Y), Smokeless Tobacco (Y), and Methamphetamine (Y); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y) ; Lifetime State Rank (L, M)
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Inhalants (Y) , Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y) , Pipes (Y), Smokeless Tobacco (Y), and Methamphetamine (Y) ; Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y) ; Lifetime State Rank (L, M)

Table D.29 Lifetime Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Inhalants (Y) , Alcohol (Y), Cigars (Y), Cocaine (Y) , Crack (Y) , Pipes (Y), Smokeless Tobacco (Y), and Methamphetamine (Y); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Cigarette Lifetime Indicator (Y); Drug-Specific Lifetime Indicator for Alcohol (Y) , Cocaine (Y), Crack (Y) , Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Lifetime State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.30 Hallucinogen Child Drugs Lifetime Response Propensity Models

Age Group	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
12-17	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Cigarette Lifetime Indicator (Y); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y)	N/A
18-25	Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Cigarette Lifetime Indicator (Y); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y)	N/A
26+	Gender (M); Age Category (A1, A2); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Cigarette Lifetime Indicator (Y); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y)	N/A

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

N/A = not applicable.

Table D.31 Hallucinogen Child Drugs Lifetime Use Predictive Mean Models: People 12 to 17 Years Old

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
LSD	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Hallucinogen Lifetime State Rank (L, M)	N/A
PCP	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for LSD (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Alcohol (Y); Hallucinogen Lifetime State Rank (L, M)	N/A
Ecstasy	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for LSD (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and PCP (Y); Hallucinogen Lifetime State Rank (L, M)	N/A

Table D.31 Hallucinogen Child Drugs Lifetime Use Predictive Mean Models: People 12 to 17 Years Old (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Ketamine	No Model Used; No Nonrespondents	N/A
DMT/AMT/Foxy	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Ecstasy (Y), LSD (Y), PCP (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Ketamine (Y); Hallucinogen Lifetime State Rank (L, M)	N/A
<i>Salvia divinorum</i>	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for DMT/AMT/Foxy (Y), Ketamine (Y), LSD (Y), PCP (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Ecstasy (Y); Hallucinogen Lifetime State Rank (L, M)	N/A

Table D.31 Hallucinogen Child Drugs Lifetime Use Predictive Mean Models: People 12 to 17 Years Old (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Other Hallucinogens	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for DMT/AMT/Foxy (Y), Ketamine (Y), LSD (Y), PCP (Y), Salvia (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Ecstasy (Y); Hallucinogen Lifetime State Rank (L, M)	N/A

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

N/A = not applicable.

Table D.32 Hallucinogen Child Drugs Lifetime Use Predictive Mean Models: People 18 to 25 Years Old

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
LSD	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Hallucinogen Lifetime State Rank (L, M)	N/A
PCP	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for LSD (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Alcohol (Y); Hallucinogen Lifetime State Rank (L, M)	N/A

Table D.32 Hallucinogen Child Drugs Lifetime Use Predictive Mean Models: People 18 to 25 Years Old (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Ecstasy	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for LSD (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and PCP (Y); Hallucinogen Lifetime State Rank (L, M)	N/A
Ketamine	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Ecstasy (Y), PCP (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and LSD (Y); Hallucinogen Lifetime State Rank (L, M)	N/A

Table D.32 Hallucinogen Child Drugs Lifetime Use Predictive Mean Models: People 18 to 25 Years Old (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
DMT/AMT/Foxy	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Ecstasy (Y), LSD (Y), PCP (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Ketamine (Y); Hallucinogen Lifetime State Rank (L, M)	N/A
<i>Salvia divinorum</i>	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for DMT/AMT/Foxy (Y), Ketamine (Y), LSD (Y), PCP (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Ecstasy (Y); Hallucinogen Lifetime State Rank (L, M)	N/A

Table D.32 Hallucinogen Child Drugs Lifetime Use Predictive Mean Models: People 18 to 25 Years Old (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Other Hallucinogens	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for DMT/AMT/Foxy (Y), Ketamine (Y), LSD (Y), PCP (Y), Salvia (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Ecstasy (Y); Hallucinogen Lifetime State Rank (L, M)	N/A

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

N/A = not applicable.

Table D.33 Hallucinogen Child Drugs Lifetime Use Predictive Mean Models: People 26 Years or Older

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
LSD	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Drug-Specific Lifetime Indicator for Alcohol (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cigars (Y); Hallucinogen Lifetime State Rank (L, M)	N/A
PCP	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for LSD (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Alcohol (Y); Hallucinogen Lifetime State Rank (L, M)	N/A

Table D.33 Hallucinogen Child Drugs Lifetime Use Predictive Mean Models: People 26 Years or Older (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Ecstasy	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for LSD (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and PCP (Y); Hallucinogen Lifetime State Rank (L, M)	N/A
Ketamine	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Ecstasy (Y), PCP (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and LSD (Y); Hallucinogen Lifetime State Rank (L, M)	N/A

Table D.33 Hallucinogen Child Drugs Lifetime Use Predictive Mean Models: People 26 Years or Older (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
DMT/AMT/Foxy	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for Ecstasy (Y), LSD (Y), PCP (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Ketamine (Y); Hallucinogen Lifetime State Rank (L, M)	N/A
<i>Salvia divinorum</i>	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for DMT/AMT/Foxy (Y), Ketamine (Y), LSD (Y), PCP (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Ecstasy (Y); Hallucinogen Lifetime State Rank (L, M)	N/A

Table D.33 Hallucinogen Child Drugs Lifetime Use Predictive Mean Models: People 26 Years or Older (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Other Hallucinogens	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Cigarette Lifetime Indicator (Y); Intermediate Drug-Specific Lifetime Indicator for DMT/AMT/Foxy (Y), Ketamine (Y), LSD (Y), PCP (Y), Salvia (Y), Alcohol (Y), Cigars (Y), Cocaine (Y), Crack (Y), Heroin (Y), Inhalants (Y), Marijuana (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Ecstasy (Y); Hallucinogen Lifetime State Rank (L, M)	N/A

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

N/A = not applicable.

Table D.34 Recency-of-Any-Use Response Propensity Models

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Stimulants	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Any Use Recency State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF); Any Use Recency State Rank (L, M)
Sedatives	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Stimulants (PY, LF); Any Use Recency State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)

Table D.34 Recency-of-Any-Use Response Propensity Models (continued)

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Tranquilizers	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), and Sedatives (PY, LF); Any Use Recency State Rank (L, M)
Pain Relievers	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.35 Recency-of-Any-Use Predictive Mean Models

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Stimulants		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Any Use Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF); Any Use Recency State Rank (L, M)

Table D.35 Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Any Use Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF); Any Use Recency State Rank (L, M)

Table D.35 Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Any Use Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF); Any Use Recency State Rank (L, M)

Table D.35 Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Any Use Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF); Any Use Recency State Rank (L, M)

Table D.35 Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT , UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Any Use Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF); Any Use Recency State Rank (L, M)

Table D.35 Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Sedatives		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Stimulants (PY, LF); Any Use Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)

Table D.35 Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Stimulants (PY, LF); Any Use Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)

Table D.35 Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Stimulants (PY, LF); Any Use Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)

Table D.35 Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Stimulants (PY, LF); Any Use Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)

Table D.35 Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Stimulants (PY, LF); Any Use Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)

Table D.35 Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Tranquilizers		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), and Sedatives (PY, LF); Any Use Recency State Rank (L, M)

Table D.35 Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), and Sedatives (PY, LF); Any Use Recency State Rank (L, M)

Table D.35 Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), and Sedatives (PY, LF); Any Use Recency State Rank (L, M)

Table D.35 Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), and Sedatives (PY, LF); Any Use Recency State Rank (L, M)

Table D.35 Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), and Sedatives (PY, LF); Any Use Recency State Rank (L, M)

Table D.35 Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Pain Relievers		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)

Table D.35 Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)

Table D.35 Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)

Table D.35 Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)

Table D.35 Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.36 OxyContin® Recency-of-Any-Use Response Propensity Models

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
OxyContin®	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)	N/A

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

N/A = not applicable.

Table D.37 OxyContin® Recency-of-Any-Use Predictive Mean Models

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
OxyContin®		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)	N/A

Table D.37 OxyContin® Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)	N/A

Table D.37 OxyContin® Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)	N/A

Table D.37 OxyContin® Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)	N/A

Table D.37 OxyContin® Recency-of-Any-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Any Use Recency State Rank (L, M)	N/A

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

N/A = not applicable.

Table D.38 Recency-of-Use/Misuse Response Propensity Models

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Methamphetamine	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF); Recency State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)
Hallucinogens	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Inhalants (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Methamphetamine (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF); Recency State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Inhalants (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.38 Recency-of-Use/Misuse Response Propensity Models (continued)

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Inhalants	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF); Recency State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)
Pain Relievers	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Recency State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.38 Recency-of-Use/Misuse Response Propensity Models (continued)

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Stimulants	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), and Tranquilizers (PY, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)
Sedatives	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Tranquilizers (PY, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.38 Recency-of-Use/Misuse Response Propensity Models (continued)

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Tranquilizers	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.39 Recency-of-Use/Misuse Predictive Mean Models

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Methamphetamine		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF); Recency State Rank (L, M)	No Model Used; No Nonrespondents

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Hallucinogens		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Inhalants (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Methamphetamine (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Inhalants (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S) ; Orthogonal Age * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Inhalants (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Methamphetamine (PM, PY, LF), and Marijuana (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF) ; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Inhalants (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM) ; Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Inhalants (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Methamphetamine (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Inhalants (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM) ; Census Region (N, M, S); Orthogonal Age * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y) , Crack (Y), Heroin (Y), Inhalants (Y), Pipes (Y), Smokeless Tobacco (Y) , and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF) , Methamphetamine (PM, PY, LF), and Marijuana (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF) , Stimulants (PY, LF) , Tranquilizers (PY, LF) , and Sedatives (PY, LF) ; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF) , Marijuana (PM, PY, LF) , Methamphetamine (PM, PY, LF), and Inhalants (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM) , Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM) ; Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S) ; Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Inhalants (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Methamphetamine (PM, PY, LF), and Marijuana (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF) ; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Inhalants (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM) ; Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Inhalants		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF) ; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S) ; Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM) ; Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF) ; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM) ; Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Pain Relievers (PY, LF), Stimulants (PY, LF), Tranquilizers (PY, LF), and Sedatives (PY, LF); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Pain Relievers		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S) ; Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), Tranquilizers (PY, LF), and Stimulants (PY, LF); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Stimulants		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), and Tranquilizers (PY, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), and Tranquilizers (PY, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), and Tranquilizers (PY, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), and Tranquilizers (PY, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Sedatives (PY, LF), and Tranquilizers (PY, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Sedatives		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Tranquilizers (PY, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Tranquilizers (PY, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Tranquilizers (PY, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Tranquilizers (PY, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Any Use Recency, Domain Not Restricted to Lifetime Users (2 Levels) for Tranquilizers (PY, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Tranquilizers		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S) ; Orthogonal Age * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM) ; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM) ; Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)

Table D.39 Recency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M) ; Orthogonal Age; Orthogonal Age Squared ; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S) ; Orthogonal Age * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y) , Crack (Y) , Heroin (Y) , Pipes (Y) , Smokeless Tobacco (Y) , and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF) , Inhalants (PM, PY, LF) , Marijuana (PM, PY, LF) , Methamphetamine (PM, PY, LF) , and Hallucinogens (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM) , Stimulants (PMM, PYM, LFM, LNM) , and Sedatives (PMM, PYM, LFM, LNM) ; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y) , Crack (Y) , Heroin (Y) , Pipes (Y) , Smokeless Tobacco (Y) , and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF) , Inhalants (PM, PY, LF) , Marijuana (PM, PY, LF) , Methamphetamine (PM, PY, LF) , and Hallucinogens (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM) , Stimulants (PMM, PYM, LFM, LNM) , and Sedatives (PMM, PYM, LFM, LNM) ; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.40 OxyContin® Recency-of-Misuse Response Propensity Models

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
OxyContin®	Gender (M); Age Category (T, A1, A2, A3) ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Drug-Specific Lifetime Indicator for Cigars (Y) , Crack (Y) , Hallucinogens (Y) , Heroin (Y) , Inhalants (Y) , Methamphetamine (Y) , Pipes (Y) , Smokeless Tobacco (Y) , and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF) , and Marijuana (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM) , Tranquilizers (PMM, PYM, LFM, LNM) , and Stimulants (PMM, PYM, LFM, LNM) ; Recency State Rank (L, M)	N/A

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

N/A = not applicable.

Table D.41 OxyContin® Recency-of-Misuse Predictive Mean Models

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
OxyContin®		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	N/A

Table D.41 OxyContin® Recency-of-Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	N/A

Table D.41 OxyContin® Recency-of-Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM) ; Recency State Rank (L, M)	N/A
Bin 4	No Model Used; No Nonrespondents	N/A

Table D.41 OxyContin® Recency-of-Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Hallucinogens (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), and Marijuana (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	N/A

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

N/A = not applicable.

Table D.42 Hallucinogen Child Drugs Recency-of-Use Response Propensity Models

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
<i>Salvia divinorum</i>	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), DMT/AMT/Foxy (Y), Ecstasy (Y), Heroin (Y), Other Hallucinogen (Y), Ketamine (Y), LSD (Y), PCP (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Inhalants (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Ecstasy (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and DMT/AMT/Foxy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)

Table D.42 Hallucinogen Child Drugs Recency-of-Use Response Propensity Models (continued)

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
LSD	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), DMT/AMT/Foxy (Y), Ecstasy (Y), Heroin (Y), Other Hallucinogen (Y), Ketamine (Y), PCP (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), Salvia (PM, PY, LF), and Inhalants (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Ecstasy (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and DMT/AMT/Foxy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)

Table D.42 Hallucinogen Child Drugs Recency-of-Use Response Propensity Models (continued)

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Ketamine	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), DMT/AMT/Foxy (Y), Ecstasy (Y), Heroin (Y), Other Hallucinogen (Y), PCP (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), Salvia (PM, PY, LF), and Inhalants (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Ecstasy (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and DMT/AMT/Foxy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)

Table D.42 Hallucinogen Child Drugs Recency-of-Use Response Propensity Models (continued)

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Ecstasy	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), DMT/AMT/Foxy (Y), Heroin (Y), Other Hallucinogen (Y), PCP (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), Salvia (PM, PY, LF), and Inhalants (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and DMT/AMT/Foxy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)

Table D.42 Hallucinogen Child Drugs Recency-of-Use Response Propensity Models (continued)

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
PCP	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), DMT/AMT/Foxy (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), Salvia (PM, PY, LF), and Ecstasy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Ecstasy (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), Salvia (PM, PY, LF), and DMT/AMT/Foxy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)

Table D.42 Hallucinogen Child Drugs Recency-of-Use Response Propensity Models (continued)

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
DMT/AMT/Foxy	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and Ecstasy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and Ecstasy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.43 Hallucinogen Child Drugs Recency-of-Use Predictive Mean Models

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
<i>Salvia divinorum</i>		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 2	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 3	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), DMT/AMT/Foxy (Y), Ecstasy (Y), Heroin (Y), Other Hallucinogen (Y), Ketamine (Y), LSD (Y), PCP (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Inhalants (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Ecstasy (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and DMT/AMT/Foxy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)

Table D.43 Hallucinogen Child Drugs Recency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), DMT/AMT/Foxy (Y), Ecstasy (Y), Heroin (Y), Other Hallucinogen (Y), Ketamine (Y), LSD (Y), PCP (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Inhalants (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Ecstasy (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and DMT/AMT/Foxy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)
LSD		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents

Table D.43 Hallucinogen Child Drugs Recency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), DMT/AMT/Foxy (Y), Ecstasy (Y), Heroin (Y), Other Hallucinogen (Y), Ketamine (Y), PCP (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), Salvia (PM, PY, LF), and Inhalants (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	Gender (M) ; Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Ecstasy (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and DMT/AMT/Foxy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)

Table D.43 Hallucinogen Child Drugs Recency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	No Model Used; No Nonrespondents	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Ecstasy (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and DMT/AMT/Foxy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)

Table D.43 Hallucinogen Child Drugs Recency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), DMT/AMT/Foxy (Y), Ecstasy (Y), Heroin (Y), Other Hallucinogen (Y), Ketamine (Y), PCP (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), Salvia (PM, PY, LF), and Inhalants (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Ecstasy (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and DMT/AMT/Foxy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)

Table D.43 Hallucinogen Child Drugs Recency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S) ; Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), DMT/AMT/Foxy (Y), Ecstasy (Y), Heroin (Y), Other Hallucinogen (Y), Ketamine (Y), PCP (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), Salvia (PM, PY, LF), and Inhalants (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM) ; Lifetime State Rank (L, M)	Gender (M) ; Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S) ; Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Ecstasy (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and DMT/AMT/Foxy (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM) ; Lifetime State Rank (L, M)
Ketamine		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents

Table D.43 Hallucinogen Child Drugs Recency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	No Model Used; No Nonrespondents	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S) ; Orthogonal Age * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Ecstasy (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and DMT/AMT/Foxy (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM) ; Lifetime State Rank (L, M)
Bin 3	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents

Table D.43 Hallucinogen Child Drugs Recency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), DMT/AMT/Foxy (Y), Ecstasy (Y), Heroin (Y), Other Hallucinogen (Y), PCP (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), Salvia (PM, PY, LF), and Inhalants (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age ; Orthogonal Age Squared ; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Ecstasy (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and DMT/AMT/Foxy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)

Table D.43 Hallucinogen Child Drugs Recency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), DMT/AMT/Foxy (Y), Ecstasy (Y), Heroin (Y), Other Hallucinogen (Y), PCP (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), Salvia (PM, PY, LF), and Inhalants (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Ecstasy (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and DMT/AMT/Foxy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)
Ecstasy		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 2	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents

Table D.43 Hallucinogen Child Drugs Recency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), DMT/AMT/Foxy (Y), Heroin (Y), Other Hallucinogen (Y), PCP (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), Salvia (PM, PY, LF), and Inhalants (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM) ; Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared ; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and DMT/AMT/Foxy (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM) ; Lifetime State Rank (L, M)

Table D.43 Hallucinogen Child Drugs Recency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), DMT/AMT/Foxy (Y), Heroin (Y), Other Hallucinogen (Y), PCP (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), Salvia (PM, PY, LF), and Inhalants (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and DMT/AMT/Foxy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)

Table D.43 Hallucinogen Child Drugs Recency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), DMT/AMT/Foxy (Y), Heroin (Y), Other Hallucinogen (Y), PCP (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), Salvia (PM, PY, LF), and Inhalants (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM) ; Lifetime State Rank (L, M)	Gender (M); Orthogonal Age ; Orthogonal Age Squared ; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S) ; Orthogonal Age * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and DMT/AMT/Foxy (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM) ; Lifetime State Rank (L, M)
PCP		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents

Table D.43 Hallucinogen Child Drugs Recency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), DMT/AMT/Foxy (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), Salvia (PM, PY, LF), and Ecstasy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	No Model Used; No Nonrespondents

Table D.43 Hallucinogen Child Drugs Recency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), DMT/AMT/Foxy (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), Salvia (PM, PY, LF), and Ecstasy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Ecstasy (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), Salvia (PM, PY, LF), and DMT/AMT/Foxy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)

Table D.43 Hallucinogen Child Drugs Recency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), DMT/AMT/Foxy (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), Salvia (PM, PY, LF), and Ecstasy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Ecstasy (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), Salvia (PM, PY, LF), and DMT/AMT/Foxy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)

Table D.43 Hallucinogen Child Drugs Recency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), DMT/AMT/Foxy (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), Salvia (PM, PY, LF), and Ecstasy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Ecstasy (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), Salvia (PM, PY, LF), and DMT/AMT/Foxy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)
DMT/AMT/Foxy		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents

Table D.43 Hallucinogen Child Drugs Recency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and Ecstasy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and Ecstasy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)

Table D.43 Hallucinogen Child Drugs Recency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and Ecstasy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and Ecstasy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)

Table D.43 Hallucinogen Child Drugs Recency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and Ecstasy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared ; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and Ecstasy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)

Table D.43 Hallucinogen Child Drugs Recency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and Ecstasy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Other Hallucinogen (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Ketamine (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), Salvia (PM, PY, LF), and Ecstasy (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Lifetime State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.44 12-Month Frequency-of-Use Response Propensity Models

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Methamphetamine	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens, and Inhalants
Hallucinogens	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants, and Methamphetamine

Table D.44 12-Month Frequency-of-Use Response Propensity Models (continued)

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Inhalants	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens, and Methamphetamine	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens, and Methamphetamine

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.45 12-Month Frequency-of-Use Predictive Mean Models

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Methamphetamine		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens, and Inhalants; Recency State Rank (L, M)

Table D.45 12-Month Frequency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens, and Inhalants; Recency State Rank (L, M)

Table D.45 12-Month Frequency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens, and Inhalants; Recency State Rank (L, M)

Table D.45 12-Month Frequency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens, and Inhalants; Recency State Rank (L, M)

Table D.45 12-Month Frequency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens, and Inhalants; Recency State Rank (L, M)

Table D.45 12-Month Frequency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Hallucinogens		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants, and Methamphetamine; Recency State Rank (L, M)

Table D.45 12-Month Frequency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants, and Methamphetamine; Recency State Rank (L, M)

Table D.45 12-Month Frequency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants, and Methamphetamine; Recency State Rank (L, M)

Table D.45 12-Month Frequency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM) ; Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants, and Methamphetamine; Recency State Rank (L, M)

Table D.45 12-Month Frequency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants, and Methamphetamine; Recency State Rank (L, M)

Table D.45 12-Month Frequency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Inhalants		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens, and Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens, and Methamphetamine; Recency State Rank (L, M)

Table D.45 12-Month Frequency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens, and Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens, and Methamphetamine; Recency State Rank (L, M)

Table D.45 12-Month Frequency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens, and Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens, and Methamphetamine; Recency State Rank (L, M)

Table D.45 12-Month Frequency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens, and Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens, and Methamphetamine; Recency State Rank (L, M)

Table D.45 12-Month Frequency-of-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H) ; Pipe Recency (PM, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM) ; Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens, and Methamphetamine ; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H) ; Pipe Recency (PM, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM) ; Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens, and Methamphetamine ; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.46 30-Day Frequency-of-Use/Misuse Response Propensity Models

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Methamphetamine	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Drug-Specific 30-Day Frequency for Hallucinogens, Pain Relievers, Sedatives, Stimulants, Tranquilizers, and Inhalants

Table D.46 30-Day Frequency-of-Use/Misuse Response Propensity Models (continued)

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Hallucinogens	Gender (M); Age Category (T, A1, A2, A3) ; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Methamphetamine	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Inhalants, Pain Relievers, Sedatives, Stimulants, Tranquilizers, and Methamphetamine

Table D.46 30-Day Frequency-of-Use/Misuse Response Propensity Models (continued)

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Inhalants	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Hallucinogens, and Methamphetamine	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Hallucinogens, Pain Relievers, Sedatives, Stimulants, Tranquilizers, and Methamphetamine

Table D.46 30-Day Frequency-of-Use/Misuse Response Propensity Models (continued)

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Stimulants	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, and Inhalants	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Pain Relievers, Sedatives, Tranquilizers, and Inhalants
Sedatives	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), and Tranquilizers (PMM, PYM, LFM, LNM); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Stimulants, and Inhalants	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Pain Relievers, Stimulants, Tranquilizers, and Inhalants

Table D.46 30-Day Frequency-of-Use/Misuse Response Propensity Models (continued)

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Tranquilizers	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Sedatives, Stimulants, and Inhalants	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Pain Relievers, Sedatives, Stimulants, and Inhalants
Pain Relievers	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Sedatives, Stimulants, Tranquilizers, and Inhalants	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Sedatives, Stimulants, Tranquilizers, and Inhalants

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Methamphetamine		
Bin 1	No Model Used; No Nonrespondents	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Drug-Specific 30-Day Frequency for Hallucinogens, Pain Relievers, Sedatives, Stimulants, Tranquilizers, and Inhalants; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Drug-Specific 30-Day Frequency for Hallucinogens, Pain Relievers, Sedatives, Stimulants, Tranquilizers, and Inhalants; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Drug-Specific 30-Day Frequency for Hallucinogens, Pain Relievers, Sedatives, Stimulants, Tranquilizers, and Inhalants; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Drug-Specific 30-Day Frequency for Hallucinogens, Pain Relievers, Sedatives, Stimulants, Tranquilizers, and Inhalants; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Drug-Specific 30-Day Frequency for Hallucinogens, Pain Relievers, Sedatives, Stimulants, Tranquilizers, and Inhalants; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Hallucinogens		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Inhalants, Pain Relievers, Sedatives, Stimulants, Tranquilizers, and Methamphetamine; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Inhalants, Pain Relievers, Sedatives, Stimulants, Tranquilizers, and Methamphetamine; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Inhalants, Pain Relievers, Sedatives, Stimulants, Tranquilizers, and Methamphetamine ; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens ; Drug-Specific 30-Day Frequency for Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Inhalants, Pain Relievers, Sedatives, Stimulants, Tranquilizers, and Methamphetamine ; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Inhalants, Pain Relievers, Sedatives, Stimulants, Tranquilizers, and Methamphetamine; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Inhalants		
Bin 1	No Model Used; No Nonrespondents	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Hallucinogens, Pain Relievers, Sedatives, Stimulants, Tranquilizers, and Methamphetamine; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Hallucinogens, and Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Hallucinogens, Pain Relievers, Sedatives, Stimulants, Tranquilizers, and Methamphetamine; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Hallucinogens, and Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Hallucinogens, Pain Relievers, Sedatives, Stimulants, Tranquilizers, and Methamphetamine; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Hallucinogens, and Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Hallucinogens, Pain Relievers, Sedatives, Stimulants, Tranquilizers, and Methamphetamine; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Hallucinogens, and Methamphetamine; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Hallucinogens, Pain Relievers, Sedatives, Stimulants, Tranquilizers, and Methamphetamine; Recency State Rank (L, M)
Stimulants		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 2	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, and Inhalants; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Pain Relievers, Sedatives, Tranquilizers, and Inhalants; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, and Inhalants; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Pain Relievers, Sedatives, Tranquilizers, and Inhalants; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, and Inhalants; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Pain Relievers, Sedatives, Tranquilizers, and Inhalants; Recency State Rank (L, M)
Sedatives		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 2	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 3	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	No Model Used; No Nonrespondents	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S) ; Orthogonal Age * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Pipe Recency (PM, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Pain Relievers, Stimulants, Tranquilizers, and Inhalants ; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), and Tranquilizers (PMM, PYM, LFM, LNM); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Stimulants, and Inhalants; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Pain Relievers, Stimulants, Tranquilizers, and Inhalants; Recency State Rank (L, M)
Tranquilizers		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Sedatives, Stimulants, and Inhalants; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Pain Relievers, Sedatives, Stimulants, and Inhalants; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Pipe Recency (PM, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF) , Crack (PM, PY, LF) , Heroin (PM, PY, LF) , Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF) , Smokeless Tobacco (PM, PY, P3Y, LF) , and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Sedatives, Stimulants, and Inhalants; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF) , Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Pain Relievers, Sedatives, Stimulants, and Inhalants; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Sedatives, Stimulants, and Inhalants; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Pain Relievers, Sedatives, Stimulants, and Inhalants; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Sedatives, Stimulants, and Inhalants; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Pain Relievers, Sedatives, Stimulants, and Inhalants; Recency State Rank (L, M)
Pain Relievers		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Sedatives, Stimulants, Tranquilizers, and Inhalants; Recency State Rank (L, M)	No Model Used; No Nonrespondents
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Sedatives, Stimulants, Tranquilizers, and Inhalants; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Sedatives, Stimulants, Tranquilizers, and Inhalants; Recency State Rank (L, M)

Table D.47 30-Day Frequency-of-Use/Misuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Sedatives, Stimulants, Tranquilizers, and Inhalants; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Sedatives, Stimulants, Tranquilizers, and Inhalants; Recency State Rank (L, M)
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Sedatives, Stimulants, Tranquilizers, and Inhalants; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific 30-Day Frequency for Hallucinogens, Methamphetamine, Sedatives, Stimulants, Tranquilizers, and Inhalants; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.48 Age-at-First-Use Response Propensity Models

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Methamphetamine	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Methamphetamine (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Methamphetamine (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM)
Hallucinogens	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM)

Table D.48 Age-at-First-Use Response Propensity Models (continued)

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Inhalants	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Inhalants (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Inhalants (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.49 Age-at-First-Use Predictive Mean Models

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Methamphetamine		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Methamphetamine (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Drug-Specific 30-Day Frequency for Methamphetamine; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Heroin, Marijuana, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Methamphetamine (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Drug-Specific 30-Day Frequency for Methamphetamine; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Inhalants, Marijuana, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

Table D.49 Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Methamphetamine (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Drug-Specific 30-Day Frequency for Methamphetamine; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Heroin, Marijuana, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Methamphetamine (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Drug-Specific 30-Day Frequency for Methamphetamine; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Inhalants, Marijuana, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

Table D.49 Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Methamphetamine (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Drug-Specific 30-Day Frequency for Methamphetamine; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Heroin, Marijuana, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Methamphetamine (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Drug-Specific 30-Day Frequency for Methamphetamine; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Inhalants, Marijuana, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

Table D.49 Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Methamphetamine (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Drug-Specific 30-Day Frequency for Methamphetamine; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Heroin, Marijuana, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Methamphetamine (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Drug-Specific 30-Day Frequency for Methamphetamine; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Inhalants, Marijuana, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

Table D.49 Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Methamphetamine (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Drug-Specific 30-Day Frequency for Methamphetamine; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Heroin, Marijuana, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Methamphetamine (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Methamphetamine; Drug-Specific 30-Day Frequency for Methamphetamine; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Inhalants, Marijuana, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

Table D.49 Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Hallucinogens		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Hallucinogens; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Heroin, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Hallucinogens; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Heroin, Inhalants, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

Table D.49 Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Hallucinogens; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Heroin, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Hallucinogens; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Heroin, Inhalants, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

Table D.49 Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Hallucinogens; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Heroin, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Hallucinogens; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Heroin, Inhalants, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

Table D.49 Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Hallucinogens; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Heroin, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Hallucinogens; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Heroin, Inhalants, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

Table D.49 Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Hallucinogens; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Heroin, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Hallucinogens; Drug-Specific 30-Day Frequency for Hallucinogens; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Heroin, Inhalants, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

Table D.49 Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Inhalants		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Inhalants (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Inhalants; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Inhalants (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Inhalants; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

Table D.49 Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Inhalants (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Inhalants; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Inhalants (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Inhalants; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

Table D.49 Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Inhalants (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Inhalants; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Inhalants (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Inhalants; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

Table D.49 Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Inhalants (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Inhalants; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Inhalants (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Inhalants; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

Table D.49 Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Inhalants (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Inhalants; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Hallucinogens (PM, PY, LF), Heroin (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Inhalants (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific 12-Month Frequency, Domain Restricted to Lifetime Users for Inhalants; Drug-Specific 30-Day Frequency for Inhalants; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Recency State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.50 Hallucinogen Child Drugs Age-at-First-Use Response Propensity Models

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Ecstasy	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), PCP (PM, PY, LF), and LSD (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Ecstasy (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), PCP (PM, PY, LF), and LSD (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Ecstasy (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM)
LSD	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), PCP (PM, PY, LF), and Ecstasy (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for LSD (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Marijuana (PM, PY, LF), PCP (PM, PY, LF), and Ecstasy (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for LSD (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM)

Table D.50 Hallucinogen Child Drugs Age-at-First-Use Response Propensity Models (continued)

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
PCP	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), and Ecstasy (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for PCP (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Inhalants (Y), Methamphetamine (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), and Ecstasy (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for PCP (PM, PY) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.51 Hallucinogen Child Drugs Age-at-First-Use Predictive Mean Models

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Ecstasy		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Ecstasy (PM, PY), and Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Inhalants, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Ecstasy (PM, PY), and Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Inhalants, LSD, Marijuana, Methamphetamine, PCP, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)

Table D.51 Hallucinogen Child Drugs Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Ecstasy (PM, PY), and Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Inhalants, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Ecstasy (PM, PY), and Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Inhalants, LSD, Marijuana, Methamphetamine, PCP, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)

Table D.51 Hallucinogen Child Drugs Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Ecstasy (PM, PY), and Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Inhalants, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Ecstasy (PM, PY), and Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Inhalants, LSD, Marijuana, Methamphetamine, PCP, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)

Table D.51 Hallucinogen Child Drugs Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Ecstasy (PM, PY), and Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Inhalants, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Ecstasy (PM, PY), and Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Inhalants, LSD, Marijuana, Methamphetamine, PCP, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)

Table D.51 Hallucinogen Child Drugs Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Ecstasy (PM, PY), and Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Inhalants, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Ecstasy (PM, PY), and Hallucinogens (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Hallucinogens, Heroin, Inhalants, LSD, Marijuana, Methamphetamine, PCP, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)

Table D.51 Hallucinogen Child Drugs Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
LSD		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Ecstasy (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY), and LSD (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Ecstasy, Hallucinogens, Heroin, Inhalants, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Ecstasy (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY), and LSD (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Ecstasy, Hallucinogens, Heroin, Inhalants, Marijuana, Methamphetamine, PCP, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)

Table D.51 Hallucinogen Child Drugs Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Ecstasy (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY), and LSD (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Ecstasy, Hallucinogens, Heroin, Inhalants, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Ecstasy (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY), and LSD (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Ecstasy, Hallucinogens, Heroin, Inhalants, Marijuana, Methamphetamine, PCP, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)

Table D.51 Hallucinogen Child Drugs Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Ecstasy (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY), and LSD (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Ecstasy, Hallucinogens, Heroin, Inhalants, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Ecstasy (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY), and LSD (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Ecstasy, Hallucinogens, Heroin, Inhalants, Marijuana, Methamphetamine, PCP, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)

Table D.51 Hallucinogen Child Drugs Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Ecstasy (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY), and LSD (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Ecstasy, Hallucinogens, Heroin, Inhalants, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Ecstasy (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY), and LSD (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Ecstasy, Hallucinogens, Heroin, Inhalants, Marijuana, Methamphetamine, PCP, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)

Table D.51 Hallucinogen Child Drugs Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Ecstasy (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and Cocaine (PM, PY, LF) ; Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY), and LSD (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Ecstasy, Hallucinogens, Heroin, Inhalants, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Ecstasy (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), PCP (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY), and LSD (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM) ; Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Ecstasy, Hallucinogens, Heroin, Inhalants, Marijuana, Methamphetamine, PCP, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)
PCP		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents

Table D.51 Hallucinogen Child Drugs Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Ecstasy (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY), and PCP (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Ecstasy, Hallucinogens, Heroin, Inhalants, LSD, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Ecstasy (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY), and PCP (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Ecstasy, Hallucinogens, Heroin, Inhalants, LSD, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)

Table D.51 Hallucinogen Child Drugs Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Ecstasy (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY), and PCP (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Ecstasy, Hallucinogens, Heroin, Inhalants, LSD, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Ecstasy (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY), and PCP (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Ecstasy, Hallucinogens, Heroin, Inhalants, LSD, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)

Table D.51 Hallucinogen Child Drugs Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Ecstasy (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY), and PCP (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Ecstasy, Hallucinogens, Heroin, Inhalants, LSD, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Ecstasy (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY), and PCP (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Ecstasy, Hallucinogens, Heroin, Inhalants, LSD, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)

Table D.51 Hallucinogen Child Drugs Age-at-First-Use Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Ecstasy (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY), and PCP (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Ecstasy, Hallucinogens, Heroin, Inhalants, LSD, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Orthogonal Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Orthogonal Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Daily Cigarette Use (Y); Pipe Recency (PM, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Crack (PM, PY, LF), Ecstasy (PM, PY, LF), Heroin (PM, PY, LF), Inhalants (PM, PY, LF), LSD (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Cocaine (PM, PY, LF); Drug-Specific Recency, Domain Restricted to Lifetime Users (3 Levels) for Hallucinogens (PM, PY), and PCP (PM, PY); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigars (PM, PY, P3Y, LF), Smokeless Tobacco (PM, PY, P3Y, LF), and Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Age at First Use for Alcohol, Cigars, Cigarettes, Cocaine, Crack, Ecstasy, Hallucinogens, Heroin, Inhalants, LSD, Marijuana, Methamphetamine, Smokeless Tobacco, and Daily Cigarette Use; Lifetime State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.52 Past Year Initiation Response Propensity Models

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Sedatives	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Initiation State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Pain Relievers (Y), Tranquilizers (Y), and Stimulants (Y); Initiation State Rank (L, M)

Table D.52 Past Year Initiation Response Propensity Models (continued)

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Tranquilizers	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Sedatives (Y); Initiation State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Pain Relievers (Y), Stimulants (Y), and Sedatives (Y); Initiation State Rank (L, M)

Table D.52 Past Year Initiation Response Propensity Models (continued)

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Stimulants	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Sedatives (Y), and Tranquilizers (Y); Initiation State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Pain Relievers (Y), Tranquilizers (Y), and Sedatives (Y); Initiation State Rank (L, M)

Table D.52 Past Year Initiation Response Propensity Models (continued)

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Pain Relievers	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Sedatives (Y), Tranquilizers (Y), and Stimulants (Y); Initiation State Rank (L, M)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Sedatives (Y), Tranquilizers (Y), and Stimulants (Y); Initiation State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.53 Past Year Initiation Predictive Mean Models

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Sedatives		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 2	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 3	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 4	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S) ; Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF) ; Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF) ; Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM) ; Initiation State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Pain Relievers (Y), Tranquilizers (Y), and Stimulants (Y); Initiation State Rank (L, M)
Tranquilizers		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents

Table D.53 Past Year Initiation Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Sedatives (Y); Initiation State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Pain Relievers (Y), Stimulants (Y), and Sedatives (Y); Initiation State Rank (L, M)

Table D.53 Past Year Initiation Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Sedatives (Y); Initiation State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Pain Relievers (Y), Stimulants (Y), and Sedatives (Y); Initiation State Rank (L, M)

Table D.53 Past Year Initiation Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Sedatives (Y); Initiation State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Pain Relievers (Y), Stimulants (Y), and Sedatives (Y); Initiation State Rank (L, M)

Table D.53 Past Year Initiation Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM) ; Drug-Specific Past Year Initiation for Sedatives (Y); Initiation State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S) ; Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Stimulants (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM) ; Drug-Specific Past Year Initiation for Pain Relievers (Y), Stimulants (Y), and Sedatives (Y); Initiation State Rank (L, M)

Table D.53 Past Year Initiation Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Stimulants		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Sedatives (Y), and Tranquilizers (Y); Initiation State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Pain Relievers (Y), Tranquilizers (Y), and Sedatives (Y); Initiation State Rank (L, M)

Table D.53 Past Year Initiation Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Sedatives (Y), and Tranquilizers (Y); Initiation State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Pain Relievers (Y), Tranquilizers (Y), and Sedatives (Y); Initiation State Rank (L, M)

Table D.53 Past Year Initiation Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	No Model Used; No Nonrespondents	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Pain Relievers (Y), Tranquilizers (Y), and Sedatives (Y); Initiation State Rank (L, M)

Table D.53 Past Year Initiation Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Sedatives (Y), and Tranquilizers (Y); Initiation State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM) ; Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Pain Relievers (Y), Tranquilizers (Y), and Sedatives (Y); Initiation State Rank (L, M)

Table D.53 Past Year Initiation Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Sedatives (Y), and Tranquilizers (Y); Initiation State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed ; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Pain Relievers (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Sedatives (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Pain Relievers (Y), Tranquilizers (Y), and Sedatives (Y); Initiation State Rank (L, M)
Pain Relievers		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents

Table D.53 Past Year Initiation Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Sedatives (Y), Tranquilizers (Y), and Stimulants (Y); Initiation State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Sedatives (Y), Tranquilizers (Y), and Stimulants (Y); Initiation State Rank (L, M)

Table D.53 Past Year Initiation Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Sedatives (Y), Tranquilizers (Y), and Stimulants (Y); Initiation State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Sedatives (Y), Tranquilizers (Y), and Stimulants (Y); Initiation State Rank (L, M)

Table D.53 Past Year Initiation Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Sedatives (Y), Tranquilizers (Y), and Stimulants (Y); Initiation State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Sedatives (Y), Tranquilizers (Y), and Stimulants (Y); Initiation State Rank (L, M)

Table D.53 Past Year Initiation Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Sedatives (Y), Tranquilizers (Y), and Stimulants (Y); Initiation State Rank (L, M)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Orthogonal Age * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Orthogonal Age * Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Lifetime Indicator for Cigars (Y), Crack (Y), Heroin (Y), Pipes (Y), Smokeless Tobacco (Y), and Cocaine (Y); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (3 Levels) for Alcohol (PM, PY, LF), Inhalants (PM, PY, LF), Marijuana (PM, PY, LF), Methamphetamine (PM, PY, LF), and Hallucinogens (PM, PY, LF); Drug-Specific Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Cigarettes (PM, PY, P3Y, LF); Drug-Specific Misuse Recency, Domain Not Restricted to Lifetime Users (4 Levels) for Sedatives (PMM, PYM, LFM, LNM), Tranquilizers (PMM, PYM, LFM, LNM), and Stimulants (PMM, PYM, LFM, LNM); Drug-Specific Past Year Initiation for Sedatives (Y), Tranquilizers (Y), and Stimulants (Y); Initiation State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.54 Hallucinogens Dependence and Abuse Response Propensity Models

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Abuse	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Hallucinogens (Y)
Dependence	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Hallucinogens (Y)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Hallucinogens (Y)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.55 Hallucinogens Dependence and Abuse Predictive Mean Models

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Abuse		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Hallucinogens (Y)
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Hallucinogens (Y)
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Hallucinogens (Y)
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Hallucinogens (Y)

Table D.55 Hallucinogens Dependence and Abuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Dependence		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Hallucinogens (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Hallucinogens (Y)
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Hallucinogens (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Hallucinogens (Y)
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Hallucinogens (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Hallucinogens (Y)
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Hallucinogens (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Hallucinogens (Y)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.56 Inhalants Dependence and Abuse Response Propensity Models

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Abuse	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Inhalants (Y)
Dependence	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Inhalants (Y)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Inhalants (Y)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.57 Inhalants Dependence and Abuse Predictive Mean Models

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Abuse		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Inhalants (Y)
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Inhalants (Y)
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Inhalants (Y)
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Inhalants (Y)
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Inhalants (Y)

Table D.57 Inhalants Dependence and Abuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Dependence		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Inhalants (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Inhalants (Y)
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Inhalants (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Inhalants (Y)
Bin 3	No Model Used; Too Few Positive Respondents	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Inhalants (Y)
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Inhalants (Y)	No Model Used; Too Few Positive Respondents

Table D.57 Inhalants Dependence and Abuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Inhalants (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Inhalants (Y)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.58 Stimulants and Methamphetamine Dependence and Abuse Response Propensity Models

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Stimulant Abuse	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Methamphetamine (Y), and Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y)
Stimulant Dependence	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Stimulants (Y)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Methamphetamine (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y), and Stimulants (Y)
Methamphetamine Abuse	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Stimulants (Y)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Methamphetamine (Y), and Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Stimulants (Y)
Methamphetamine Dependence	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y), and Stimulants (Y)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y), and Stimulants (Y)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.59 Stimulants and Methamphetamine Dependence and Abuse Predictive Mean Models

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Stimulant Abuse		
Bin 1	No Model Used; No Nonrespondents	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Methamphetamine (Y), and Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y)
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Methamphetamine (Y), and Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y)
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Methamphetamine (Y), and Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y)
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Methamphetamine (Y), and Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y)

Table D.59 Stimulants and Methamphetamine Dependence and Abuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Methamphetamine (Y), and Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y)
Stimulant Dependence		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Stimulants (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Methamphetamine (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y), and Stimulants (Y)
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Stimulants (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Methamphetamine (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y), and Stimulants (Y)

Table D.59 Stimulants and Methamphetamine Dependence and Abuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Stimulants (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Methamphetamine (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y), and Stimulants (Y)
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Stimulants (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Methamphetamine (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y), and Stimulants (Y)
Methamphetamine Abuse		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Stimulants (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Methamphetamine (Y), and Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Stimulants (Y)
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Stimulants (Y)	No Model Used; No Nonrespondents

Table D.59 Stimulants and Methamphetamine Dependence and Abuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Stimulants (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Methamphetamine (Y), and Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Stimulants (Y)
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Stimulants (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Methamphetamine (Y), and Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Stimulants (Y)
Methamphetamine Dependence		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y), and Stimulants (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y), and Stimulants (Y)
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y), and Stimulants (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y), and Stimulants (Y)

Table D.59 Stimulants and Methamphetamine Dependence and Abuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y), and Stimulants (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y), and Stimulants (Y)
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y), and Stimulants (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Stimulants (Y); Drug-Specific Past Year Abuse Indicator for Methamphetamine (Y), and Stimulants (Y)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.60 Pain Relievers Dependence and Abuse Response Propensity Models

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Abuse	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Pain Relievers (Y)
Dependence	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Pain Relievers (Y)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Pain Relievers (Y)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.61 Pain Relievers Dependence and Abuse Predictive Mean Models

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Abuse		
Bin 1	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Pain Relievers (Y)
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Pain Relievers (Y)
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Pain Relievers (Y)
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Pain Relievers (Y)
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Pain Relievers (Y)

Table D.61 Pain Relievers Dependence and Abuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Dependence		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Pain Relievers (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Pain Relievers (Y)
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Pain Relievers (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Pain Relievers (Y)
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Pain Relievers (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Pain Relievers (Y)
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Pain Relievers (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Pain Relievers (Y)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.62 Sedatives and Tranquilizers Dependence and Abuse Response Propensity Models

Imputation Step	Variables Included in Response Propensity Model	
	Cycle 1	Cycle 2
Sedative Abuse	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y), and Tranquilizers (Y); Drug-Specific Past Year Abuse Indicator for Tranquilizers (Y)
Sedative Dependence	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Tranquilizers (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y), and Tranquilizers (Y)
Tranquilizer Abuse	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y), and Tranquilizers (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y)
Tranquilizer Dependence	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y), and Tranquilizers (Y)	Gender (M); Age Category (T, A1, A2, A3); Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y), and Tranquilizers (Y)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.63 Sedatives and Tranquilizers Dependence and Abuse Predictive Mean Models

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Sedative Abuse		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 2	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 3	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	No Model Used; No Nonrespondents
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y), and Tranquilizers (Y); Drug-Specific Past Year Abuse Indicator for Tranquilizers (Y)
Sedative Dependence		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 2	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 3	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Bin 4	No Model Used; No Nonrespondents	No Model Used; Too Few Positive Respondents
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Tranquilizers (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y), and Tranquilizers (Y)
Tranquilizer Abuse		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents

Table D.63 Sedatives and Tranquilizers Dependence and Abuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y), and Tranquilizers (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y)
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y), and Tranquilizers (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y)
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y), and Tranquilizers (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y)
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y), and Tranquilizers (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y)
Tranquilizer Dependence		
Bin 1	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents

Table D.63 Sedatives and Tranquilizers Dependence and Abuse Predictive Mean Models (continued)

Imputation Step	Variables Included in Predictive Mean Model	
	Cycle 1	Cycle 2
Bin 2	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y), and Tranquilizers (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y), and Tranquilizers (Y)
Bin 3	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y), and Tranquilizers (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y), and Tranquilizers (Y)
Bin 4	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y), and Tranquilizers (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y), and Tranquilizers (Y)
Bin 5	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y), and Tranquilizers (Y)	Gender (M); Orthogonal Age; Orthogonal Age Squared; Orthogonal Age Cubed; Race/Hispanic Recode (4 Levels) (B, O, H); Drug-Specific Past Year Indicator for Alcohol (Y), Cigarettes (Y), Cocaine (Y), Crack (Y), Heroin (Y), Marijuana (Y), Smokeless Tobacco (Y), and Cigars (Y); Drug-Specific Past Year Dependence Indicator for Sedatives (Y); Drug-Specific Past Year Abuse Indicator for Sedatives (Y), and Tranquilizers (Y)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.2](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

D.5 Household Composition Variables

Exhibit D.3 Definitions of Levels for Categorical Variables

Gender

M: Male, F: Female¹

Race/Hispanic Recode (4 Levels)

W: Not Hispanic/Latino and White Only,¹ B: Not Hispanic/Latino and Black/African American Only, O: Not Hispanic/Latino and Other, H: Hispanic/Latino

Education Level (4 Levels)

L: Less than High School, HS: High School Graduate, SC: Some College, C: College Graduate¹

Employment Status

FT: Full Time, PT: Part Time, UN: Unemployed, OE: Other Employment¹

Marital Status (3 Levels)

M: Married, WM: Was Married, NM: Never Been Married¹

Marital Status (4 Levels)

M: Married, W: Widowed, DS: Divorced/Separated, NM: Never Been Married¹

Population Density (3 Levels)

R: Segment Not in CBSA, SM: Segment in CBSA with Fewer than 1 Million People, L: Segment in CBSA with 1 Million or More People¹

Census Region

N: Northeast, M: Midwest, S: South, W: West¹

Percentage Hispanic/Latino in Block

H: $\geq 71\%$, M: $[20,71)\%$, L: $< 20\%$ ¹

Percentage Owner Occupied in Block

H: $\geq 50\%$, M: $[10,50)\%$, L: $< 10\%$ ¹

Gender * Race/Hispanic Recode (4 Levels)

MW: Male Not Hispanic/Latino and White Only,¹ MB: Male Not Hispanic/Latino and Black/African American Only, MO: Male Not Hispanic/Latino and Other, MH: Male Hispanic/Latino, FW: Female Not Hispanic/Latino and White Only,¹ FB: Female Not Hispanic/Latino and Black/African American Only,¹ FO: Female Not Hispanic/Latino and Other,¹ FH: Female Hispanic/Latino¹

Age * Race/Hispanic Recode (4 Levels)

W: Age * Not Hispanic/Latino and White Only,¹ B: Age * Not Hispanic/Latino and Black/African American Only, O: Age * Not Hispanic/Latino and Other, H: Age * Hispanic/Latino

Age Squared * Race/Hispanic Recode (4 Levels)

W: Age Squared * Not Hispanic/Latino and White Only,¹ B: Age Squared * Not Hispanic/Latino and Black/African American Only, O: Age Squared * Not Hispanic/Latino and Other, H: Age Squared * Hispanic/Latino

Other Family Members in Household

N: No, Y: Yes¹

NOTE: An asterisk "*" represents an interaction between two variables.

CBSA = core-based statistical area.

¹ This is the reference level for this variable, against which effects of other factor levels are measured.

Table D.64 Household Composition: People 12 to 17 Years Old

Variable Requiring Imputation	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Household Size (TOTPEOP)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Total People in Household (Screener)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Total People in Household (Screener)
Number of People Younger than 18 in Household (KID17)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Household Size; Number of Eligible 12 to 17 in Household (Screener)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Household Size; Number of Eligible 12 to 17 in Household (Screener)
Number of People Older than 64 in Household (HH65)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Number of People Younger than 18 Years Old in Household; Household Size	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Number of People Younger than 18 Years Old in Household; Household Size

Table D.64 Household Composition: People 12 to 17 Years Old (continued)

Variable Requiring Imputation	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Other Family Present in Household (FAMSKIP)	Gender (M); Age; Age Squared ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S) ; Percentage Hispanic/Latino in Block (H, M) ; Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H) ; Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household ; Household Size	Gender (M); Age; Age Squared ; Race/Hispanic Recode (4 Levels) (B, O, H) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S) ; Percentage Hispanic/Latino in Block (H, M) ; Percentage Owner Occupied in Block (H, M) ; Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H) ; Number of People Younger than 18 Years Old in Household ; Number of People Older than 64 Years Old in Household; Household Size
Number of Respondent's Family Members in Household Including Foster Relationships (FMLYSIZE)	Gender (M); Age; Age Squared ; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM) ; Census Region (N, M, S) ; Percentage Hispanic/Latino in Block (H, M) ; Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H) ; Other Family Members in Household (N) ; Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household; Household Size	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Other Family Members in Household (N); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household; Household Size
Number of Respondent's Family Members in Household Younger than 18 Including Foster Relationships (KIDFMLY)	Gender (M); Age; Age Squared ; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M) ; Percentage Owner Occupied in Block (H, M) ; Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H) ; Other Family Members in Household (N) ; Number of People Younger than 18 Years Old in Household ; Number of People Older than 64 Years Old in Household ; Household Size ; Number of Respondent's Family Members in Household (Includes Foster Relationships)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Other Family Members in Household (N); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household; Household Size; Number of Respondent's Family Members in Household (Includes Foster Relationships)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.3](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.65 Household Composition: People 18 to 25 Years Old

Variable Requiring Imputation	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Household Size (TOTPEOP)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Total People in Household (Screener)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Total People in Household (Screener)
Number of People Younger than 18 in Household (KID17)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Household Size; Number of Eligible 12 to 17 in Household (Screener)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Household Size; Number of Eligible 12 to 17 in Household (Screener)
Number of People Older than 64 in Household (HH65)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Number of People Younger than 18 Years Old in Household; Household Size	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Number of People Younger than 18 Years Old in Household; Household Size

Table D.65 Household Composition: People 18 to 25 Years Old (continued)

Variable Requiring Imputation	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Other Family Present in Household (FAMSKIP)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household; Household Size	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household; Household Size
Number of Respondent's Family Members in Household Including Foster Relationships (FMLYSIZE)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Other Family Members in Household (N); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household; Household Size	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Other Family Members in Household (N); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household; Household Size

Table D.65 Household Composition: People 18 to 25 Years Old (continued)

Variable Requiring Imputation	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Number of Respondent's Family Members in Household Younger than 18 Including Foster Relationships (KIDFMLY)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Other Family Members in Household (N); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household; Household Size; Number of Respondent's Family Members in Household (Includes Foster Relationships)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Other Family Members in Household (N); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household; Household Size; Number of Respondent's Family Members in Household (Includes Foster Relationships)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.3](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.66 Household Composition: People 26 to 64 Years Old

Variable Requiring Imputation	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Household Size (TOTPEOP)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Total People in Household (Screener)	Gender (M); Age; Age Squared ; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H) ; Total People in Household (Screener)
Number of People Younger than 18 in Household (KID17)	Gender (M); Age; Age Squared ; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H) ; Household Size; Number of Eligible 12 to 17 in Household (Screener)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Household Size; Number of Eligible 12 to 17 in Household (Screener)
Number of People Older than 64 in Household (HH65)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H) ; Number of People Younger than 18 Years Old in Household; Household Size	Gender (M); Age; Age Squared ; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H) ; Number of People Younger than 18 Years Old in Household; Household Size

Table D.66 Household Composition: People 26 to 64 Years Old (continued)

Variable Requiring Imputation	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Other Family Present in Household (FAMSKIP)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household; Household Size	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M) ; Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household; Household Size
Number of Respondent's Family Members in Household Including Foster Relationships (FMLYSIZE)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Other Family Members in Household (N); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household; Household Size	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Other Family Members in Household (N); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household; Household Size

Table D.66 Household Composition: People 26 to 64 Years Old (continued)

Variable Requiring Imputation	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Number of Respondent's Family Members in Household Younger than 18 Including Foster Relationships (KIDFMLY)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Other Family Members in Household (N); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household; Household Size; Number of Respondent's Family Members in Household (Includes Foster Relationships)	Gender (M); Age; Age Squared ; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S) ; Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M) ; Age * Gender; Age Squared * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H) ; Other Family Members in Household (N); Number of People Younger than 18 Years Old in Household ; Number of People Older than 64 Years Old in Household; Household Size; Number of Respondent's Family Members in Household (Includes Foster Relationships)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.3](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.67 Household Composition: People 65 Years or Older

Variable Requiring Imputation	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Household Size (TOTPEOP)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Total People in Household (Screener)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Total People in Household (Screener)
Number of People Younger than 18 in Household (KID17)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Household Size; Number of Eligible 12 to 17 in Household (Screener)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Household Size; Number of Eligible 12 to 17 in Household (Screener)
Number of People Older than 64 in Household (HH65)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Number of People Younger than 18 Years Old in Household; Household Size	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Number of People Younger than 18 Years Old in Household; Household Size

Table D.67 Household Composition: People 65 Years or Older (continued)

Variable Requiring Imputation	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Other Family Present in Household (FAMSKIP)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S) ; Percentage Hispanic/Latino in Block (H, M) ; Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H) ; Number of People Younger than 18 Years Old in Household ; Number of People Older than 64 Years Old in Household ; Household Size	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S) ; Percentage Hispanic/Latino in Block (H, M) ; Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H) ; Number of People Younger than 18 Years Old in Household ; Number of People Older than 64 Years Old in Household ; Household Size
Number of Respondent's Family Members in Household Including Foster Relationships (FMLYSIZE)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S) ; Percentage Hispanic/Latino in Block (H, M) ; Percentage Owner Occupied in Block (H, M) ; Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H) ; Other Family Members in Household (N) ; Number of People Younger than 18 Years Old in Household ; Number of People Older than 64 Years Old in Household ; Household Size	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S) ; Percentage Hispanic/Latino in Block (H, M) ; Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H) ; Other Family Members in Household (N) ; Number of People Younger than 18 Years Old in Household ; Number of People Older than 64 Years Old in Household ; Household Size

Table D.67 Household Composition: People 65 Years or Older (continued)

Variable Requiring Imputation	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model
Number of Respondent's Family Members in Household Younger than 18 Including Foster Relationships (KIDFMLY)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH) ; Age * Race/Hispanic Recode (4 Levels) (B, O, H) ; Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H) ; Other Family Members in Household (N); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household; Household Size; Number of Respondent's Family Members in Household (Includes Foster Relationships)	Gender (M); Age; Age Squared; Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H) ; Other Family Members in Household (N) ; Number of People Younger than 18 Years Old in Household ; Number of People Older than 64 Years Old in Household; Household Size ; Number of Respondent's Family Members in Household (Includes Foster Relationships)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.3](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

D.6 Income Variables

Exhibit D.4 Definitions of Levels for Categorical Variables

Gender

M: Male, F: Female¹

Race/Hispanic Recode (4 Levels)

W: Not Hispanic/Latino and White Only, B: Not Hispanic/Latino and Black/African American Only, O: Not Hispanic/Latino and Other,¹ H: Hispanic/Latino

Education Level (4 Levels)

L: Less than High School, HS: High School Graduate, SC: Some College, C: College Graduate¹

Employment Status

FT: Full Time, PT: Part Time, UN: Unemployed, OE: Other Employment¹

Marital Status (3 Levels)

M: Married, WM: Was Married, NM: Never Been Married¹

Marital Status (4 Levels)

M: Married, W: Widowed, DS: Divorced/Separated, NM: Never Been Married¹

Population Density (3 Levels)

R: Segment Not in CBSA, SM: Segment in CBSA with Fewer than 1 Million People, L: Segment in CBSA with 1 Million or More People¹

Census Region

N: Northeast, M: Midwest, S: South, W: West¹

Percentage Hispanic/Latino in Block

H: $\geq 71\%$, M: $[20,71)\%$, L: $< 20\%$ ¹

Percentage Non-Hispanic/Latino Black/African American in Block

H: $\geq 50\%$, M: $[10,50)\%$, L: $< 10\%$ ¹

Percentage Owner Occupied in Block

H: $\geq 50\%$, M: $[10,50)\%$, L: $< 10\%$ ¹

Gender * Race/Hispanic Recode (4 Levels)

MW: Male Not Hispanic/Latino and White Only, MB: Male Not Hispanic/Latino and Black/African American Only, MO: Male Not Hispanic/Latino and Other,¹ MH: Male Hispanic/Latino, FW: Female Not Hispanic/Latino and White Only,¹ FB: Female Not Hispanic/Latino and Black/African American Only,¹ FO: Female Not Hispanic/Latino and Other,¹ FH: Female Hispanic/Latino¹

Age * Race/Hispanic Recode (4 Levels)

W: Age * Not Hispanic/Latino and White Only, B: Age * Not Hispanic/Latino and Black/African American Only, O: Age * Not Hispanic/Latino and Other,¹ H: Age * Hispanic/Latino

Age Squared * Race/Hispanic Recode (4 Levels)

W: Age Squared * Not Hispanic/Latino and White Only, B: Age Squared * Not Hispanic/Latino and Black/African American Only, O: Age Squared * Not Hispanic/Latino and Other,¹ H: Age Squared * Hispanic/Latino

Age Cubed * Race/Hispanic Recode (4 Levels)

W: Age Cubed * Not Hispanic/Latino and White Only, B: Age Cubed * Not Hispanic/Latino and Black/African American Only, O: Age Cubed * Not Hispanic/Latino and Other,¹ H: Age Cubed * Hispanic/Latino

Intermediate Family Received Food Stamps

Y: Yes, N: No¹

Family Received Food Stamps

N: No, Y: Yes¹

Intermediate Family Received Public Assistance

Y: Yes, N: No¹

Family Received Public Assistance

N: No, Y: Yes¹

Intermediate Family Received Social Security or Railroad Payments

Y: Yes, N: No¹

Family Received Social Security or Railroad Payments

N: No, Y: Yes¹

Intermediate Family Received Supplemental Security Income

Y: Yes, N: No¹

Exhibit D.4 Definitions of Levels for Categorical Variables (continued)

Family Received Supplemental Security Income

N: No, Y: Yes¹

Intermediate Family Received Welfare/Job Placement/Childcare Services

Y: Yes, N: No¹

Family Received Welfare/Job Placement/Childcare Services

N: No, Y: Yes¹

Total Family Income > or < \$20,000

LT: < \$20,000, GT: ≥ \$20,000¹

Income State Rank

L: Low State Income Rank (low tertile), M: Middle State Income Rank (middle tertile), H: High State Income Rank (high tertile)¹

NOTE: An asterisk "*" represents an interaction between two variables.

CBSA = core-based statistical area.

¹ This is the reference level for this variable, against which effects of other factor levels are measured.

Table D.68 Dichotomous Income Indicator Response Propensity Models

Age Group	Variables Included in Response Propensity Model (Dichotomous Income Indicators)
12-17	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Age Cubed * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Age Cubed * Race/Hispanic Recode (4 Levels) (W, B, H); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
18-25	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S) ; Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Age Cubed * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH) ; Age * Race/Hispanic Recode (4 Levels) (W, B, H) ; Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H) ; Age Cubed * Race/Hispanic Recode (4 Levels) (W, B, H) ; Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous) ; Number of Adults in Household; Income State Rank (L, M)
26-64	Age; Age Squared; Age Cubed ; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Age Cubed * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H) ; Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H) ; Age Cubed * Race/Hispanic Recode (4 Levels) (W, B, H) ; Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
65+	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H) ; Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S) ; Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Age Cubed * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH) ; Age * Race/Hispanic Recode (4 Levels) (W, B, H) ; Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H) ; Age Cubed * Race/Hispanic Recode (4 Levels) (W, B, H) ; Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous) ; Number of Adults in Household; Income State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.4](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.69 Dichotomous Income Indicator Predictive Mean Models: People 12 to 17 Years Old

Variable Requiring Imputation	Variables Included in Predictive Mean Model (Dichotomous Income Indicators)
Social Security	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
Supplemental Security	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Social Security or Railroad Payments (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
Food Stamps	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Public Assistance (Y); Intermediate Family Received Social Security or Railroad Payments (Y); Intermediate Family Received Supplemental Security Income (Y); Intermediate Family Received Welfare/Job Placement/Childcare Services (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
Welfare Payments	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Social Security or Railroad Payments (Y); Intermediate Family Received Supplemental Security Income (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)

Table D.69 Dichotomous Income Indicator Predictive Mean Models: People 12 to 17 Years Old (continued)

Variable Requiring Imputation	Variables Included in Predictive Mean Model (Dichotomous Income Indicators)
Welfare Services	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Public Assistance (Y); Intermediate Family Received Social Security or Railroad Payments (Y); Intermediate Family Received Supplemental Security Income (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
# Welfare Months	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Food Stamps (Y); Intermediate Family Received Public Assistance (Y); Intermediate Family Received Social Security or Railroad Payments (Y); Intermediate Family Received Supplemental Security Income (Y); Intermediate Family Received Welfare/Job Placement/Childcare Services (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
Total Income	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Food Stamps (Y); Intermediate Family Received Public Assistance (Y); Intermediate Family Received Social Security or Railroad Payments (Y); Intermediate Family Received Supplemental Security Income (Y); Intermediate Family Received Welfare/Job Placement/Childcare Services (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.4](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.70 Dichotomous Income Indicator Predictive Mean Models: People 18 to 25 Years Old

Variable Requiring Imputation	Variables Included in Predictive Mean Model (Dichotomous Income Indicators)
Social Security	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous) ; Number of Adults in Household; Income State Rank (L, M)
Supplemental Security	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Social Security or Railroad Payments (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
Food Stamps	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Public Assistance (Y); Intermediate Family Received Social Security or Railroad Payments (Y); Intermediate Family Received Supplemental Security Income (Y); Intermediate Family Received Welfare/Job Placement/Childcare Services (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
Welfare Payments	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Social Security or Railroad Payments (Y); Intermediate Family Received Supplemental Security Income (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)

Table D.70 Dichotomous Income Indicator Predictive Mean Models: People 18 to 25 Years Old (continued)

Variable Requiring Imputation	Variables Included in Predictive Mean Model (Dichotomous Income Indicators)
Welfare Services	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Public Assistance (Y) ; Intermediate Family Received Social Security or Railroad Payments (Y); Intermediate Family Received Supplemental Security Income (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
# Welfare Months	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Food Stamps (Y); Intermediate Family Received Public Assistance (Y); Intermediate Family Received Social Security or Railroad Payments (Y); Intermediate Family Received Supplemental Security Income (Y); Intermediate Family Received Welfare/Job Placement/Childcare Services (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
Total Income	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Food Stamps (Y); Intermediate Family Received Public Assistance (Y); Intermediate Family Received Social Security or Railroad Payments (Y); Intermediate Family Received Supplemental Security Income (Y); Intermediate Family Received Welfare/Job Placement/Childcare Services (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.4](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.71 Dichotomous Income Indicator Predictive Mean Models: People 26 to 64 Years Old

Variable Requiring Imputation	Variables Included in Predictive Mean Model (Dichotomous Income Indicators)
Social Security	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
Supplemental Security	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Social Security or Railroad Payments (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
Food Stamps	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Public Assistance (Y); Intermediate Family Received Social Security or Railroad Payments (Y); Intermediate Family Received Supplemental Security Income (Y); Intermediate Family Received Welfare/Job Placement/Childcare Services (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
Welfare Payments	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Social Security or Railroad Payments (Y); Intermediate Family Received Supplemental Security Income (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)

Table D.71 Dichotomous Income Indicator Predictive Mean Models: People 26 to 64 Years Old (continued)

Variable Requiring Imputation	Variables Included in Predictive Mean Model (Dichotomous Income Indicators)
Welfare Services	Age; Age Squared; Age Cubed ; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Public Assistance (Y) ; Intermediate Family Received Social Security or Railroad Payments (Y); Intermediate Family Received Supplemental Security Income (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
# Welfare Months	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Food Stamps (Y); Intermediate Family Received Public Assistance (Y); Intermediate Family Received Social Security or Railroad Payments (Y); Intermediate Family Received Supplemental Security Income (Y); Intermediate Family Received Welfare/Job Placement/Childcare Services (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
Total Income	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Food Stamps (Y); Intermediate Family Received Public Assistance (Y); Intermediate Family Received Social Security or Railroad Payments (Y); Intermediate Family Received Supplemental Security Income (Y); Intermediate Family Received Welfare/Job Placement/Childcare Services (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.4](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.72 Dichotomous Income Indicator Predictive Mean Models: People 65 Years or Older

Variable Requiring Imputation	Variables Included in Predictive Mean Model (Dichotomous Income Indicators)
Social Security	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
Supplemental Security	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Social Security or Railroad Payments (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
Food Stamps	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Public Assistance (Y); Intermediate Family Received Social Security or Railroad Payments (Y); Intermediate Family Received Supplemental Security Income (Y); Intermediate Family Received Welfare/Job Placement/Childcare Services (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
Welfare Payments	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Social Security or Railroad Payments (Y); Intermediate Family Received Supplemental Security Income (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)

Table D.72 Dichotomous Income Indicator Predictive Mean Models: People 65 Years or Older (continued)

Variable Requiring Imputation	Variables Included in Predictive Mean Model (Dichotomous Income Indicators)
Welfare Services	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H) ; Education Level (4 Levels) (L, HS, SC) ; Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S) ; Percentage Hispanic/Latino in Block (H, M) ; Percentage Non-Hispanic/Latino Black/African American in Block (H, M) ; Percentage Owner Occupied in Block (H, M) ; Age * Gender ; Age Squared * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH) ; Age * Race/Hispanic Recode (4 Levels) (W, B, H) ; Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H) ; Intermediate Family Received Public Assistance (Y) ; Intermediate Family Received Social Security or Railroad Payments (Y); Intermediate Family Received Supplemental Security Income (Y); Number of People Younger than 18 Years Old in Household ; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
# Welfare Months	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Food Stamps (Y); Intermediate Family Received Public Assistance (Y); Intermediate Family Received Social Security or Railroad Payments (Y); Intermediate Family Received Supplemental Security Income (Y); Intermediate Family Received Welfare/Job Placement/Childcare Services (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)
Total Income	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Intermediate Family Received Food Stamps (Y); Intermediate Family Received Public Assistance (Y); Intermediate Family Received Social Security or Railroad Payments (Y); Intermediate Family Received Supplemental Security Income (Y); Intermediate Family Received Welfare/Job Placement/Childcare Services (Y); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Income State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.4](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.73 Income Finer Categories Response Propensity Models

Age Group	Variables Included in Response Propensity Model (Finer Categorization)
12-17	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Family Received Food Stamps (N); Family Received Public Assistance (N); Family Received Social Security or Railroad Payments (N); Family Received Supplemental Security Income (N); Family Received Welfare/Job Placement/Childcare Services (N); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Total Family Income > or < \$20,000 (LT); Income State Rank (L, M)
18-25	Age; Age Squared; Age Cubed ; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H) ; Family Received Food Stamps (N); Family Received Public Assistance (N); Family Received Social Security or Railroad Payments (N); Family Received Supplemental Security Income (N); Family Received Welfare/Job Placement/Childcare Services (N); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Total Family Income > or < \$20,000 (LT); Income State Rank (L, M)
26-64	Age; Age Squared ; Age Cubed ; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H) ; Family Received Food Stamps (N); Family Received Public Assistance (N); Family Received Social Security or Railroad Payments (N); Family Received Supplemental Security Income (N); Family Received Welfare/Job Placement/Childcare Services (N); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Total Family Income > or < \$20,000 (LT); Income State Rank (L, M)

Table D.73 Income Finer Categories Response Propensity Models (continued)

Age Group	Variables Included in Response Propensity Model (Finer Categorization)
65+	Age; Age Squared ; Age Cubed ; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN) ; Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender ; Age Squared * Gender ; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Family Received Food Stamps (N); Family Received Public Assistance (N) ; Family Received Social Security or Railroad Payments (N); Family Received Supplemental Security Income (N); Family Received Welfare/Job Placement/Childcare Services (N) ; Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Total Family Income > or < \$20,000 (LT); Income State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.4](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.74 Income Finer Categories Predictive Mean Models

Age Group	Variables Included in Predictive Mean Model (Finer Categorization)
12-17	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Family Received Food Stamps (N); Family Received Public Assistance (N); Family Received Social Security or Railroad Payments (N); Family Received Supplemental Security Income (N); Family Received Welfare/Job Placement/Childcare Services (N); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Total Family Income > or < \$20,000 (LT); Income State Rank (L, M)
18-25	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (3 Levels) (M, WM); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Family Received Food Stamps (N); Family Received Public Assistance (N); Family Received Social Security or Railroad Payments (N); Family Received Supplemental Security Income (N); Family Received Welfare/Job Placement/Childcare Services (N); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Total Family Income > or < \$20,000 (LT); Income State Rank (L, M)
26-64	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Family Received Food Stamps (N); Family Received Public Assistance (N); Family Received Social Security or Railroad Payments (N); Family Received Supplemental Security Income (N); Family Received Welfare/Job Placement/Childcare Services (N); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Total Family Income > or < \$20,000 (LT); Income State Rank (L, M)

Table D.74 Income Finer Categories Predictive Mean Models (continued)

Age Group	Variables Included in Predictive Mean Model (Finer Categorization)
65+	Age; Age Squared; Age Cubed; Gender (M); Race/Hispanic Recode (4 Levels) (W, B, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Hispanic/Latino in Block (H, M); Percentage Non-Hispanic/Latino Black/African American in Block (H, M); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MW, MB, MH); Age * Race/Hispanic Recode (4 Levels) (W, B, H); Age Squared * Race/Hispanic Recode (4 Levels) (W, B, H); Family Received Food Stamps (N); Family Received Public Assistance (N); Family Received Social Security or Railroad Payments (N); Family Received Supplemental Security Income (N); Family Received Welfare/Job Placement/Childcare Services (N); Number of People Younger than 18 Years Old in Household; Number of People Older than 64 Years Old in Household (Continuous); Number of Adults in Household; Total Family Income > or < \$20,000 (LT); Income State Rank (L, M)

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.4](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

D.7 Health Insurance Variables

Exhibit D.5 Definitions of Levels for Categorical Variables

Gender

M: Male, F: Female¹

Race/Hispanic Recode (4 Levels)

W: Not Hispanic/Latino and White Only,¹ B: Not Hispanic/Latino and Black/African American Only, O: Not Hispanic/Latino and Other, H: Hispanic/Latino

Education Level (4 Levels)

L: Less than High School, HS: High School Graduate, SC: Some College, C: College Graduate¹

Employment Status

FT: Full Time, PT: Part Time, UN: Unemployed, OE: Other Employment¹

Marital Status (2 Levels)

M: Married, NM: Never Been Married¹

Marital Status (4 Levels)

M: Married, W: Widowed, DS: Divorced/Separated, NM: Never Been Married¹

Population Density (3 Levels)

R: Segment Not in CBSA, SM: Segment in CBSA with Fewer than 1 Million People, L: Segment in CBSA with 1 Million or More People¹

Percentage Owner Occupied in Block

H: $\geq 50\%$, M: $[10-50)\%$, L: $< 10\%$ ¹

Gender * Race/Hispanic Recode (4 Levels)

MW: Male Not Hispanic/Latino and White Only,¹ MB: Male Not Hispanic/Latino and Black/African American Only, MO: Male Not Hispanic/Latino and Other, MH: Male Hispanic/Latino, FW: Female Not Hispanic/Latino and White Only,¹ FB: Female Not Hispanic/Latino and Black/African American Only,¹ FO: Female Not Hispanic/Latino and Other,¹ FH: Female Hispanic/Latino¹

Age * Race/Hispanic Recode (4 Levels)

W: Age * Not Hispanic/Latino and White Only,¹ B: Age * Not Hispanic/Latino and Black/African American Only, O: Age * Not Hispanic/Latino and Other, H: Age * Hispanic/Latino

Age Squared * Race/Hispanic Recode (4 Levels)

W: Age Squared * Not Hispanic/Latino and White Only,¹ B: Age Squared * Not Hispanic/Latino and Black/African American Only, O: Age Squared * Not Hispanic/Latino and Other, H: Age Squared * Hispanic/Latino

Other Family Members in Household

N: No, Y: Yes¹

Family Received Social Security or Railroad Payments

Y: Yes, N: No¹

Family Participation in Government Assistance Programs

N: No,¹ Y: Yes

Lifetime Military Service

N: No,¹ Y: Yes

Household Size (2 Levels)

L: At Least Four People in Household, S: Fewer than Four People in Household¹

Intermediate MEDICARE Coverage

Y: Yes, N: No¹

Intermediate CHAMPUS Coverage

Y: Yes, N: No¹

Intermediate MEDICAID/CHIP Coverage

Y: Yes, N: No¹

Family Income Recode

L: Income Less than \$20,000, M1: Income \$20,000-\$49,999, M2: Income \$50,000-\$74,999, H: Income \$75,000 or More¹

NOTE: An asterisk "*" represents an interaction between two variables.

CBSA = core-based statistical area.

¹ This is the reference level for this variable, against which effects of other factor levels are measured.

Table D.75 Health Insurance: Response Propensity Models

Age Group	Set of Variables Used to Determine Nonresponse	Variables Included in Response Propensity Model
12-17	Medicaid/CHIP, Medicare, CHAMPUS, Private Health Insurance	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Family Income Recode (L, M1, M2)
	Other Health Insurance	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Family Income Recode (L, M1, M2)
18-25	Medicaid/CHIP, Medicare, CHAMPUS, Private Health Insurance	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (2 Levels) (M); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Family Income Recode (L, M1, M2)
	Other Health Insurance	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (2 Levels) (M); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Family Income Recode (L, M1, M2)

Table D.75 Health Insurance: Response Propensity Models (continued)

Age Group	Set of Variables Used to Determine Nonresponse	Variables Included in Response Propensity Model
26-64	Medicaid/CHIP, Medicare, CHAMPUS, Private Health Insurance	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Family Income Recode (L, M1, M2)
	Other Health Insurance ¹	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Family Income Recode (L, M1, M2)
65+	Medicaid/CHIP, Medicare, CHAMPUS, Private Health Insurance	Age; Age Squared ; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Marital Status (2 Levels) (M); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Family Income Recode (L)
	Other Health Insurance ¹	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Family Income Recode (L, M1, M2)

NOTE: An asterisk "*" represents an interaction between two variables.

¹ The 26-64 and 65+ age groups were included in the same response propensity model for other health insurance.

Table D.76 Health Insurance: Predictive Mean Models, People 12 to 17 Years Old

Variable Requiring Imputation	Variables Included in Predictive Mean Model
Medicaid/CHIP	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Family Received Social Security or Railroad Payments (Y); Family Participation in Government Assistance Programs (Y); Household Size; Family Income Recode (L, M1, M2)
Medicare	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Family Received Social Security or Railroad Payments (Y); Intermediate MEDICAID/CHIP Coverage (Y)
CHAMPUS	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Intermediate MEDICARE Coverage (Y); Intermediate MEDICAID/CHIP Coverage (Y); Family Income Recode (L)
Private Health Insurance	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Family Received Social Security or Railroad Payments (Y); Family Participation in Government Assistance Programs (Y); Household Size; Intermediate MEDICARE Coverage (Y); Intermediate CHAMPUS Coverage (Y); Intermediate MEDICAID/CHIP Coverage (Y); Family Income Recode (L, M1, M2)
Other Health Insurance	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Family Received Social Security or Railroad Payments (Y); Family Participation in Government Assistance Programs (Y); Household Size; Family Income Recode (L, M1, M2)

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.77 Health Insurance: Predictive Mean Models, People 18 to 25 Years Old

Variable Requiring Imputation	Variables Included in Predictive Mean Model
Medicaid/CHIP	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (2 Levels) (M); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Other Family Members in Household (N); Family Received Social Security or Railroad Payments (Y); Family Participation in Government Assistance Programs (Y); Household Size; Family Income Recode (L, M1, M2)
Medicare¹	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Family Received Social Security or Railroad Payments (Y); Intermediate MEDICAID/CHIP Coverage (Y)
CHAMPUS	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (2 Levels) (M); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Lifetime Military Service (Y) ; Intermediate MEDICARE Coverage (Y); Intermediate MEDICAID/CHIP Coverage (Y); Family Income Recode (L)
Private Health Insurance	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (2 Levels) (M); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Other Family Members in Household (N); Family Received Social Security or Railroad Payments (Y); Family Participation in Government Assistance Programs (Y); Household Size; Intermediate MEDICARE Coverage (Y); Intermediate CHAMPUS Coverage (Y); Intermediate MEDICAID/CHIP Coverage (Y); Family Income Recode (L, M1, M2)
Other Health Insurance	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (2 Levels) (M); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Other Family Members in Household (N); Family Received Social Security or Railroad Payments (Y); Family Participation in Government Assistance Programs (Y); Household Size; Family Income Recode (L, M1, M2)

NOTE: An asterisk "*" represents an interaction between two variables.

¹ The 18-25 and 26-64 age groups were included in the same predictive mean model for Medicare.

Table D.78 Health Insurance: Predictive Mean Models, People 26 to 64 Years Old

Variable Requiring Imputation	Variables Included in Predictive Mean Model
Medicaid/CHIP	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Other Family Members in Household (N); Family Received Social Security or Railroad Payments (Y); Family Participation in Government Assistance Programs (Y); Household Size; Family Income Recode (L, M1, M2)
Medicare¹	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Family Received Social Security or Railroad Payments (Y); Intermediate MEDICAID/CHIP Coverage (Y)
CHAMPUS	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Lifetime Military Service (Y) ; Intermediate MEDICARE Coverage (Y); Intermediate MEDICAID/CHIP Coverage (Y); Family Income Recode (L)
Private Health Insurance	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Employment Status (FT, PT, UN); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Other Family Members in Household (N); Family Received Social Security or Railroad Payments (Y); Family Participation in Government Assistance Programs (Y); Household Size; Intermediate MEDICARE Coverage (Y); Intermediate CHAMPUS Coverage (Y); Intermediate MEDICAID/CHIP Coverage (Y); Family Income Recode (L, M1, M2)
Other Health Insurance²	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Other Family Members in Household (N); Family Received Social Security or Railroad Payments (Y); Family Participation in Government Assistance Programs (Y); Household Size; Family Income Recode (L, M1, M2)

NOTE: An asterisk "*" represents an interaction between two variables.

¹ The 18-25 and 26-64 age groups were included in the same predictive mean model for Medicare.

² The 26-64 and 65+ age groups were included in the same predictive mean model for other health insurance.

Table D.79 Health Insurance: Predictive Mean Models, People 65 Years or Older

Variable Requiring Imputation	Variables Included in Predictive Mean Model
Medicaid/CHIP	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Marital Status (2 Levels) (M); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Other Family Members in Household (N); Family Received Social Security or Railroad Payments (Y); Family Participation in Government Assistance Programs (Y); Household Size (2 Levels) (L); Family Income Recode (L)
Medicare	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Marital Status (2 Levels) (M); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Family Received Social Security or Railroad Payments (Y); Intermediate MEDICAID/CHIP Coverage (Y)
CHAMPUS	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Marital Status (2 Levels) (M); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Lifetime Military Service (Y); Intermediate MEDICARE Coverage (Y); Intermediate MEDICAID/CHIP Coverage (Y); Family Income Recode (L)
Private Health Insurance	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Marital Status (2 Levels) (M); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Other Family Members in Household (N); Family Received Social Security or Railroad Payments (Y); Family Participation in Government Assistance Programs (Y); Household Size (2 Levels) (L); Intermediate MEDICARE Coverage (Y); Intermediate CHAMPUS Coverage (Y); Intermediate MEDICAID/CHIP Coverage (Y); Family Income Recode (L)
Other Health Insurance¹	Age; Age Squared; Gender (M); Race/Hispanic Recode (4 Levels) (B, O, H); Education Level (4 Levels) (L, HS, SC); Marital Status (4 Levels) (M, W, DS); Population Density (3 Levels) (R, SM); Percentage Owner Occupied in Block (H, M); Age * Gender; Age Squared * Gender; Gender * Race/Hispanic Recode (4 Levels) (MB, MO, MH); Age * Race/Hispanic Recode (4 Levels) (B, O, H); Age Squared * Race/Hispanic Recode (4 Levels) (B, O, H); Other Family Members in Household (N); Family Received Social Security or Railroad Payments (Y); Family Participation in Government Assistance Programs (Y); Household Size; Family Income Recode (L, M1, M2)

NOTE: An asterisk "*" represents an interaction between two variables.

¹ The 26-64 and 65+ age groups were included in the same predictive mean model for other health insurance.

D.8 Pair Relationship Variables

Exhibit D.6 Definitions of Levels for Categorical Variables

Gender - Non-Pair

M: Male,¹ F: Female

Gender (Younger)

F: Female, M: Male¹

Gender (Older)

F: Female, M: Male¹

Age Category - Non-Pair

Y: 12-14, T1: 15-17, T2: 18-20, T3: 21-25, A1: 26-34, A2: 35-49, A3: 50+, UK: Unknown¹

Age Category (Older)

Y: 12-14, T1: 15-17, T2: 18-20, T3: 21-25, A1: 26-34, A2: 35-49, A3: 50+, UK: Unknown¹

Race/Hispanic Recode (4 Levels) - Non-Pair

W: Not Hispanic/Latino and White Only,¹ B: Not Hispanic/Latino and Black/African American Only, O: Not Hispanic/Latino and Other, H: Hispanic/Latino

Race/Hispanic Recode (Older) (4 Levels)

H: Hispanic/Latino, O: Not Hispanic/Latino and Other, B: Not Hispanic/Latino and Black/African American Only, W: Not Hispanic/Latino and White Only¹

Education Level (4 Levels) - Non-Pair

L: Less than High School, HS: High School Graduate, SC: Some College, C: College Graduate¹

Education Level (Younger) (4 Levels)

L: Less than High School, HS: High School Graduate, SC: Some College, C: College Graduate¹

Education Level (Older) (4 Levels)

L: Less than High School, HS: High School Graduate, SC: Some College, C: College Graduate¹

Employment Status - Non-Pair

FT: Full Time, PT: Part Time, UN: Unemployed, OE: Other Employment¹

Employment Status (Younger)

FT: Full Time, PT: Part Time, UN: Unemployed, OE: Other Employment¹

Employment Status (Older)

FT: Full Time, PT: Part Time, UN: Unemployed, OE: Other Employment¹

Marital Status (2 Levels) - Non-Pair

MWD: Married, Widowed or Divorced, NM: Never Been Married¹

Marital Status (Older) (2 Levels)

MWD: Married, Widowed or Divorced, NM: Never Been Married¹

Marital Status (Younger) (4 Levels)

M: Married, W: Widowed, D: Divorced, NM: Never Been Married¹

Marital Status (Older) (4 Levels)

M: Married, W: Widowed, D: Divorced, NM: Never Been Married¹

Population Density (3 Levels)

R: Segment Not in CBSA, SM: Segment in CBSA with Fewer than 1 Million People, L: Segment in CBSA with 1 Million or More People¹

Census Region

N: Northeast, M: Midwest, S: South, W: West¹

Percentage Black/African American in Block

H: $\geq 40\%$, M: $[10,40)\%$, L: $< 10\%$ ¹

Percentage Hispanic/Latino in Block

H: $\geq 71\%$, M: $[20,71)\%$, L: $< 20\%$ ¹

Percentage Owner Occupied in Block

H: $\geq 50\%$, M: $[10,50)\%$, L: $< 10\%$ ¹

NOTE: An asterisk "*" represents an interaction between two variables.

CBSA = core-based statistical area.

¹ This is the reference level for this variable, against which effects of other factor levels are measured.

Table D.80 Model Summaries (Pair Relationships)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
0 (12-14, 12-14)	Gender (Younger) (F); Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+
1 (12-14, 15-17)	Gender (Younger) (F); Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

Table D.80 Model Summaries (Pair Relationships) (continued)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
2 (12,14, 18-25)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (4 Levels) (M, W, D); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (4 Levels) (M, W, D) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (4 Levels) (M, W, D) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+
3 (15-17, 15-17)	Gender (Younger) (F); Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F) ; Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M) ; Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

Table D.80 Model Summaries (Pair Relationships) (continued)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
4 (15-17, 18-25)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (4 Levels) (M, W, D) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (4 Levels) (M, W, D) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (4 Levels) (M, W, D) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

Table D.80 Model Summaries (Pair Relationships) (continued)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
5 (18-20, 18-25)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Younger) (4 Levels) (L, HS, SC) ; Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Younger) (FT, PT, UN); Employment Status (Older) (FT, PT, UN); Marital Status (Younger) (4 Levels) (M, W, D); Marital Status (Older) (4 Levels) (M, W, D) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F) ; Gender (Older) (F); Age Category (Older) (T3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B) ; Education Level (Younger) (4 Levels) (L, HS, SC) ; Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Younger) (FT, PT, UN) ; Employment Status (Older) (FT, PT, UN) ; Marital Status (Younger) (4 Levels) (M, W, D) ; Marital Status (Older) (4 Levels) (M, W, D) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F) ; Gender (Older) (F); Age Category (Older) (T3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Younger) (4 Levels) (L, HS, SC) ; Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Younger) (FT, PT, UN) ; Employment Status (Older) (FT, PT, UN) ; Marital Status (Younger) (4 Levels) (M, W, D) ; Marital Status (Older) (4 Levels) (M, W, D) ; Population Density (3 Levels) (R, SM) ; Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M) ; Percentage Owner Occupied in Block (H, M) ; Number in Household Aged 0-11 ; Number in Household Aged 12-17 ; Number in Household Aged 18-25 ; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

Table D.80 Model Summaries (Pair Relationships) (continued)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
6 (21-25, 21-25)	Gender (Younger) (F); Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Younger) (4 Levels) (L, HS, SC); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Younger) (FT, PT, UN); Employment Status (Older) (FT, PT, UN); Marital Status (Younger) (4 Levels) (M, W, D); Marital Status (Older) (4 Levels) (M, W, D); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Younger) (4 Levels) (L, HS, SC); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Younger) (FT, PT, UN); Employment Status (Older) (FT, PT, UN); Marital Status (Younger) (4 Levels) (M, W, D); Marital Status (Older) (4 Levels) (M, W, D); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Younger) (4 Levels) (L, HS, SC); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Younger) (FT, PT, UN); Employment Status (Older) (FT, PT, UN); Marital Status (Younger) (4 Levels) (M, W, D); Marital Status (Older) (4 Levels) (M, W, D); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+
7 (12-14, 26+)	Gender (Younger) (F); Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (4 Levels) (M, W, D); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (4 Levels) (M, W, D); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

Table D.80 Model Summaries (Pair Relationships) (continued)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
8 (15-17, 26+)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (4 Levels) (M, W, D); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (4 Levels) (M, W, D); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (4 Levels) (M, W, D); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

Table D.80 Model Summaries (Pair Relationships) (continued)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
9 (18-20, 26+)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Younger) (4 Levels) (L, HS, SC); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Younger) (FT, PT, UN); Employment Status (Older) (FT, PT, UN); Marital Status (Younger) (4 Levels) (M, W, D) ; Marital Status (Older) (4 Levels) (M, W, D); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Younger) (4 Levels) (L, HS, SC) ; Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Younger) (FT, PT, UN); Employment Status (Older) (FT, PT, UN); Marital Status (Younger) (4 Levels) (M, W, D) ; Marital Status (Older) (4 Levels) (M, W, D); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Younger) (4 Levels) (L, HS, SC) ; Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Younger) (FT, PT, UN); Employment Status (Older) (FT, PT, UN); Marital Status (Younger) (4 Levels) (M, W, D) ; Marital Status (Older) (4 Levels) (M, W, D); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

Table D.80 Model Summaries (Pair Relationships) (continued)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
10 (21+, 26+)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Younger) (4 Levels) (L, HS, SC); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Younger) (FT, PT, UN); Employment Status (Older) (FT, PT, UN); Marital Status (Younger) (4 Levels) (M, W, D); Marital Status (Older) (4 Levels) (M, W, D); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Younger) (4 Levels) (L, HS, SC); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Younger) (FT, PT, UN); Employment Status (Older) (FT, PT, UN); Marital Status (Younger) (4 Levels) (M, W, D) ; Marital Status (Older) (4 Levels) (M, W, D); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous) ; Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (A2, A3) ; Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Younger) (4 Levels) (L, HS, SC); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Younger) (FT, PT, UN); Employment Status (Older) (FT, PT, UN); Marital Status (Younger) (4 Levels) (M, W, D) ; Marital Status (Older) (4 Levels) (M, W, D) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11 ; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.6](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.81 Model Summaries (Multiplicities)

Pair Domain	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
Parent-Child (12-20) Child Focus	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3, A1, A2, A3) ; Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T2, T3, A1, A2) ; Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T2, T3, A1, A2) ; Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49 ; Number in Household Aged 50+
Parent-Child (12-20) Parent Focus	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents	No Model Used; No Nonrespondents
Sibling (12-14) Sibling (15-17) Older Sibling Focus	Gender (Younger) (F); Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

Table D.81 Model Summaries (Multiplicities) (continued)

Pair Domain	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
Sibling (12-14) Sibling (15-17) Younger Sibling Focus	Gender (Younger) (F); Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+
Sibling (12-17) Sibling (18-25) Older Sibling Focus	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T2); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

Table D.81 Model Summaries (Multiplicities) (continued)

Pair Domain	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
Sibling (12-17) Sibling (18-25) Younger Sibling Focus	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T2) ; Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN) ; Marital Status (Older) (2 Levels) (MWD) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M) ; Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous) ; Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T2); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T2); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.6](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.82 Model Summaries (Household-Level Person Counts of Pair Domains When Respondent Is in a Responding Pair)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
Parent-Child (12-20) Child Focus, Both Pair Members Younger than 18	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (Y); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (Y); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (Y); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+
Parent-Child (12-20) Child Focus, at Least One Pair Member Older than 18	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

Table D.82 Model Summaries (Household-Level Person Counts of Pair Domains When Respondent Is in a Responding Pair) (continued)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
Parent-Child (12-20) Parent Focus, Both Pair Members Younger than 18	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (Y); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (Y); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (Y); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49 ; Number in Household Aged 50+
Parent-Child (12-20) Parent Focus, at Least One Pair Member Older than 18	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17 ; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49 ; Number in Household Aged 50+

Table D.82 Model Summaries (Household-Level Person Counts of Pair Domains When Respondent Is in a Responding Pair) (continued)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
Sibling (12-14) Sibling (15-17), Older Sibling Focus, Both Pair Members Younger than 18	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (Y); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (Y); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (Y); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+
Sibling (12-14) Sibling (15-17), Older Sibling Focus, at Least One Pair Member Older than 18	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

Table D.82 Model Summaries (Household-Level Person Counts of Pair Domains When Respondent Is in a Responding Pair) (continued)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
Sibling (12-17) Sibling (18-25), Older Sibling Focus, Both Pair Members Younger than 18	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (Y); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (Y); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (Y); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+
Sibling (12-17) Sibling (18-25), Older Sibling Focus, at Least One Pair Member Older than 18	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

Table D.82 Model Summaries (Household-Level Person Counts of Pair Domains When Respondent Is in a Responding Pair) (continued)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
Spouse-Spouse, Both Pair Members Younger than 18	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (Y); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (Y); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (Y); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+
Spouse-Spouse, at Least One Pair Member Older than 18	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

Table D.82 Model Summaries (Household-Level Person Counts of Pair Domains When Respondent Is in a Responding Pair) (continued)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
Spouse-Spouse with Children, Both Pair Members Younger than 18	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T1, T2, T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T1, T2, T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T1, T2, T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

Table D.82 Model Summaries (Household-Level Person Counts of Pair Domains When Respondent Is in a Responding Pair) (continued)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
Spouse-Spouse with Children, at Least One Pair Member Older than 18	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T1, T2, T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T1, T2, T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender (Younger) (F); Gender (Older) (F); Age Category (Older) (T1, T2, T3, A1, A2, A3); Race/Hispanic Recode (Older) (4 Levels) (H, O, B); Education Level (Older) (4 Levels) (L, HS, SC); Employment Status (Older) (FT, PT, UN); Marital Status (Older) (2 Levels) (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.6](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Table D.83 Model Summaries (Household-Level Person Counts of Pair Domains When Respondent Is Not in a Responding Pair)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
Parent-Child (12-20) Child Focus, Younger than 18	Gender - Non-Pair (F); Age Category - Non-Pair (Y); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (Y); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (Y); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+
Parent-Child (12-20) Child Focus, Older than 18	Gender - Non-Pair (F); Age Category - Non-Pair (T3, A1, A2, A3); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Education Level (4 Levels) - Non-Pair (L, HS, SC); Employment Status - Non-Pair (FT, PT, UN); Marital Status (2 Levels) - Non-Pair (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (T3, A1, A2, A3); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Education Level (4 Levels) - Non-Pair (L, HS, SC); Employment Status - Non-Pair (FT, PT, UN); Marital Status (2 Levels) - Non-Pair (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (T3, A1, A2, A3); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Education Level (4 Levels) - Non-Pair (L, HS, SC); Employment Status - Non-Pair (FT, PT, UN); Marital Status (2 Levels) - Non-Pair (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

Table D.83 Model Summaries (Household-Level Person Counts of Pair Domains When Respondent Is Not in a Responding Pair) (continued)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
Parent-Child (12-20) Parent Focus, Younger than 18	Gender - Non-Pair (F); Age Category - Non-Pair (Y); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (Y); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (Y); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49 ; Number in Household Aged 50+
Parent-Child (12-20) Parent Focus, Older than 18	Gender - Non-Pair (F); Age Category - Non-Pair (T3, A1, A2, A3); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Education Level (4 Levels) - Non-Pair (L, HS, SC); Employment Status - Non-Pair (FT, PT, UN); Marital Status (2 Levels) - Non-Pair (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (T3, A1, A2, A3) ; Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Education Level (4 Levels) - Non-Pair (L, HS, SC); Employment Status - Non-Pair (FT, PT, UN); Marital Status (2 Levels) - Non-Pair (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous) ; Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (T3, A1, A2, A3) ; Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Education Level (4 Levels) - Non-Pair (L, HS, SC); Employment Status - Non-Pair (FT, PT, UN); Marital Status (2 Levels) - Non-Pair (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25 ; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

Table D.83 Model Summaries (Household-Level Person Counts of Pair Domains When Respondent Is Not in a Responding Pair) (continued)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
Sibling (12-14) Sibling (15-17), Older Sibling Focus, Younger than 18	Gender - Non-Pair (F); Age Category - Non-Pair (Y); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (Y); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (Y); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17 ; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+
Sibling (12-14) Sibling (15-17), Older Sibling Focus, Older than 18	Gender - Non-Pair (F); Age Category - Non-Pair (T3, A1, A2, A3); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Education Level (4 Levels) - Non-Pair (L, HS, SC); Employment Status - Non-Pair (FT, PT, UN); Marital Status (2 Levels) - Non-Pair (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (T3, A1, A2, A3); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Education Level (4 Levels) - Non-Pair (L, HS, SC); Employment Status - Non-Pair (FT, PT, UN); Marital Status (2 Levels) - Non-Pair (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (T3, A1, A2, A3); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Education Level (4 Levels) - Non-Pair (L, HS, SC); Employment Status - Non-Pair (FT, PT, UN); Marital Status (2 Levels) - Non-Pair (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

Table D.83 Model Summaries (Household-Level Person Counts of Pair Domains When Respondent Is Not in a Responding Pair) (continued)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
Sibling (12-17) Sibling (18-25), Older Sibling Focus, Younger than 18	Gender - Non-Pair (F); Age Category - Non-Pair (Y); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (Y); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (Y); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+
Sibling (12-17) Sibling (18-25), Older Sibling Focus, Older than 18	Gender - Non-Pair (F); Age Category - Non-Pair (T3, A1, A2, A3); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Education Level (4 Levels) - Non-Pair (L, HS, SC); Employment Status - Non-Pair (FT, PT, UN); Marital Status (2 Levels) - Non-Pair (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (T3, A1, A2, A3); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Education Level (4 Levels) - Non-Pair (L, HS, SC); Employment Status - Non-Pair (FT, PT, UN); Marital Status (2 Levels) - Non-Pair (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (T3, A1, A2, A3); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Education Level (4 Levels) - Non-Pair (L, HS, SC); Employment Status - Non-Pair (FT, PT, UN); Marital Status (2 Levels) - Non-Pair (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

Table D.83 Model Summaries (Household-Level Person Counts of Pair Domains When Respondent Is Not in a Responding Pair) (continued)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
Spouse-Spouse, Younger than 18	Gender - Non-Pair (F); Age Category - Non-Pair (Y); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (Y); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (Y); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+
Spouse-Spouse, Older than 18	Gender - Non-Pair (F); Age Category - Non-Pair (T3, A1, A2, A3); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Education Level (4 Levels) - Non-Pair (L, HS, SC); Employment Status - Non-Pair (FT, PT, UN); Marital Status (2 Levels) - Non-Pair (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (T3, A1, A2, A3); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Education Level (4 Levels) - Non-Pair (L, HS, SC); Employment Status - Non-Pair (FT, PT, UN); Marital Status (2 Levels) - Non-Pair (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (T3, A1, A2, A3); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Education Level (4 Levels) - Non-Pair (L, HS, SC); Employment Status - Non-Pair (FT, PT, UN); Marital Status (2 Levels) - Non-Pair (MWD) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

Table D.83 Model Summaries (Household-Level Person Counts of Pair Domains When Respondent Is Not in a Responding Pair) (continued)

Model Group	Variables Included in Response Propensity Model	Variables Included in Predictive Mean Model	
		Including Household Size	Not Including Household Size
Spouse-Spouse with Children	Gender - Non-Pair (F); Age Category - Non-Pair (T1, T2, T3, A1, A2, A3); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Education Level (4 Levels) - Non-Pair (L, HS, SC); Employment Status - Non-Pair (FT, PT, UN); Marital Status (2 Levels) - Non-Pair (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous); Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (T1, T2, T3, A1, A2, A3); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Education Level (4 Levels) - Non-Pair (L, HS, SC); Employment Status - Non-Pair (FT, PT, UN); Marital Status (2 Levels) - Non-Pair (MWD); Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Household Size, Reconciled across Pairs (Continuous) ; Household Size, Reconciled across Pairs (5 Levels)	Gender - Non-Pair (F); Age Category - Non-Pair (T1, T2, T3, A1, A2, A3); Race/Hispanic Recode (4 Levels) - Non-Pair (B, O, H); Education Level (4 Levels) - Non-Pair (L, HS, SC); Employment Status - Non-Pair (FT, PT, UN); Marital Status (2 Levels) - Non-Pair (MWD) ; Population Density (3 Levels) (R, SM); Census Region (N, M, S); Percentage Black/African American in Block (H, M); Percentage Hispanic/Latino in Block (H, M); Percentage Owner Occupied in Block (H, M); Number in Household Aged 0-11; Number in Household Aged 12-17; Number in Household Aged 18-25; Number in Household Aged 26-34; Number in Household Aged 35-49; Number in Household Aged 50+

NOTE: Boldface of variables and levels indicates that they were dropped from the model.

NOTE: See [Exhibit D.6](#) for definitions of levels for categorical variables.

NOTE: An asterisk "*" represents an interaction between two variables.

Appendix E: Hot-Deck/Collocated Stochastic Imputation Procedure Summaries

This page intentionally left blank

Appendix E: Hot-Deck/Collocated Stochastic Imputation Procedure Summaries

E.1 Introduction

For the majority of variables that had missing values imputed in the 2016 National Survey on Drug Use and Health (NSDUH), the imputation method used was predictive mean neighborhood (PMN). This appendix summarizes the predictive mean vectors and the constraints applied during the PMN hot-deck step.¹ It is organized by groups of variables requiring imputation: demographics, household composition (roster), lifetime use of drugs, recency and frequency of drug use, age at first drug use, income, health insurance, and pair. There are three types of tables associated with each variable or set of variables imputed using the PMN method: (1) Logical Constraints, (2) Likeness Constraints, and (3) Constraints and Portion of the Predictive Mean Vector.

For variables that do not apply any logical constraints in the PMN process, the "Logical Constraints" table is not applicable. The "Constraints and Portion of the Predictive Mean Vector" table specifies the following for each missingness pattern:

1. the number of item nonrespondents exhibiting the pattern ("Total Number of Cases");
2. the set of logical constraints applied to the potential donors ("Logical Constraints");
3. the elements of the predictive mean vector ("Predictive Mean Vector") used to calculate the Mahalanobis distance from recipient to potential donor and to restrict the donor set via the delta constraint; and
4. the set of likeness constraints used in each try and the number of item nonrespondents who found donors on each try by different age groups ("Likeness Constraints: Number of Cases").²

In the tables that follow, the phrase "Donor's predicted means each must be within x percent of recipient's predicted means" appears in each of the multivariate imputation tables, and the phrase "Donor's predicted mean must be within x percent of recipient's predicted mean" appears in each of the univariate imputation tables. In either case, it represents one of the likeness constraints, known as the "delta constraint," and also defines the neighborhood. When this constraint is loosened, the neighborhood of potential donors is abandoned, and the candidate with the predicted mean closest to the recipient's (subject to the constraints that are still on the pool of donors) is chosen as the donor.

Although statistical imputation of the drug use or income variables could not have proceeded separately within each state because of insufficient pools of donors, the PMN procedure does incorporate information about the state of residence of each respondent. For the

¹ See Chapter 3 for details on PMN imputation.

² If a cell contains the text "No Cases" within "Likeness Constraints: Number of Cases," this means that there were no cases requiring imputation in that category.

drug use variables, in the hot-deck step of PMN, respondents were separated into three state usage-level categories for each drug, depending on the response variable of interest. Respondents from states with high usage of a given drug were placed in one category, respondents from medium-usage states were placed in another category, and respondents from low-usage states were placed in a third category. For the income variables, the states were separated into three income groups, depending upon the proportion of families within those states with incomes greater than or equal to \$20,000. As with the drug use variables, respondents from high-income states (by this measure) were placed in one category, respondents from medium-income states were placed in another category, and respondents from low-income states were placed in a third category. In the tables that follow, this variable is identified as the "state rank" for the drug use and income variables. It was used as a likeness constraint, where the set of eligible donors for each recipient was restricted so that donors and recipients were both from states with the same state rank.

E.2 Demographics

Tables E.1 through E.11 present information on the imputation procedures for the interviewer-administered demographic variables: race, Hispanic/Latino origin, Hispanic/Latino group, and education level. Tables E.12 through E.19 present information for the self-administered demographic variables: marital status, employment status, indicator of birth in the United States, and immigrant age of entry into the United States. In several instances, variable names are used without description for the purposes of brevity (see Chapter 4 for details). The segment-level variable, SEGID (segment ID), was used only in the likeness constraints for demographic imputation.

As described in the sample design report in the 2016 NSDUH methodological resource book (Center for Behavioral Health Statistics and Quality, 2017c) within each state, state sampling (SS) regions were formed, which were further partitioned into clusters of adjacent blocks called "segments." The segment ID number was a two-letter state abbreviation followed by a two-digit SS region and a two-digit segment identifier, which uniquely identifies each segment. Although the segment identifier was not used as a covariate because of the large number of levels, it was used as a constraint in the hot-deck step of the PMN procedure for race, Hispanicity, education, and employment status, as noted in Chapters 4 and 5.

E.2.1 Marital Status Variables

Table E.1 Likeness Constraints for Marital Status

Constraint #	Likeness Constraint
LikC1	Donor's age must be within 3 years of recipient's age
LikC2	Donor's predicted means must be within 5 percent of recipient's predicted means

Table E.2 Constraints and Portion of the Predictive Mean Vector for Marital Status

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
1	Completely missing	10	None	1. M1 2. M2 3. M3	No Cases	No Cases	1,2: 10

¹ The predictive mean vector components are defined by the following:

1. M1 = P(Married)
2. M2 = P(Widowed)
3. M3 = P(Divorced or separated)

E.2.2 Race Variables

Table E.3 Logical Constraints for Race

Constraint #	Logical Constraint
LogC1	Donor must be Asian, in part or in full
LogC2	Donor must be more than one race
LogC3	Donor must not be white only
LogC4	Donor must be white only or white and American Indian/Alaska Native only
LogC5	Donor must not be American Indian/Alaska Native, in part or in full

Table E.4 Likeness Constraints for Race

Constraint #	Likeness Constraint
LikC1	Segment of donor = segment of recipient
LikC2	Donor's predicted means must be within 5 percent of recipient's predicted means
LikC3	If recipient was Hispanic/Latino nonspecific, then donor must be of Hispanic/Latino origin
LikC4	If recipient selected one or more Hispanic/Latino categories, including Mexican, Puerto Rican, Central or South American, Cuban, Dominican, and Spaniard, then donor's Hispanic/Latino group value must be equal to one of the Hispanic/Latino groups mentioned by recipient
LikC5	Donor must be Mexican (Hispanic/Latino or non-Hispanic/Latino)
LikC6	Donor must be Cuban (Hispanic/Latino or non-Hispanic/Latino)
LikC7	Donor must be Central or South American (Hispanic/Latino or non-Hispanic/Latino)
LikC8	Donor must be Dominican (Hispanic/Latino or non-Hispanic/Latino)
LikC9	Donor must be Spanish (Hispanic/Latino or non-Hispanic/Latino)
LikC10	Donor must be Puerto Rican (Hispanic/Latino or non-Hispanic/Latino)
LikC11	Donor must be Puerto Rican or Cuban (Hispanic/Latino or non-Hispanic/Latino)

Table E.5 Constraints and Portion of the Predictive Mean Vector for Race

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
1	Completely missing	2,571	None	1. R1 2. R2 3. R3 4. R4	1-4: 87 2-4: 433 3,4: 260	1-4: 130 2-4: 390 3,4: 167	1-4: 64 2-4: 507 3,4: 533
2	Known to be Asian	13	1	1. R1 2. R2 3. R3 4. R4	1-4: 1 2-4: 2 3,4: 1	1-4: 0 2-4: 1 3,4: 2	1-4: 1 2-4: 3 3,4: 2
3	Known to be multiple race, but no other information	6	2	1. R1 2. R2 3. R3 4. R4	1-4: 0 2-4: 0 3,4: 1	1-4: 0 2-4: 0 3,4: 1	1-4: 0 2-4: 0 3,4: 4
4	Known not to be white, but no other information	12	3	1. R2/(1-R1) 2. R3/(1-R1) 3. R4/(1-R1)	1-4: 0 2-4: 2	1-4: 0 2-4: 1 3,4: 1	1-4: 0 2-4: 4 3,4: 4
5	Known to be white or both white and American Indian/Alaska Native	1	4	1. R1/(1-R2-R3-R4)	No Cases	1-4: 0 2-4: 1	No Cases
6	Known not to be American Indian/Alaska Native, in part or in full	0	5	1. R1/(1-R3) 2. R2/(1-R3) 3. R4/(1-R3)	No Cases	No Cases	No Cases
7	Known to be non-Hispanic Mexican	4	None	1. R1 2. R2 3. R3 4. R4	1-5: 1 2-5: 1	No Cases	1-5: 1 2-5: 1
8	Known to be non-Hispanic Cuban	0	None	1. R1 2. R2 3. R3 4. R4	No Cases	No Cases	No Cases
9	Known to be non-Hispanic Central or South American	0	None	1. R1 2. R2 3. R3 4. R4	No Cases	No Cases	No Cases

Table E.5 Constraints and Portion of the Predictive Mean Vector for Race (continued)

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
10	Known to be non-Hispanic Dominican	2	None	1. R1 2. R2 3. R3 4. R4	1-4,8: 0 2-4,8: 1	1-4,8: 0 2-4,8: 1	No Cases
11	Known to be non-Hispanic Spanish	1	None	1. R1 2. R2 3. R3 4. R4	1-4,9: 0 2-4,9: 0 3,4,9: 1	No Cases	No Cases
12	Known to be non-Hispanic Puerto Rican	0	None	1. R1 2. R2 3. R3 4. R4	No Cases	No Cases	No Cases
13	Known to be non-Hispanic Cuban and Puerto Rican	0	None	1. R1 2. R2 3. R3 4. R4	No Cases	No Cases	No Cases

¹ The predictive mean vector components are defined by the following:

1. R1 = P(White only)
2. R2 = P(Black only)
3. R3 = P(American Indian/Alaska Native only)
4. R4 = P(Asian only)

E.2.3 Hispanic/Latino Origin Variables

Table E.6 Likeness Constraints for Hispanic/Latino Origin

Constraint #	Likeness Constraint
LikC1	Segment of donor = segment of recipient
LikC2	Donor's predicted mean must be within 5 percent of recipient's predicted mean

Table E.7 Constraints and Portion of the Predictive Mean Vector for Hispanic/Latino Origin

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
1	Completely missing	222	None	1. H1	1,2: 100 2: 90	1,2: 5 2: 9	1,2: 6 2: 12

¹ The predictive mean vector components are defined by the following:

1. H1 = P(Hispanic/Latino origin)

E.2.4 Hispanic/Latino Group Variables

Table E.8 Likeness Constraints for Hispanic/Latino Group

Constraint #	Likeness Constraint
LikC1	Segment of donor = segment of recipient
LikC2	Donor's predicted mean must be within 5 percent of recipient's predicted mean
LikC3	If recipient had $8 \leq \text{EDHOGRP} \leq 21$, then donor's values for IRRACWH, IRRACEBK, IRRACENA, IRRACEAS, IRRACENH, and IRRACEPI must indicate a subset of the racial categories mentioned by recipient

Table E.9 Constraints and Portion of the Predictive Mean Vector for Hispanic/Latino Group

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases ²
1	Completely missing	83	None	1. H1 2. H2 3. H3	1-3: 8 2,3: 63 3: 12

¹ The predictive mean vector components are defined by the following:

1. H1 = P(Mexican)
2. H2 = P(Puerto Rican)
3. H3 = P(Central or South American)

² The hot-deck program for the Hispanic/Latino group is not separated into age groups.

E.2.5 Education Variables

Table E.10 Likeness Constraints for Education Level

Constraint #	Likeness Constraint
LikC1	Segment of donor = segment of recipient
LikC2	Donor's predicted mean must be within 5 percent of recipient's predicted mean
LikC3	Age of donor = age of recipient

Table E.11 Constraints and Portion of the Predictive Mean Vector for Education Level

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector, by Age Group ¹		Likeness Constraints: Number of Cases, by Age Group		
				12-17	18+	12-17	18-25	26+
1	Missing	64	None	1. E1 2. E2 3. E3 4. E4	1. E1 2. E2 3. E3	1-3: 1 2,3: 8 3: 1	1-3: 3 2,3: 25 3: 3	1,2: 2 2: 21

¹ The predictive mean vector components are defined by the following:

12-17 Age Group

1. E1 = P(Less than elementary school)
2. E2 = P(Elementary school)
3. E3 = P(Middle school)
4. E4 = P(Some high school)

18+ Age Group

1. E1 = P(Less than high school)
2. E2 = P(High school)
3. E3 = P(Some college)

E.2.6 Employment Variables

Table E.12 Logical Constraints for Employment Status

Constraint #	Logical Constraint
LogC1	Donor must be employed

Table E.13 Likeness Constraints for Employment Status

Constraint #	Likeness Constraint
LikC1	Segment of donor = segment of recipient
LikC2	Donor's predicted mean must be within 5 percent of recipient's predicted mean
LikC3	Donor's age must be within 5 years of recipient's age

Table E.14 Constraints and Portion of the Predictive Mean Vector for Employment Status

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
1	Completely missing	625	None	1. E1 2. E2 3. E3	1-3: 13 2,3: 87 3: 10	1-3: 12 2,3: 168 3: 20	1-3: 2 2,3: 238 3: 75
2	Known to be employed; part-time vs. full-time status unknown	100	1	1. E1/ (E1+E2)	1-3: 0 2,3: 14	1-3: 2 2,3: 50	1-3: 6 2,3: 28

¹ The predictive mean vector components are defined by the following:

1. E1 = P(Employed full time)
2. E2 = P(Employed part time)
3. E3 = P(Unemployed)

E.2.7 Immigrant Variables

Table E.15 Likeness Constraints for Indicator of Birth in the United States

Constraint #	Likeness Constraint
LikC1	Segment of donor = segment of recipient
LikC2	Donor's predicted mean must be within 5 percent of recipient's predicted mean

Table E.16 Constraints and Portion of the Predictive Mean Vector for Indicator of Birth in the United States

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
1	Completely missing	289	None	1. B1	1,2: 28 2: 75	1,2: 12 2: 52	1,2: 36 2: 86

¹ The predictive mean vector components are defined by the following:

1. B1 = P(Born in the United States)

Table E.17 Logical Constraints for Age of Entry in the United States

Constraint #	Likeness Constraint
LogC1	Donor's age of entry must be less than recipient's current age
LogC2	Difference between recipient's current age and donor's age of entry must be less than or equal to 1 if recipient lived in the United States less than a year, or difference must be greater than 1 if recipient lived in the United States more than a year

Table E.18 Likeness Constraints for Age of Entry in the United States

Constraint #	Likeness Constraint
LikC1	Segment of donor = segment of recipient
LikC2	Donor's predicted mean must be within 5 percent of recipient's predicted mean

Table E.19 Constraints and Portion of the Predictive Mean Vector for Age of Entry in the United States

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases ²
1	Completely missing	252	1,2	1. PrAgeEntry	1,2: 54 2: 195 None: 3

¹ The predictive mean vector components are defined by the following:

1. PrAgeEntry = Predicted age of entry

² The hot-deck program for immigrant age of entry is not separated into age groups.

E.3 Drug Variables

Drug variables were imputed with two methods, PMN and modified PMN (modPMN) (see Chapter 3 for details). Cigarettes, cigars, smokeless tobacco, pipes, alcohol, heroin, marijuana, cocaine, and crack are presented in Section E.3.1, PMN. Inhalants, methamphetamine, hallucinogens, hallucinogen child drugs, pain relievers, tranquilizers, sedatives, stimulants, and OxyContin are presented in Section E.3.2, modPMN.

Section E.3.1 presents information on the missingness patterns, constraints, and predictive mean vectors applied during PMN imputation procedures for lifetime drug use, recency and frequency of use, and age at first use. [Tables E.20](#) through [E.54](#) contain information on the imputation procedures for drugs that continued to be imputed using PMN.

Section E.3.2 presents information on the imputation procedures applied under the modPMN method. For continuous variables, information on the missingness patterns, constraints, and predictive mean vectors used for hot-deck imputation is presented. For categorical variables, information on the edited values, restrictions from previously imputed variables, and predictive mean vectors used for collocated stochastic imputation is presented. [Tables E.55](#) through [E.124](#) contain information by measure on the imputation procedures for drugs that were imputed using modPMN.

E.3.1 PMN

E.3.1.1 Lifetime Drug Use

There were a large number of missingness patterns for lifetime drug use. The usable case rule (Section 2.2.3) required a certain minimum response pattern, but other than that any combination of the lifetime drug variables may have been missing for a given respondent record.

The probabilities associated with the eight lifetime drug variables listed in [Table E.20](#) formed the full predictive mean vector. Only the probabilities associated with the lifetime drug variables for which values were missing were used in the predictive mean vector for each item nonrespondent. The predicted mean for cigarette lifetime use was never calculated because no respondents were missing the lifetime variable for cigarette use.³

Table E.20 Elements of Full Predictive Mean Vector for Lifetime Drug Use

Lifetime Drug Use	Predicted Mean
Heroin Lifetime	P(Lifetime User)
Crack Lifetime	P(Lifetime User Lifetime User of Cocaine), if cocaine lifetime use is known P(Lifetime User of Cocaine) * P(Lifetime User Lifetime User of Cocaine), if cocaine lifetime use is unknown ¹
Cocaine Lifetime	P(Lifetime User)
Marijuana Lifetime	P(Lifetime User)
Alcohol Lifetime	P(Lifetime User)
Pipes Lifetime	P(Lifetime User)
Cigars Lifetime	P(Lifetime User)
Smokeless Tobacco Lifetime	P(Lifetime User)

¹ The predicted mean for individuals missing crack lifetime where cocaine lifetime use is unknown can also be written as the joint probability of P(Lifetime User of Cocaine and Lifetime User of Crack).

Table E.21 Likeness Constraints for Lifetime

Constraint #	Likeness Constraint
LikC1	State rank of donor = state rank of recipient
LikC2	Donor's predicted means each must be within 5 percent of recipient's predicted means
LikC3	Lifetime use of donor = lifetime use of recipient, for each nonmissing lifetime indicator
LikC4	If recipient was missing the lifetime indicator(s) for any member of a family of drugs, ¹ then donor's lifetime indicator(s) agreed with recipient's nonmissing lifetime indicator(s) within that family

¹ The cocaine family includes crack.

³ The lifetime variable for cigarette use was never missing because the usable case rule required a nonmissing response to the cigarette gate question.

Table E.22 Constraints and Portion of the Predictive Mean Vector for Lifetime

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
1	Missing one or more of the lifetime drug use variables	180	None	1. LF for gate question associated with each missing variable	1-3: 90 1,2,4: 2 1,4: 3	1-3: 26 1,2,4: 2 1,4: 2	1-3: 47 1,2,4: 5 1,4: 3

¹ The predictive mean vector components are defined by the following:

1. LF = P(Lifetime user of x), where x is the drug associated with the missing gate question

Table E.23 Likeness Constraints for Cigarette Ever Daily Used

Constraint #	Likeness Constraint
LikC1	Age of donor = age of recipient
LikC2	State rank of donor = state rank of recipient
LikC3	Donor's predicted mean must be within 5 percent of recipient's predicted mean
LikC4	If the recipient was a past year cigarette user, then donor must also be a past year cigarette user
LikC5	If the recipient was a past 3 years but not past year or lifetime but not past 3 years cigarette user, then donor's cigarette recency must match the recipient's cigarette recency
LikC6	If the recipient was a past 3 years but not past year cigarette user, then donor cannot have used cigarettes in the past year
LikC7	If the recipient was a lifetime but not past 3 years cigarette user, then donor cannot have used cigarettes in the past 3 years
LikC8	Age of donor must be greater than or equal to age of recipient

Table E.24 Constraints and Portion of the Predictive Mean Vector for Cigarette Ever Daily Used

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
1	Missing CIGDLYMO	11	None	1. CIG	1-5: 1 1,3-5: 2 1,4,5: 1	1-5: 3 1,3-5: 1	1-5: 2 1,3-5: 1

¹ The predictive mean vector components are defined by the following:

1. CIG = P(Daily use at some point in lifetime | respondent was a lifetime user but not a current daily user)

E.3.1.2 Drug Recency and Frequency

Many tables in this section abbreviate certain words. "Recency" is an abbreviation for "recency of use," "frequency" or "Freq." is an abbreviation for "frequency of use," and "30-day binge alcohol frequency" or "ALCBNG30D" is an abbreviation for the "number of days in the past 30 days when the respondent consumed five or more alcoholic drinks for males and four or more alcoholic drinks for females."

An empty cell within "Missingness Pattern" indicates that all information was available. An entry of "Missing" in a cell indicates that all information was missing. Other entries for missingness patterns provide the available information, indicating that this information was partially missing. For example, an entry of "Past year" indicates that the individual used the drug of interest in the past year, but the specific recency of past month versus past year not past month must be imputed. However, if the entry is shown in parentheses, then the specific recency is known and no imputation is required. For example, an entry of "(Past year)" indicates that the individual used the drug of interest in the past year and the specific recency of past month versus past year not past month is already known and does not require imputation.

Table E.25 Logical Constraints for Cigarette Recency and Frequency

Constraint #	Logical Constraint
LogC1	Donor must have used cigarettes within the past 3 years (a cigarette recency category of 1, 2, or 3)
LogC2	Donor cannot be a past month cigarette user (cigarette recency $\neq$ 1)
LogC3	Donor cannot be a past year cigarette user (cigarette recency $\neq$ 1 or 2)
LogC4	Donor must be a past year cigarette user (cigarette recency = 1 or 2)
LogC5	Donor must be a past month cigarette user (cigarette recency = 1)
LogC6	If recipient was never a daily user of cigarettes (CIGDLYMO = 2), then donor's 30-day cigarette frequency cannot equal 30
LogC7	If recipient's age at first use equals his or her current age, then donor's 30-day cigarette frequency (1) cannot be greater than the number of days between recipient's interview date and his or her date of first cigarette use (inclusive) and (2) cannot be greater than the number of days between recipient's interview date and his or her birthday (inclusive)
LogC8	Donor must be a past year but not past month cigarette user or a past 3 years but not past year cigarette user (cigarette recency = 2 or 3)

Table E.26 Likeness Constraints for Cigarette Recency and Frequency

Constraint #	Likeness Constraint
LikC1	State rank of donor = state rank of recipient
LikC2	Donor's predicted means each must be within 5 percent of recipient's predicted means

Table E.27 Constraints and Portion of the Predictive Mean Vector for Cigarette Recency and Frequency

#	Missingness Pattern ¹	Total Number of Cases	Logical Constraints	Predictive Mean Vector ²	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
1	Past year, 30-day frequency missing	7	4,6	1. $R1/(R1+R2)$ 2. $(R1*D)/(R1+R2)$ 3. $R1*(1-D)*PM/(R1+R2)$	1,2: 0 1: 6	1,2: 0 1: 1	No Cases
2	Recency and 30-day frequency missing	8	6	1. $R1$ 2. $R2$ 3. $R3$ 4. $R1*D$ 5. $R1*(1-D)*PM$	1,2: 1 1: 2	1,2: 0 1: 3	1,2: 1 1: 1
3	Past month, 30-day frequency missing	23	5-7	1. D 2. PM	1,2: 1 1: 3	1,2: 6	1,2: 13
4	Not past year	136	3,6	1. $R3/(R3+R4)$	1,2: 71 1: 1	1,2: 60	1,2: 4
5	Not past month	71	2,6	1. $R2/(R2+R3+R4)$ 2. $R3/(R2+R3+R4)$	1,2: 31 1: 12	1,2: 16	1,2: 12
6	Past year but not past month, or past 3 years but not past year	85	6,8	1. $R2/(R2+R3)$	1,2: 48	1,2: 33	1,2: 4
7	Past 3 years, 30-day frequency missing	3	1,6	1. $R1/(R1+R2+R3)$ 2. $R2/(R1+R2+R3)$ 3. $(R1*D)/(R1+R2+R3)$ 4. $R1*(1-D)*PM/(R1+R2+R3)$	1,2: 0 1: 3	No Cases	No Cases

¹ The response to CIGDLYMO, the edited response to the "ever daily used" question, technically could be used to subdivide each of the first three missingness patterns into two: one for respondents with CIGDLYMO = 2, and the other for respondents with CIGDLYMO ≠ 2. This was not done, because the benefit derived from this change would likely be insignificant. Respondents with CIGDLYMO = 2 technically have zero probability of being a daily user, so the predictive mean vectors could be simplified by setting D = 0. For example, the predictive mean vector for respondents in Missingness Pattern 2 with CIGDLYMO = 2 might look like this: (1) $R1$; (2) $R2$; (3) $R3$; (4) $R1*PM$.

² The predictive mean vector components are defined by the following:

1. $R1 = P(\text{Past month cigarette use} \mid \text{lifetime cigarette use})$
2. $R2 = P(\text{Past year but not past month cigarette use} \mid \text{lifetime cigarette use})$
3. $R3 = P(\text{Past 3 years but not past year cigarette use} \mid \text{lifetime cigarette use})$
4. $R4 = P(\text{Lifetime but not past 3 years cigarette use} \mid \text{lifetime cigarette use})$
5. $D = P(\text{Daily cigarette use} \mid \text{past month cigarette use})$
6. $PM = P(\text{Cigarette use on a given day in the past month} \mid \text{past month use})$

Table E.28 Logical Constraints for Cigar and Smokeless Tobacco Recency and Frequency

Constraint #	Logical Constraint
LogC1	Donor must have used within the past 3 years (a recency category of 1, 2, or 3)
LogC2	Donor cannot be a past month user (recency $\neq$ 1)
LogC3	Donor must be a past year user (recency = 1 or 2)
LogC4	Donor must be a past month user (recency = 1)
LogC5	Donor cannot be a past month or past year user (recency $\neq$ 1 or 2)
LogC6	If recipient's age at first use equals his or her current age, then donor's 30-day frequency (1) cannot be greater than the number of days between recipient's interview date and his or her date of first use (inclusive) and (2) cannot be greater than the number of days between recipient's interview date and his or her birthday (inclusive)
LogC7	Donor must be a past year but not past month user or a past 3 years but not past year user (recency = 2 or 3)

Table E.29 Likeness Constraints for Cigar and Smokeless Tobacco Recency and Frequency

Constraint #	Likeness Constraint
LikC1	State rank of donor = state rank of recipient
LikC2	Donor's predicted means each must be within 5 percent of recipient's predicted means

Table E.30 Constraints and Portion of the Predictive Mean Vector for Cigar Recency and Frequency

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
1	Past year, 30-day frequency missing	9	3	1. $R1/(R1+R2)$ 2. $R1*PM/(R1+R2)$	1,2: 1 1: 6	1,2: 2	No Cases
2	Recency and 30-day frequency missing	9	None	1. R1 2. R2 3. R3 4. $R1*PM$	1,2: 0 1: 1	1,2: 1 1: 1	1,2: 1 1: 5
3	Past month, 30-day frequency missing	6	4,6	1. PM	1,2: 2	1,2: 4	No Cases
4	Not past year	91	5	1. $R3/(R3+R4)$	1,2: 14 1: 4	1,2: 59 1: 1	1,2: 13
5	Not past month	52	2	1. $R2/(R2+R3+R4)$ 2. $R3/(R2+R3+R4)$	1,2: 16 1: 6	1,2: 18 1: 3	1,2: 8 1: 1
6	Past year but not past month, or past 3 years but not past year	76	7	1. $R2/(R2+R3)$	1,2: 31	1,2: 39	1,2: 6
7	Past 3 years, 30-day frequency missing	2	1	1. $R1/(R1+R2+R3)$ 2. $R2/(R1+R2+R3)$ 3. $(R1*PM)/(R1+R2+R3)$	No Cases	1,2: 1 1: 1	No Cases

¹ The predictive mean vector components are defined by the following:

1. $R1 = P(\text{Past month cigar use} \mid \text{lifetime cigar use})$
2. $R2 = P(\text{Past year but not past month cigar use} \mid \text{lifetime cigar use})$
3. $R3 = P(\text{Past 3 years but not past year cigar use} \mid \text{lifetime cigar use})$
4. $R4 = P(\text{Lifetime but not past 3 years cigar use} \mid \text{lifetime cigar use})$
5. $PM = P(\text{Cigar use on a given day in the past month} \mid \text{past month cigar use})$

Table E.31 Constraints and Portion of the Predictive Mean Vector for Smokeless Tobacco Recency and Frequency

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
1	Past year, 30-day frequency missing	4	3	1. $R1/(R1+R2)$ 2. $R1*PM/(R1+R2)$	1,2: 3 1: 1	No Cases	No Cases
2	Recency and 30-day frequency missing	14	None	1. $R1$ 2. $R2$ 3. $R3$ 4. $R1*PM$	1,2: 0 1: 6	1,2: 0 1: 1	1,2: 0 1: 7
3	Past month, 30-day frequency missing	8	4,6	1. PM	1,2: 4	No Cases	1,2: 4
4	Not past year	72	5	1. $R3/(R3+R4)$	1,2: 18 1: 5	1,2: 28 1: 1	1,2: 20
5	Not past month	45	2	1. $R2/(R2+R3+R4)$ 2. $R3/(R2+R3+R4)$	1,2: 12 1: 9	1,2: 10 1: 4	1,2: 4 1: 6
6	Past year but not past month, or past 3 years but not past year	51	7	1. $R2/(R2+R3)$	1,2: 22 1: 1	1,2: 17	1,2: 11
7	Past 3 years, 30-day frequency missing	0	1	1. $R1/(R1+R2+R3)$ 2. $R2/(R1+R2+R3)$ 3. $(R1*PM)/(R1+R2+R3)$	No Cases	No Cases	No Cases

¹ The predictive mean vector components are defined by the following:

1. $R1 = P(\text{Past month smokeless tobacco use} \mid \text{lifetime smokeless tobacco use})$
2. $R2 = P(\text{Past year but not past month smokeless tobacco use} \mid \text{lifetime smokeless tobacco use})$
3. $R3 = P(\text{Past 3 years but not past year smokeless tobacco use} \mid \text{lifetime smokeless tobacco use})$
4. $R4 = P(\text{Lifetime but not past 3 years smokeless tobacco use} \mid \text{lifetime smokeless tobacco use})$
5. $PM = P(\text{Smokeless tobacco use on a given day in the past month} \mid \text{past month smokeless tobacco use})$

Table E.32 Likeness Constraints for Pipe Recency and Frequency

Constraint #	Likeness Constraint
LikC1	State rank of donor = state rank of recipient
LikC2	Donor's predicted means each must be within 5 percent of recipient's predicted means

Table E.33 Constraints and Portion of the Predictive Mean Vector for Pipe Recency and Frequency

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
1	Missing	1	None	1. R1	No Cases	No Cases	1,2: 1

¹ The predictive mean vector components are defined by the following:

1. R1 = P(Past month pipe use | lifetime pipe use)

Table E.34 Logical Constraints for Various Drugs Recency and Frequency

Constraint #	Logical Constraint
LogC1	Donor's proportion of past year drug use * recipient's maximum number of days could have used drug in the past year must be less than or equal to recipient's maximum possible past year drug frequency of use Recipient's maximum possible drug frequency of use in the past year is limited by the following factors: (1) it must be less than or equal to the maximum period recipient could have used drug, as determined by the month of first drug use (2) if the maximum period recipient could have used drug is greater than 30, but recipient is a past month drug user with a nonmissing 30-day drug frequency, then the past year drug frequency must be less than or equal to the maximum period (minus the number of days recipient did not use drug in the past month) (3) if recipient is not a past month drug user, then the past year drug frequency must be less than or equal to the maximum period (minus 30)
LogC2	Donor's proportion of past year drug use * recipient's maximum number of days could have used drug in the past year must be greater than or equal to recipient's minimum possible past year drug frequency of use Recipient's minimum possible drug frequency of use in the past year is limited by the following factors: (1) if recipient is a past month drug user, it must be at least as much as the 30-day drug frequency (2) if recipient is not a past month drug user but is a past year drug user, it must be at least 1
LogC3	Donor's proportion of past year drug use * recipient's maximum number of days could have used drug in the past year must be greater than or equal to recipient's 30-day drug use
LogC4	Donor's 30-day drug use must be less than or equal to recipient's maximum number of days could have used drug in past 30 days
LogC5	Donor's 30-day drug use must be greater than or equal to recipient's minimum number of days could have used drug in past 30 days
LogC6	Donor's 30-day drug use must be greater than or equal to recipient's ALCBNG30D if nonmissing
LogC7	Donor's 30-day drug use must be greater than or equal to donor's proportion of past year drug use * recipient's maximum number of days could have used drug in past year [minus 335]
LogC8	Donor must be a past month drug user (drug recency = 1)

Table E.34 Logical Constraints for Various Drugs Recency and Frequency (continued)

Constraint #	Logical Constraint
LogC9	If recipient's age at first drug use equals his or her current age, then donor's 30-day drug frequency (1) cannot be greater than recipient's days between his or her interview date and date of first drug use (inclusive) and (2) cannot be greater than recipient's days between his or her interview date and birthday (inclusive)
LogC10	If recipient's age at first drug use equals his or her current age, then (1) donor's proportion of past year drug use * recipient's maximum number of days could have used drug in the past year cannot be greater than recipient's days between his or her interview date and date of first drug use (inclusive) and (2) donor's proportion of past year drug use * recipient's maximum number of days could have used drug in the past year cannot be greater than recipient's days between his or her interview date and birthday (inclusive)
LogC11	Recipient's estimated 30-day drug frequency is not given/legitimately skipped (estimated drug frequency $\neq$ 1-6)
LogC12	If recipient's age at first drug use equals his or her current age, then (1) donor's proportion of past year drug use * recipient's maximum number of days could have used drug in the past year cannot be greater than recipient's days between his or her interview date and date of first drug use (minus 30) and (2) donor's proportion of past year drug use * recipient's maximum number of days could have used drug in the past year cannot be greater than recipient's days between the interview date and birthday (minus 30)
LogC13	Donor must be a past year but not past month drug user (drug recency = 2)
LogC14	Donor's 30-day binge alcohol frequency (ALCBNG30D) value must be less than or equal to recipient's 30-day drug frequency
LogC15	Donor must have a 30-day binge alcohol frequency (ALCBNG30D) greater than zero if recipient is (1) female and reports usually having four or more drinks on each day she drank during the past 30 days or (2) male and reports usually having five or more drinks on each day he drank during the past 30 days
LogC16	Donor must be a past month or past year but not past month drug user (drug recency = 1 or 2)
LogC17	Donor's proportion of past year drug use * recipient's maximum number of days could have used drug in the past year must be greater than or equal to donor's 30-day drug frequency
LogC18	If recipient's month/year of first drug use data indicate that he or she must have used at least once in the interval (1 year before interview, 30 days before interview), then donor's proportion of past year drug use * recipient's maximum number of days could have used drug in the past year must be greater than recipient's 30-day drug frequency
LogC19	If recipient's month/year of first drug use data indicate that he or she must have used at least once in the interval (1 year before interview, 30 days before interview), then donor's proportion of past year drug use * recipient's maximum number of days could have used drug in the past year must be greater than recipient's ALCBNG30D value if nonmissing
LogC20	If recipient's month/year of first drug use data indicate that he or she must have used drug at least once in the interval (1 year before interview, 30 days before interview), and if donor is a past month drug user, then donor's proportion of past year drug use * recipient's maximum number of days could have used drug in the past year must be greater than donor's 30-day drug frequency

Table E.35 Likeness Constraints for Various Drugs Recency and Frequency

Constraint #	Likeness Constraint
LikC1	State rank of donor = state rank of recipient
LikC2	Donor's predicted means each must be within 5 percent of recipient's predicted means
LikC3	Donor's drug recency must match recipient's drug recency (when nonmissing) ¹

¹ Although this constraint also is used as a logical constraint for some missingness patterns, it is included for clarity.

Table E.36 Constraints and Portion of the Predictive Mean Vector for Alcohol Recency and Frequency

#	Missingness Pattern			Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
	Recency	12-Mo. Freq.	30-Day Freq.				12-17	18-25	26+
1	(Past month)	Missing	Missing	37	1,2,4-11,17,19,20	1. PM 2. PY	1-3: 6 1,3: 7	1-3: 4 1,3: 1	1-3: 18 1,3: 1
2	(Past month)		Missing	214	4-6,8,9,11	1. PM	1-3: 21 1,3: 3	1-3: 84 1,3: 5	1-3: 100 1,3: 1
3	(Past month)	Missing		253	1-3,8,10,18	1. PY	1-3: 44	1-3: 64 1,3: 1	1-3: 143 1,3: 1
4	(Past year but not past month)	Missing		139	1,2,12,13	1. PY	1-3: 57	1-3: 30	1-3: 52
5	Past year		Missing	324	4,5,9,11,16	1. R1/(R1+R2) 2. R1*PM/(R1+R2)	1-3: 45 1,3: 4	1-3: 79 1,3: 3	1-3: 193
6	Past year	Missing	Missing	101	1,2,4,5,7,9-12,16,17,20	1. R1/(R1+R2) 2. R1*PM/(R1+R2) 3. PY	1-3: 9 1,3: 11	1-3: 12 1,3: 11	1-3: 31 1,3: 27
7	Lifetime (known)	Missing	Missing	358	1,2,4,5,7,9-12,17	1. R1 2. R2 3. R1*PM 4. (R1+R2)*PY	1-3: 89 1,3: 153	1-3: 25 1,3: 68	1-3: 7 1,3: 16

¹ The predictive mean vector components are defined by the following:

1. R1 = P(Past month alcohol use | lifetime alcohol use)
2. R2 = P(Past year but not past month alcohol use | lifetime alcohol use)
3. PM = P(Alcohol use on a given day in the past month | past month alcohol use)
4. PY = P(Alcohol use on a given day in the past year | past year alcohol use)

Table E.37 Constraints and Portion of the Predictive Mean Vector for 30-Day Binge Alcohol Frequency

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group					
					Males			Females		
					12-17	18-25	26+	12-17	18-25	26+
1	Missing	948	14,15	1. PMB	1,2: 69 1: 4	1,2: 156	1,2: 248	1,2: 84 1: 4	1,2: 148	1,2: 235

¹ The predictive mean vector components are defined by the following:

1. PMB = P(Past month binge alcohol use | past month alcohol use)

Table E.38 Constraints and Portion of the Predictive Mean Vector for Marijuana Recency and Frequency

#	Missingness Pattern			Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
							12-17	18-25	26+
	Recency	12-Month Freq.	30-Day Freq.						
1	(Past month)	Missing	Missing	9	1,2,4,5,7-11, 17,20	1. PM 2. PY	1-3: 1 1,3: 1	1-3: 2	1-3: 4 1,3: 1
2	(Past month)		Missing	10	4,5,8,9,11	1. PM	1-3: 1 1,3: 2	1-3: 7	No Cases
3	(Past month)	Missing		93	1,2,3,8,10,18	1. PY	1-3: 26 1,3: 1	1-3: 39	1-3: 27
4	(Past year but not past month)	Missing		52	1,2,12,13	1. PY	1-3: 22 1,3: 1	1-3: 15	1-3: 14
5	Past year		Missing	86	4,5,9,11,16	1. R1/(R1+R2) 2. R1*PM/(R1+R2)	1-3: 10 1,3: 5	1-3: 23 1,3: 6	1-3: 38 1,3: 4
6	Past year	Missing	Missing	97	1,2,4,5,7,9-12, 16,17,20	1. R1/(R1+R2) 2. R1*PM/(R1+R2) 3. PY	1-3: 0 1,3: 22	1-3: 12 1,3: 42	1-3: 3 1,3: 18
7	Missing	Missing	Missing	208	1,2,4,5,7,9-12, 17	1. R1 2. R2 3. R1*PM 4. (R1+R2)*PY	1-3: 12 1,3: 86	1-3: 4 1,3: 56	1-3: 2 1,3: 48

¹ The predictive mean vector components are defined by the following:

1. R1 = P(Past month marijuana use | lifetime marijuana use)
2. R2 = P(Past year but not past month marijuana use | lifetime marijuana use)
3. PM = P(Marijuana use on a given day in the past month | past month marijuana use)
4. PY = P(Marijuana use on a given day in the past year | past year marijuana use)

Table E.39 Constraints and Portion of the Predictive Mean Vector for Heroin Recency and Frequency

#	Missingness Pattern			Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
							12-17	18-25	26+
	Recency	12-Month Freq.	30-Day Freq.						
1	(Past month)	Missing	Missing	0	1,2,4,5,7-11,17,20	1. PM 2. PY	No Cases	No Cases	No Cases
2	(Past month)		Missing	0	4,5,8,9,11	1. PM	No Cases	No Cases	No Cases
3	(Past month)	Missing		2	1-3,8,10,18	1. PY	No Cases	1-3: 0 1,3: 2	No Cases
4	(Past year but not past month)	Missing		2	1,2,12,13	1. PY	No Cases	1-3: 1	1-3: 1
5	Past year		Missing	2	4,5,9,11,16	1. R1/(R1+R2) 2. R1*PM/(R1+R2)	No Cases	No Cases	1-3: 0 1,3: 2
6	Past year	Missing	Missing	15	1,2,4,5,7,9-12,16,17,20	1. R1/(R1+R2) 2. R1*PM/(R1+R2) 3. PY	No Cases	1-3: 0 1,3: 6	1-3: 0 1,3: 9
7	Missing	Missing	Missing	11	1,2,4,5,7,9-12,17	1. R1 2. R2 3. R1*PM 4. (R1+R2)*PY	1-3: 0 1,3: 1	1-3: 0 1,3: 2	1-3: 0 1,3: 8

¹ The predictive mean vector components are defined by the following:

1. R1 = P(Past month heroin use | lifetime heroin use)
2. R2 = P(Past year but not past month heroin use | lifetime heroin use)
3. PM = P(Heroin use on a given day in the past month | past month heroin use)
4. PY = P(Heroin use on a given day in the past year | past year heroin use)

Table E.40 Logical Constraints for Cocaine Recency and Frequency

Constraint #	Logical Constraint
LogC1	Donor must be a past month cocaine user (cocaine recency = 1)
LogC2	Donor's proportion of past year cocaine use * recipient's maximum number of days could have used cocaine in the past year must be less than or equal to recipient's maximum possible past year cocaine frequency of use Recipient's maximum possible cocaine frequency of use in the past year is limited by the following factors: (1) it must be less than or equal to the maximum period recipient could have used cocaine, as determined by the month of first cocaine use (2) if the maximum period recipient could have used cocaine is greater than 30, but recipient is a past month cocaine user with a nonmissing 30-day cocaine frequency, then the past year cocaine frequency must be less than or equal to the maximum period (the number of days recipient did not use cocaine in the past month) (3) if recipient is not a past month cocaine user, then the past year cocaine frequency must be less than or equal to the maximum period (minus 30)
LogC3	Donor's proportion of past year cocaine use * recipient's minimum number of days could have used cocaine in the past year must be greater than or equal to recipient's minimum possible past year cocaine frequency of use Recipient's minimum possible cocaine frequency of use in the past year is limited by the following factors: (1) if recipient is a past month cocaine user, it must be at least as much as the 30-day cocaine frequency (2) if recipient is not a past month cocaine user but is a past year cocaine user, it must be at least 1 (3) if recipient has a valid value for the 12-month crack frequency, it must be at least as much as the recipient's 12-month crack frequency
LogC4	If recipient's age at first cocaine use equals his or her current age, and (1) donor is a past month cocaine user, then donor's proportion of past year cocaine use * recipient's maximum number of days could have used cocaine in the past year must be less than or equal to the number of days between recipient's interview date and birthday (inclusive) (2) donor is a past month cocaine user, then donor's proportion of past year cocaine use * recipient's maximum number of days could have used cocaine in the past year must be less than or equal to the number of days between recipient's interview date and date of first drug use (inclusive) (3) donor is a past year but not past month cocaine user, then donor's proportion of past year cocaine use * recipient's maximum number of days could have used cocaine in the past year must be less than or equal to the number of days between recipient's interview date and birthday (minus 30) (4) donor is a past year but not past month cocaine user, then donor's proportion of past year cocaine use * recipient's maximum number of days could have used cocaine in the past year must be less than or equal to the number of days between recipient's interview date and date of first drug use (minus 30)
LogC5	If recipient's age at first cocaine use equals his or her current age, then (1) donor's 30-day cocaine use must be less than or equal to the number of days between recipient's interview date and birthday (inclusive) (2) donor's 30-day cocaine use must be less than or equal to the number of days between recipient's interview date and date of first drug use (inclusive)

Table E.40 Logical Constraints for Cocaine Recency and Frequency (continued)

Constraint #	Logical Constraint
LogC6	Donor's 30-day cocaine use must be less than or equal to recipient's maximum number of days could have used cocaine in past 30 days
LogC7	Donor's 30-day cocaine use must be greater than or equal to recipient's minimum number of days could have used cocaine in past 30 days
LogC8	Donor's 30-day cocaine use must be greater than or equal to donor's proportion of past year drug use * recipient's maximum number of days could have used drug in past year (minus 335)
LogC9	Recipient's estimated cocaine 30-day frequency is not given/legitimately skipped (estimated cocaine frequency $\neq$ 1-6)
LogC10	Donor must be a past month or past year but not past month cocaine user (cocaine recency = 1 or 2)
LogC11	Donor's proportion of past year crack use * recipient's maximum number of days could have used crack in the past year must be less than or equal to recipient's maximum possible past year crack frequency of use Recipient's maximum possible crack frequency of use in the past year is limited by the following factors: (1) it must be less than or equal to the maximum period recipient could have used crack, as determined by the month of first crack use (2) if the maximum period recipient could have used crack is greater than 30, but recipient is a past month crack user with a nonmissing 30-day crack frequency, then the past year crack frequency must be less than or equal to the maximum period (the number of days recipient did not use in the past month) (3) if recipient is not a past month crack user, then the past year crack frequency must be less than or equal to the maximum period (minus 30) (4) if recipient has a valid value for the 12-month cocaine frequency, it must be less than or equal to the recipient's 12-month cocaine frequency
LogC12	Donor's proportion of past year crack use * recipient's minimum number of days could have used crack in the past year must be greater than or equal to recipient's minimum possible past year crack frequency of use Recipient's minimum possible crack frequency of use in the past year is limited by the following factors: (1) if recipient is a past month crack user, it must be at least as much as the 30-day crack frequency (2) if recipient is not a past month crack user but is a past year crack user, it must be at least 1

Table E.40 Logical Constraints for Cocaine Recency and Frequency (continued)

Constraint #	Logical Constraint
LogC13	If recipient's age at first crack use equals his or her current age, and (1) donor is a past month crack user, then donor's proportion of past year crack use * recipient's maximum number of days could have used cocaine in the past year must be less than or equal to the number of days between recipient's interview date and birthday (inclusive) (2) donor is a past month crack user, then donor's proportion of past year crack use * recipient's maximum number of days could have used cocaine in the past year must be less than or equal to the number of days between recipient's interview date and date of first drug use (inclusive) (1) donor is a past year but not past month crack user, then donor's proportion of past year crack use * recipient's maximum number of days could have used cocaine in the past year must be less than or equal to the number of days between recipient's interview date and birthday (minus 30) (2) donor is a past year but not past month crack user, then donor's proportion of past year crack use * recipient's maximum number of days could have used cocaine in the past year must be less than or equal to the number of days between recipient's interview date and date of first drug use (minus 30)
LogC14	Donor's 30-day crack use must be less than or equal to the number of days between recipient's interview date and birthday (inclusive)
LogC15	Donor's 30-day crack use must be less than or equal to recipient's maximum number of days could have used crack in past 30 days
LogC16	Donor's 30-day crack use must be greater than or equal to donor's proportion of past year drug use * recipient's maximum number of days could have used drug in past year (minus 335)
LogC17	Donor's 30-day crack use must be greater than or equal to recipient's minimum number of days could have used crack in past 30 days
LogC18	Recipient's estimated 30-day crack frequency is not given/legitimately skipped (estimated crack frequency $\neq$ 1-6)
LogC19	Donor must be a past month crack user (crack recency = 1)
LogC20	Donor must be a past month or past year not past month crack user (crack recency = 1, 2)
LogC21	Donor must be a crack user (crack recency $\neq$ 91)
LogC22	If recipient's month/year of first use data for cocaine indicate that he or she must have used cocaine at least once in the interval (1 year before interview, 30 days before interview), and recipient is a past month user, then donor's proportion of past year cocaine use * recipient's maximum number of days could have used cocaine in the past year must be greater than recipient's 30-day cocaine frequency
LogC23	If recipient's month/year of first use data for cocaine indicate that he or she must have used cocaine at least once in the interval (1 year before interview, 30 days before interview), and donor is a past month cocaine user, then donor's proportion of past year cocaine use * recipient's maximum number of days could have used cocaine in the past year must be greater than donor's 30-day cocaine frequency
LogC24	If recipient's month/year of first use data for crack indicate that he or she must have used crack at least once in the interval (1 year before interview, 30 days before interview), and recipient is a past month user, then donor's proportion of past year crack use * recipient's maximum number of days could have used crack in the past year must be greater than recipient's 30-day crack frequency

Table E.40 Logical Constraints for Cocaine Recency and Frequency (continued)

Constraint #	Logical Constraint
LogC25	If recipient's month/year of first use data for crack indicate that he or she must have used crack at least once in the interval (1 year before interview, 30 days before interview), and donor is a past month crack user, then donor's proportion of past year crack use * recipient's maximum number of days could have used in the past year must be greater than donor's 30-day crack frequency
LogC26	Donor's proportion of past year crack use * recipient's maximum number of days could have used crack in the past year must be less than or equal to donor's proportion of past year cocaine use * recipient's maximum number of days could have used cocaine in the past year
LogC27	If recipient is a past month cocaine user or cocaine recency is missing, donor is a past month cocaine user, and donor is a past year not past month crack user, then donor's proportion of days used crack in the last year * recipient's maximum period used crack in the past year cannot be larger than recipient's 12-month frequency minus donor's 30-day frequency

Table E.41 Likeness Constraints for Cocaine Recency and Frequency

Constraint #	Likeness Constraint
LikC1	State rank of donor = state rank of recipient
LikC2	Donor's predicted means each must be within 5 percent of recipient's predicted means
LikC3	Donor's cocaine recency must match recipient's drug recency (when nonmissing) ¹
LikC4	Donor's crack recency agrees with recipient's crack recency (when nonmissing)

¹ Although this constraint also is used as a logical constraint for some missingness patterns, it is included for clarity.

Table E.42 Constraints and Portion of the Predictive Mean Vector for Cocaine Recency and Frequency

#	Missingness Pattern						Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
	Cocaine Recency	Crack Recency	Cocaine 12-Mo. Freq.	Crack 12-Mo. Freq.	Cocaine 30-Day Freq.	Crack 30-Day Freq.				12-17	18-25	26+
1	(Past year)		Missing				34	2-4,10,22	1. PY	1-4: 3 1-3: 0 1,3: 1 3: 1	1-4: 7 1-3: 6 1,3: 2	1-4: 5 1-3: 2 1,3: 7
2	(Past year)	(Past year)		Missing			1	11-13,20,24	1. PY	No Cases	No Cases	1-4: 0 1-3: 0 1,3: 1
3	(Past year)	(Past year)	Missing	Missing			1	2-4,10-13,20,22,24,26	1. PY	No Cases	No Cases	1-4: 0 1-3: 0 1,3: 1
4	(Past month)				Missing		11	1,5-7,9	1. PM	No Cases	1-4: 0 1-3: 3 1,3: 2 3: 0 None: 0 None*: 1	1-4: 2 1-3: 2 1,3: 1
5	(Past month)	(Past month)				Missing	0	14-16,18,19	1. PM	No Cases	No Cases	No Cases
6	(Past month)	(Past month)			Missing	Missing	0	1,5-7,9,14-16,18,19	1. PM	No Cases	No Cases	No Cases
7	(Past month)		Missing		Missing		2	1-9,23	1. PM 2. PY	No Cases	1-4: 1	1-4: 0 1-3: 0 1,3: 1
8	(Past month)	(Past year)		Missing	Missing		0	1,5-7,9,11-13,20,24	1. PM 2. PY	No Cases	No Cases	No Cases
9	(Past month)	(Past year)	Missing	Missing	Missing		0	1-9,11-13,20,23,24,26	1. PM 2. PY	No Cases	No Cases	No Cases
10	(Past month)	(Past month)	Missing			Missing	0	2-4,14-16,18,19,22	1. PM 2. PY	No Cases	No Cases	No Cases

Table E.42 Constraints and Portion of the Predictive Mean Vector for Cocaine Recency and Frequency (continued)

#	Missingness Pattern						Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
	Cocaine Recency	Crack Recency	Cocaine 12-Mo. Freq.	Crack 12-Mo. Freq.	Cocaine 30-Day Freq.	Crack 30-Day Freq.				12-17	18-25	26+
11	(Past month)	(Past month)		Missing		Missing	0	11-19,25	1. PM 2. PY	No Cases	No Cases	No Cases
12	(Past month)	(Past month)	Missing	Missing		Missing	0	2-4,11-19,22,25,26	1. PM 2. PY	No Cases	No Cases	No Cases
13	(Past month)	(Past month)	Missing		Missing	Missing	0	1-9,14-16,18,19,23	1. PM 2. PY	No Cases	No Cases	No Cases
14	(Past month)	(Past month)		Missing	Missing	Missing	0	1,5-7,9,11-19,25	1. PM 2. PY	No Cases	No Cases	No Cases
15	(Past month)	(Past month)	Missing	Missing	Missing	Missing	1	1-9,11-19,23,25,26	1. PM 2. PY	No Cases	No Cases	1-4: 0 1-3: 0 1,3: 1
16	(Past month)	Past year				Missing	0	14-16,18,20	1. R1/(R1+R2) 2. R1*PM/(R1+R2)	No Cases	No Cases	No Cases
17	(Past month)	Past year			Missing	Missing	0	1,5-7,9,14-16,18,20	1. R1/(R1+R2) 2. R1*PM/(R1+R2)	No Cases	No Cases	No Cases
18	(Past month)	Past year		Missing		Missing	1	11-18,20,25	1. R1/(R1+R2) 2. R1*PM/(R1+R2) 3. PY	No Cases	1-4: 0 1-3: 0 1,3: 1	No Cases
19	(Past month)	Past year		Missing	Missing	Missing	0	1,5-7,9,11-18,20,25,27	1. R1/(R1+R2) 2. R1*PM/(R1+R2) 3. PY	No Cases	No Cases	No Cases
20	(Past month)	Missing		Missing		Missing	0	11-18,21	1. R1 2. R2 3. R1*PM 4. (R1+R2)*PY	No Cases	No Cases	No Cases
21	(Past month)	Missing		Missing	Missing	Missing	0	1,5-7,9,11-18,21,27	1. R1 2. R2 3. R1*PM 4. (R1+R2)*PY	No Cases	No Cases	No Cases

Table E.42 Constraints and Portion of the Predictive Mean Vector for Cocaine Recency and Frequency (continued)

#	Missingness Pattern						Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
	Cocaine Recency	Crack Recency	Cocaine 12-Mo. Freq.	Crack 12-Mo. Freq.	Cocaine 30-Day Freq.	Crack 30-Day Freq.				12-17	18-25	26+
22	(Past month)	Missing	Missing	Missing	Missing	Missing	0	1-9,11-18,21,23,26	1. R1 2. R2 3. R1*PM 4. (R1+R2)*PY	No Cases	No Cases	No Cases
23	Missing		Missing		Missing		73	2-9	1. R1 2. R2 3. R1*PM 4. (R1+R2)*PY	1-4: 0 1-3: 0 1,3: 10	1-4: 0 1-3: 0 1,3: 46	1-4: 1 1-3: 0 1,3: 16
24	Missing	Missing	Missing	Missing	Missing	Missing	6	2-9,11-18,21,26	1. R1 2. R2 3. R1*PM 4. (R1+R2)*PY	No Cases	1-4: 0 1-3: 0 1,3: 1	1-4: 0 1-3: 0 1,3: 5
25	Past year				Missing		12	5-7,9,10	1. R1/(R1+R2) 2. R1*PM/(R1+R2)	No Cases	1-4: 2 1-3: 0 1,3: 4	1-4: 3 1-3: 0 1,3: 3
26	Past year	Past year			Missing	Missing	1	5-7,9,10,14-16,18,20	1. R1/(R1+R2) 2. R1*PM/(R1+R2)	1-4: 0 1-3: 0 1,3: 1	No Cases	No Cases
27	Past year		Missing		Missing		7	2-10,23	1. R1/(R1+R2) 2. R1*PM/(R1+R2) 3. PY	1-4: 0 1-3: 0 1,3: 1	1-4: 0 1-3: 0 1,3: 4	1-4: 0 1-3: 0 1,3: 2
28	Past year	Past year	Missing		Missing	Missing	1	2-10,14-16,18,20,23	1. R1/(R1+R2) 2. R1*PM/(R1+R2) 3. PY	No Cases	1-4: 0 1-3: 0 1,3: 1	No Cases
29	Past year	Past year		Missing	Missing	Missing	2	5-7,9-18,20,25,27	1. R1/(R1+R2) 2. R1*PM/(R1+R2) 3. PY	1-4: 0 1-3: 0 1,3: 0 3: 1	1-4: 0 1-3: 0 1,3: 0 3: 1	No Cases

Table E.42 Constraints and Portion of the Predictive Mean Vector for Cocaine Recency and Frequency (continued)

#	Missingness Pattern						Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
	Cocaine Recency	Crack Recency	Cocaine 12-Mo. Freq.	Crack 12-Mo. Freq.	Cocaine 30-Day Freq.	Crack 30-Day Freq.				12-17	18-25	26+
30	Past year	Past year	Missing	Missing	Missing	Missing	0	2-18,20,23,25,26	1. R1/(R1+R2) 2. R1*PM/(R1+R2) 3. PY	No Cases	No Cases	No Cases
31	Past year	Missing		Missing	Missing	Missing	0	5-7,9-18,21,27	1. R1/(R1+R2) 2. R1*PM/(R1+R2) 3. PY	No Cases	No Cases	No Cases
32	Past year	Missing	Missing	Missing	Missing	Missing	0	2-18,21,23,26	1. R1/(R1+R2) 2. R1*PM/(R1+R2) 3. PY	No Cases	No Cases	No Cases
33	(Past month)	Missing	Missing	Missing		Missing	0	2-4,10-18,21,22,26	1. R1 2. R2 3. R1*PM 4. (R1+R2)*PY	No Cases	No Cases	No Cases
34	Past year	(Past year but not past month)	Missing	Missing	Missing		1	2-13,20,23,26	1. R1/(R1+R2) 2. R1*PM/(R1+R2) 3. PY	No Cases	1-4: 0 1-3: 0 1,3: 1	No Cases

*A donor could not be found for these cases. Imputed values were randomly assigned within the bounds created during the editing process.

NOTE: Cocaine users included crack users and cocaine users who were not crack users.

¹ The predictive mean vector components are defined by the following:

1. R1 = P(Past month cocaine use | lifetime cocaine use)
2. R2 = P(Past year but not past month cocaine use | lifetime cocaine use)
3. PM = P(Cocaine use on a given day in the past month | past month use of cocaine)
4. PY = P(Cocaine use on a given day in the past year | past year use of cocaine)

E.3.1.3 Age at First Use

Table E.43 Logical Constraints for Age at First Use, Univariate Assignment

Constraint #	Logical Constraint
LogC1	Donor's age at first drug use must be less than or equal to recipient's age
LogC2	If recipient was not a past year drug user, then donor cannot have age at first drug use equal to recipient's age
LogC3	If recipient was not a past 3 years drug user, then donor cannot have age at first drug use equal to recipient's age, recipient's age minus 1, or recipient's age minus 2
LogC4	If recipient's 12-month drug frequency was greater than the number of days since recipient's most recent birthday, then donor cannot have age at first drug use equal to recipient's age
LogC5	If recipient's 30-day drug frequency was greater than the number of days since recipient's most recent birthday, then donor cannot have age at first drug use equal to recipient's age
LogC6	If recipient's most recent birthday was within the past 30 days, and recipient's drug recency was past year but not past month, then donor cannot have age at first drug use equal to recipient's age
LogC7	If recipient's drug recency was past year but not past month, and the difference between the number of days since recipient's most recent birthday and recipient's 12-month drug frequency was less than 30, then donor cannot have age at first drug use equal to recipient's age. (These recipients are missed by LogC3 and LogC4, but the idea is the same: the date of first drug use must be earlier than recipient's most recent birthday.)
LogC8	Donor's age at first cigarette use must be less than or equal to recipient's age at first daily cigarette use (if existing)
LogC9	Donor's age at first daily cigarette use must be greater than or equal to recipient's imputation-revised age at first cigarette use

Table E.44 Likeness Constraints for Age at First Use, Univariate Assignment

Constraint #	Likeness Constraint
LikC1	State rank of donor = state rank of recipient
LikC2	Donor's predicted mean must be within 5 percent of recipient's predicted mean
LikC3	If recipient was a past year drug user, then donor was a past year drug user; if recipient was a past 3 years but not past year drug user, then donor was the same; if recipient was a lifetime but not past 3 years drug user, then donor was the same
LikC4	Age of donor = age of recipient
LikC5	Age of donor must be greater than or equal to age of recipient
LikC6	If recipient was a past year drug user, then donor was a past year drug user; if recipient was not a past year drug user, then donor was the same
LikC7	If recipient was a past 3 years but not past year drug user, then donor was not a past year drug user; if recipient was a lifetime but not past 3 years drug user, then donor was the same
LikC8	If recipient was not a past year drug user, then donor was the same

Table E.45 Constraints and Portion of the Predictive Mean Vector for Cigarette Age at First Use

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
1	Missing	426	1-3,5,6,8	1. PrAFU	1-4: 136 2-4: 3 3,4: 3	1-4: 112	1-4: 165 2-4: 5 3,4: 0 4,7: 1 5,7: 1

¹ The predictive mean vector components are defined by the following:

1. PrAFU = Predicted cigarette age at first use

Table E.46 Constraints and Portion of the Predictive Mean Vector for Cigarette Age at First Daily Use

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
1	Missing	116	1-3,5,6,9	1. PrCigAFU	1-4: 5	1-4: 21 2-4: 2	1-4: 80 2-4: 5 3,4: 2 4,7: 0 5,7: 1

¹ The predictive mean vector components are defined by the following:

1. PrCigAFU = Predicted cigarette age at first daily use

Table E.47 Constraints and Portion of the Predictive Mean Vector for Cigar Age at First Use

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
1	Missing	387	1-3,5,6	1. PrAFU	1-4: 54 2-4: 2 3,4: 3	1-4: 109	1-4: 194 2-4: 14 3,4: 9 4,7: 2

¹ The predictive mean vector components are defined by the following:

1. PrAFU = Predicted cigar age at first use

Table E.48 Constraints and Portion of the Predictive Mean Vector for Smokeless Tobacco Age at First Use

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
1	Missing	216	1-3,5,6	1. PrAFU	1-4: 58 2-4: 6 3,4: 4	1-4: 56 2-4: 2	1-4: 64 2-4: 13 3,4: 13

¹ The predictive mean vector components are defined by the following:

1. PrAFU = Predicted smokeless tobacco age at first use

Table E.49 Constraints and Portion of the Predictive Mean Vector for Alcohol Age at First Use

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
1	Missing	703	1,2,4,6,7	1. PrAFU	1,2,4,6: 251 2,4,6: 2 4,6: 2	1,2,4,6: 126	1,2,4,6: 306 2,4,6: 10 4,6: 5 4,8: 0 5,8: 0 8: 1

¹ The predictive mean vector components are defined by the following:

1. PrAFU = Predicted alcohol age at first use

Table E.50 Constraints and Portion of the Predictive Mean Vector for Marijuana Age at First Use

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
1	Missing	234	1,2,4,6,7	1. PrAFU	1,2,4,6: 68 2,4,6: 2	1,2,4,6: 55	1,2,4,6: 100 2,4,6: 4 4,6: 4 4,8: 0 5,8: 1

¹ The predictive mean vector components are defined by the following:

1. PrAFU = Predicted marijuana age at first use

Table E.51 Constraints and Portion of the Predictive Mean Vector for Heroin Age at First Use

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
1	Missing	11	1,2,4,6,7	1. PrAFU	1,2,4,6: 0 2,4,6: 1 4,6: 1	1,2,4,6: 1	1,2,4,6: 2 2,4,6: 3 4,6: 2 4,8: 0 5,8: 1

¹ The predictive mean vector components are defined by the following:

1. PrAFU = Predicted heroin age at first use

Table E.52 Logical Constraints for Age at First Use, Multivariate Assignment

Constraint #	Logical Constraint
LogC1	For each parent or child drug for which recipient was missing age at first use, donor's age at first use must be less than or equal to recipient's age
LogC2	For each parent or child drug for which recipient was missing age at first use, if recipient was not a past year user, then donor cannot have age at first use equal to recipient's age
LogC3	For each child drug for which recipient was missing age at first use, donor must be a lifetime user
LogC4	If recipient was missing the age at first use for the parent drug, and recipient's parent drug 12-month frequency was greater than the number of days since recipient's most recent birthday, then donor cannot have parent drug age at first use equal to recipient's age

Table E.52 Logical Constraints for Age at First Use, Multivariate Assignment (continued)

Constraint #	Logical Constraint
LogC5	For each child drug for which recipient was missing age at first use, if recipient's 12-month frequency was greater than the number of days since recipient's most recent birthday, then donor cannot have age at first use equal to recipient's age
LogC6	If recipient was missing the age at first use for the parent drug, then donor must have parent drug age at first use less than or equal to the recipient's child drug(s) age(s) at first use (if existing)
LogC7	For each parent or child drug for which recipient was missing age at first use, if recipient's most recent birthday was within the past 30 days, and recipient's recency was past year but not past month, then donor cannot have age at first use equal to recipient's age
LogC8	If recipient was missing the age at first use for the parent drug, and recipient's parent drug recency was past year but not past month, and the difference between the number of days since recipient's most recent birthday and recipient's parent drug 12-month frequency was less than 30, then donor cannot have parent drug age at first use equal to recipient's age. (These recipients are missed by LogC3 and LogC4, but the idea is the same: the date of first use must be earlier than recipient's most recent birthday.)
LogC9	For each child drug for which recipient was missing age at first use, if recipient's recency was past year but not past month, and the difference between the number of days since recipient's most recent birthday and recipient's 12-month frequency was less than 30, then donor cannot have age at first use equal to recipient's age. (These recipients are missed by LogC3 and LogC4, but the idea is the same: the date of first use must be earlier than recipient's most recent birthday.)
LogC10	For each child drug for which recipient was missing age at first use, donor must have age at first use greater than or equal to recipient's parent drug age at first use (if existing)

Table E.53 Likeness Constraints for Age at First Use, Multivariate Assignment

Constraint #	Likeness Constraint
LikC1	State rank of donor = state rank of recipient
LikC2	Donor's predicted mean must be within 5 percent of recipient's predicted mean
LikC3	Age of donor = age of recipient
LikC4	Donor agrees with recipient with respect to lifetime use for child drug(s)
LikC5	For the parent drug, if recipient was a past year user, then donor was a past year user
LikC6	For the parent drug, if recipient was a lifetime but not past year user, then donor was a lifetime but not past year user
LikC7	For each child drug, if recipient was a past year user, then donor was a past year user
LikC8	For each child drug, if recipient was a lifetime but not past year user, then donor was not a past year user
LikC9	Donor was at least as old as recipient, but no more than 20 years older than recipient
LikC10	Donor was no more than 20 years older than recipient

Table E.54 Constraints and Portion of the Predictive Mean Vector for Cocaine Age at First Use

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
1	Missing	68	1-10	1. PrCocAFU	1-8: 3 2-8: 1 2,3,5-8: 0 3,5-8: 3	1-8: 23 2-8: 1	1-8: 24 2-8: 4 2,3,5-8: 1 3,5-8: 6 3,6,8: 2

¹ The predictive mean vector components are defined by the following:

1. PrCocAFU = Predicted cocaine age at first use

E.3.2 modPMN

E.3.2.1 Lifetime Drug Use

Table E.55 Imputation Details for Inhalants Lifetime

Edited Lifetime Use	Number of Cases	Predictive Mean Vector ¹
Completely missing	175	1. LF1 2. LF2
Total Cases Requiring Imputation	175	

¹ The predictive mean vector components are defined by the following:

1. LF1 = P(Lifetime use of inhalants)
2. LF2 = P(Never used inhalants)

Table E.56 Imputation Details for Methamphetamine Lifetime

Edited Lifetime Use	Number of Cases	Predictive Mean Vector ¹
Completely missing	87	1. LF1 2. LF2
Total Cases Requiring Imputation	87	

¹ The predictive mean vector components are defined by the following:

1. LF1 = P(Lifetime use of methamphetamine)
2. LF2 = P(Never used methamphetamine)

Table E.57 Imputation Details for Hallucinogens Lifetime

Edited Lifetime Use	Number of Cases	Predictive Mean Vector ¹
Completely missing	290	1. LF1 2. LF2
Total Cases Requiring Imputation	290	

¹ The predictive mean vector components are defined by the following:

1. LF1 = P(Lifetime use of hallucinogens)
2. LF2 = P(Never used hallucinogens)

There were a large number of missingness patterns for lifetime hallucinogen child drug use. The usable case rule (Section 2.2.3) required a certain minimum response pattern, but other than that any combination of the lifetime hallucinogen child drug variables may have been missing for a given respondent record.

The probabilities associated with the seven lifetime hallucinogen child drug variables listed in Table E.58 formed the full predictive mean vector. Only the probabilities associated with the lifetime hallucinogen child drug variables for which values were missing were used in the predictive mean vector for each item nonrespondent. Although lifetime hallucinogen child drug variables were imputed using PMN methodology, they have been included in this section to allow proximity to the parent measure, hallucinogens lifetime.

Table E.58 Elements of Full Predictive Mean Vector for Lifetime Hallucinogen Child Drug Use

Lifetime Drug Use	Predicted Mean
LSD Lifetime	P(Lifetime User Lifetime User of Hallucinogens)
PCP Lifetime	P(Lifetime User Lifetime User of Hallucinogens)
Ecstasy Lifetime	P(Lifetime User Lifetime User of Hallucinogens)
Ketamine Lifetime	P(Lifetime User Lifetime User of Hallucinogens)
DMT/AMT/Foxy Lifetime	P(Lifetime User Lifetime User of Hallucinogens)
<i>Salvia divinorum</i> Lifetime	P(Lifetime User Lifetime User of Hallucinogens)
Other Hallucinogens Lifetime ¹	P(Lifetime User Lifetime User of Hallucinogens)

¹ Other hallucinogens for lifetime directly include peyote, mescaline, and psilocybin and indirectly include any other hallucinogens not previously included in the lifetime predictive mean vector.

Table E.59 Logical Constraints for Hallucinogen Child Lifetime

Constraint #	Logical Constraint
LogC1	Donor must be a lifetime user of at least one hallucinogen child drug that the recipient is missing if recipient is a nonuser of all nonmissing drugs or is missing all child drugs

Table E.60 Likeness Constraints for Hallucinogen Child Lifetime

Constraint #	Likeness Constraint
LikC1	State rank of donor = state rank of recipient
LikC2	Donor's predicted means each must be within 5 percent of recipient's predicted means
LikC3	Lifetime use of donor = lifetime use of recipient, for each nonmissing lifetime indicator

Table E.61 Constraints and Portion of the Predictive Mean Vector for Hallucinogen Child Lifetime

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26+
1	Missing one or more of the lifetime drug use variables	71	1	1. LF for gate question associated with each missing variable	1-3: 4 1,2: 4 1: 2	1-3: 13 1,2: 2 1: 1	1-3: 28 1,2: 7 1: 10

¹ The predictive mean vector components are defined by the following:

1. LF = P(Lifetime user of x | lifetime user of hallucinogens), where x is the child drug associated with the missing gate question

E.3.2.2 Drug Recency and Frequency

Many tables in this section abbreviate certain words. "Recency" is an abbreviation for "recency of use," and "frequency" or "Freq." is an abbreviation for "frequency of use."

An entry of "Missing" within "Missingness Pattern" indicates that all information was missing. Other entries for missingness patterns provide the available information, indicating that this information was partially missing. For example, an entry of "Past year" indicates that the individual used the drug of interest in the past year, but the specific recency of past month versus past year not past month must be imputed.

E.3.2.2.1 Recency of Any Use

Table E.62 Imputation Details for Sedatives Recency of Any Use

Edited Recency of Any Use	Number of Cases	Predictive Mean Vector ¹
Lifetime use	18	1. RU1/(RU1+RU2) 2. RU2/(RU1+RU2)
Not past year (lifetime use unknown)	203	1. RU2/(RU2+RU3) 2. RU3/(RU2+RU3)
Completely missing	158	1. RU1 2. RU2 3. RU3
Total Cases Requiring Imputation	379	

¹ The predictive mean vector components are defined by the following:

1. RU1 = P(Past year use of sedatives)
2. RU2 = P(Lifetime use of sedatives)
3. RU3 = P(Never used sedatives)

Table E.63 Imputation Details for Stimulants Recency of Any Use

Edited Recency of Any Use	Number of Cases	Predictive Mean Vector ¹
Lifetime use	18	1. RU1/(RU1+RU2) 2. RU2/(RU1+RU2)
Not past year (lifetime use unknown)	152	1. RU2/(RU2+RU3) 2. RU3/(RU2+RU3)
Completely missing	147	1. RU1 2. RU2 3. RU3
Total Cases Requiring Imputation	317	

¹ The predictive mean vector components are defined by the following:

1. RU1 = P(Past year use of stimulants)
2. RU2 = P(Lifetime use of stimulants)
3. RU3 = P(Never used stimulants)

Table E.64 Imputation Details for Tranquilizers Recency of Any Use

Edited Recency of Any Use	Number of Cases	Predictive Mean Vector ¹
Lifetime use	30	1. RU1/(RU1+RU2) 2. RU2/(RU1+RU2)
Not past year (lifetime use unknown)	305	1. RU2/(RU2+RU3) 2. RU3/(RU2+RU3)
Completely missing	186	1. RU1 2. RU2 3. RU3
Total Cases Requiring Imputation	521	

¹ The predictive mean vector components are defined by the following:

1. RU1 = P(Past year use of tranquilizers)
2. RU2 = P(Lifetime use of tranquilizers)
3. RU3 = P(Never used tranquilizers)

Table E.65 Imputation Details for Pain Relievers Recency of Any Use

Edited Recency of Any Use	Number of Cases	Predictive Mean Vector ¹
Lifetime use	187	1. RU1/(RU1+RU2) 2. RU2/(RU1+RU2)
Not past year (lifetime use unknown)	354	1. RU2/(RU2+RU3) 2. RU3/(RU2+RU3)
Completely missing	182	1. RU1 2. RU2 3. RU3
Total Cases Requiring Imputation	723	

¹ The predictive mean vector components are defined by the following:

1. RU1 = P(Past year use of pain relievers)
2. RU2 = P(Lifetime use of pain relievers)
3. RU3 = P(Never used pain relievers)

Table E.66 Imputation Details for OxyContin Recency of Any Use

Edited Recency of Any Use	Restrictions Based on Imputation-Revised Pain Relievers Any Use	Number of Cases	Predictive Mean Vector ¹
Completely missing	None	190	1. RU1 2. RU2
	No past year any use	75	N/A
Total Cases Requiring Imputation		265	

¹ The predictive mean vector components are defined by the following:

1. RU1 = P(Past year use of OxyContin | past year use of pain relievers)
2. RU2 = P(No past year use of OxyContin | past year use of pain relievers)

E.3.2.2.2 Recency of Use/Misuse

Table E.67 Imputation Details for Inhalants Recency of Use

Edited Recency of Use	Restrictions Based on Imputation-Revised Lifetime Use	Number of Cases	Predictive Mean Vector ¹
Past year	None	17	1. RU1/(RU1+RU2) 2. RU2/(RU1+RU2)
Completely missing ²	None	247	1. RU1 2. RU2 3. RU3
	Never used	161	N/A
Total Cases Requiring Imputation		425	

N/A = not applicable.

¹ The predictive mean vector components are defined by the following:

1. RU1 = P(Past month use of inhalants | lifetime use of inhalants)
2. RU2 = P(Past year but not past month use of inhalants | lifetime use of inhalants)
3. RU3 = P(Lifetime but not past year use of inhalants | lifetime use of inhalants)

² Respondents who were logically assigned lifetime edited recency of use values are included in the completely missing row.

Table E.68 Imputation Details for Methamphetamine Recency of Use

Edited Recency of Use	Restrictions Based on Imputation-Revised Lifetime Use	Number of Cases	Predictive Mean Vector ¹
Past year	None	17	1. RU1/(RU1+RU2) 2. RU2/(RU1+RU2)
Completely missing ²	None	27	1. RU1 2. RU2 3. RU3
	Never used	85	N/A
Total Cases Requiring Imputation		129	

N/A = not applicable.

¹ The predictive mean vector components are defined by the following:

1. RU1 = P(Past month use of methamphetamine | lifetime use of methamphetamine)
2. RU2 = P(Past year but not past month use of methamphetamine | lifetime use of methamphetamine)
3. RU3 = P(Lifetime but not past year use of methamphetamine | lifetime use of methamphetamine)

² Respondents who were logically assigned lifetime edited recency of use values are included in the completely missing row.

Table E.69 Imputation Details for Hallucinogens Recency of Use

Edited Recency of Use	Restrictions Based on Imputation-Revised Lifetime Use	Number of Cases	Predictive Mean Vector ¹
Past year	None	67	1. $RU1/(RU1+RU2)$ 2. $RU2/(RU1+RU2)$
Completely missing ²	None	192	1. RU1 2. RU2 3. RU3
	Never used	280	N/A
Total Cases Requiring Imputation		539	

N/A = not applicable.

¹ The predictive mean vector components are defined by the following:

1. $RU1 = P(\text{Past month use of hallucinogens} \mid \text{lifetime use of hallucinogens})$
2. $RU2 = P(\text{Past year but not past month use of hallucinogens} \mid \text{lifetime use of hallucinogens})$
3. $RU3 = P(\text{Lifetime but not past year use of hallucinogens} \mid \text{lifetime use of hallucinogens})$

² Respondents who were logically assigned lifetime edited recency of use values are included in the completely missing row.

Table E.70 Imputation Details for LSD Recency of Use

Edited Recency of Use	Restrictions Based on Imputation-Revised Lifetime Use	Restrictions Based on Imputation-Revised Hallucinogen Recency of Use	Restrictions Based on Other Hallucinogen Child Use ¹	Number of Cases	Predictive Mean Vector ²
Past year	None	None	None	7	1. RU1/(RU1+RU2) 2. RU2/(RU1+RU2)
			Past month	1	N/A
		Past year, not past month	N/A	15	N/A
Completely missing ³	None	None	None	14	1. RU1 2. RU2 3. RU3
			Past month	3	N/A
		Not past month	None	14	1. RU2/(RU2+RU3) 2. RU3/(RU2+RU3)
			Past year, not past month	9	N/A
		Lifetime, not past year	N/A	41	N/A
	Never used	N/A	N/A	68	N/A
Total Cases Requiring Imputation				172	

N/A = not applicable.

NOTE: Other hallucinogens for recency directly include peyote, mescaline, and psilocybin and indirectly include any other hallucinogens not covered by LSD, PCP, Ecstasy, ketamine, DMT/AMT/Foxy, or *Salvia divinorum*.

¹ Respondent imputed to IRHALLUCREC, if respondent did not use any "other" hallucinogen child drugs and

1. Respondent only used LSD during lifetime and LSD recency is missing or

2. Respondent used multiple hallucinogen child drugs during lifetime, only LSD recency is missing, and the minimum of the nonmissing recencies is greater than IRHALLUCREC.

² The predictive mean vector components are defined by the following:

1. RU1 = P(Past month use of LSD | lifetime use of LSD)

2. RU2 = P(Past year but not past month use of LSD | lifetime use of LSD)

3. RU3 = P(Lifetime but not past year use of LSD | lifetime use of LSD)

³ Respondents who were logically assigned lifetime edited recency of use values are included in the completely missing row.

Table E.71 Imputation Details for PCP Recency of Use

Edited Recency of Use	Restrictions Based on Imputation-Revised Lifetime Use	Restrictions Based on Imputation-Revised Hallucinogen Recency of Use	Restrictions Based on Other Hallucinogen Child Use ¹	Number of Cases	Predictive Mean Vector ²
Past year	None	None	None	0	1. RU1/(RU1+RU2) 2. RU2/(RU1+RU2)
			Past month	1	N/A
		Past year, not past month	N/A	1	N/A
Completely missing ³	None	None	None	5	1. RU1 2. RU2 3. RU3
			Past month	0	N/A
		Not past month	None	4	1. RU2/(RU2+RU3) 2. RU3/(RU2+RU3)
			Past year, not past month	0	N/A
		Lifetime, not past year	N/A	19	N/A
	Never used	N/A	N/A	52	N/A
Total Cases Requiring Imputation				82	

N/A = not applicable.

NOTE: Other hallucinogens for recency directly include peyote, mescaline, and psilocybin and indirectly include any other hallucinogens not covered by LSD, PCP, Ecstasy, ketamine, DMT/AMT/Foxy, or *Salvia divinorum*.

¹ Respondent imputed to IRHALLUCREC, if respondent did not use any "other" hallucinogen child drugs and

1. Respondent only used PCP during lifetime and PCP recency is missing or

2. Respondent used multiple hallucinogen child drugs during lifetime, only PCP recency is missing, and the minimum of the nonmissing recencies is greater than IRHALLUCREC.

² The predictive mean vector components are defined by the following:

1. RU1 = P(Past month use of PCP | lifetime use of PCP)

2. RU2 = P(Past year but not past month use of PCP | lifetime use of PCP)

3. RU3 = P(Lifetime but not past year use of PCP | lifetime use of PCP)

³ Respondents who were logically assigned lifetime edited recency of use values are included in the completely missing row.

Table E.72 Imputation Details for Ecstasy Recency of Use

Edited Recency of Use	Restrictions Based on Imputation-Revised Lifetime Use	Restrictions Based on Imputation-Revised Hallucinogen Recency of Use	Restrictions Based on Other Hallucinogen Child Use ¹	Number of Cases	Predictive Mean Vector ²
Past year	None	None	None	5	1. RU1/(RU1+RU2) 2. RU2/(RU1+RU2)
			Past month	2	N/A
		Past year, not past month	N/A	8	N/A
Completely missing ³	None	None	None	7	1. RU1 2. RU2 3. RU3
			Past month	6	N/A
		Not past month	None	17	1. RU2/(RU2+RU3) 2. RU3/(RU2+RU3)
			Past year, not past month	9	N/A
		Lifetime, not past year	N/A	44	N/A
	Never used	N/A	N/A	74	N/A
Total Cases Requiring Imputation				172	

N/A = not applicable.

NOTE: Other hallucinogens for recency directly include peyote, mescaline, and psilocybin and indirectly include any other hallucinogens not covered by LSD, PCP, Ecstasy, ketamine, DMT/AMT/Foxy, or *Salvia divinorum*.

- ¹ Respondent imputed to IRHALLUCREC, if respondent did not use any "other" hallucinogen child drugs and
1. Respondent only used Ecstasy during lifetime and Ecstasy recency is missing or
 2. Respondent used multiple hallucinogen child drugs during lifetime, only Ecstasy recency is missing, and the minimum of the nonmissing recencies is greater than IRHALLUCREC.

- ² The predictive mean vector components are defined by the following:

1. RU1 = P(Past month use of Ecstasy | lifetime use of Ecstasy)
2. RU2 = P(Past year but not past month use of Ecstasy | lifetime use of Ecstasy)
3. RU3 = P(Lifetime but not past year use of Ecstasy | lifetime use of Ecstasy)

- ³ Respondents who were logically assigned lifetime edited recency of use values are included in the completely missing row.

Table E.73 Imputation Details for Ketamine Recency of Use

Edited Recency of Use	Restrictions Based on Imputation-Revised Lifetime Use	Restrictions Based on Imputation-Revised Hallucinogen Recency of Use	Restrictions Based on Other Hallucinogen Child Use ¹	Number of Cases	Predictive Mean Vector ²
Past year	None	None	None	4	1. RU1/(RU1+RU2) 2. RU2/(RU1+RU2)
			Past month	0	N/A
		Past year, not past month	N/A	1	N/A
Completely missing ³	None	None	None	2	1. RU1 2. RU2 3. RU3
			Past month	2	N/A
		Not past month	None	2	1. RU2/(RU2+RU3) 2. RU3/(RU2+RU3)
			Past year, not past month	1	N/A
		Lifetime, not past year	N/A	3	N/A
	Never used	N/A	N/A	103	N/A
Total Cases Requiring Imputation				118	

N/A = not applicable.

NOTE: Other hallucinogens for recency directly include peyote, mescaline, and psilocybin and indirectly include any other hallucinogens not covered by LSD, PCP, Ecstasy, ketamine, DMT/AMT/Foxy, or *Salvia divinorum*.

- ¹ Respondent imputed to IRHALLUCREC, if respondent did not use any "other" hallucinogen child drugs and
1. Respondent only used ketamine during lifetime and ketamine recency is missing or
 2. Respondent used multiple hallucinogen child drugs during lifetime, only ketamine recency is missing, and the minimum of the nonmissing recencies is greater than IRHALLUCREC.

- ² The predictive mean vector components are defined by the following:

1. RU1 = P(Past month use of ketamine | lifetime use of ketamine)
2. RU2 = P(Past year but not past month use of ketamine | lifetime use of ketamine)
3. RU3 = P(Lifetime but not past year use of ketamine | lifetime use of ketamine)

- ³ Respondents who were logically assigned lifetime edited recency of use values are included in the completely missing row.

Table E.74 Imputation Details for AMT/DMT/Foxy Recency of Use

Edited Recency of Use	Restrictions Based on Imputation-Revised Lifetime Use	Restrictions Based on Imputation-Revised Hallucinogen Recency of Use	Restrictions Based on Other Hallucinogen Child Use ¹	Number of Cases	Predictive Mean Vector ²
Past year	None	None	None	1	1. RU1/(RU1+RU2) 2. RU2/(RU1+RU2)
			Past month	0	N/A
		Past year, not past month	N/A	1	N/A
Completely missing ³	None	None	None	6	1. RU1 2. RU2 3. RU3
			Past month	0	N/A
		Not past month	None	2 ⁴	1. RU2/(RU2+RU3) 2. RU3/(RU2+RU3)
			Past year, not past month	0	N/A
		Lifetime, not past year	N/A	8	N/A
	Never used	N/A	N/A	116	N/A
Total Cases Requiring Imputation				134	

N/A = not applicable.

NOTE: Other hallucinogens for recency directly include peyote, mescaline, and psilocybin and indirectly include any other hallucinogens not covered by LSD, PCP, Ecstasy, ketamine, DMT/AMT/Foxy, or *Salvia divinorum*.

- ¹ Respondent imputed to IRHALLUCREC, if respondent did not use any "other" hallucinogen child drugs and
1. Respondent only used AMT/DMT/Foxy during lifetime and AMT/DMT/Foxy recency is missing or
 2. Respondent used multiple hallucinogen child drugs during lifetime, only AMT/DMT/Foxy recency is missing, and the minimum of the nonmissing recencies is greater than IRHALLUCREC.

- ² The predictive mean vector components are defined by the following:

1. RU1 = P(Past month use of AMT/DMT/Foxy | lifetime use of AMT/DMT/Foxy)
2. RU2 = P(Past year but not past month use of AMT/DMT/Foxy | lifetime use of AMT/DMT/Foxy)
3. RU3 = P(Lifetime but not past year use of AMT/DMT/Foxy | lifetime use of AMT/DMT/Foxy)

- ³ Respondents who were logically assigned lifetime edited recency of use values are included in the completely missing row.

- ⁴ One respondent who was logically assigned to past year, not past month use based on the edited age-at-first-use and date-at-first-use variables for overall hallucinogens is included in this count.

Table E.75 Imputation Details for Salvia Recency of Use

Edited Recency of Use	Restrictions Based on Imputation-Revised Lifetime Use	Restrictions Based on Imputation-Revised Hallucinogen Recency of Use	Restrictions Based on Other Hallucinogen Child Use ¹	Number of Cases	Predictive Mean Vector ²
Past year	None	None	None	2	1. RU1/(RU1+RU2) 2. RU2/(RU1+RU2)
			Past month	0	N/A
		Past year, not past month	N/A	0	N/A
Completely missing ³	None	None	None	1	1. RU1 2. RU2 3. RU3
			Past month	1	N/A
		Not past month	None	3	1. RU2/(RU2+RU3) 2. RU3/(RU2+RU3)
			Past year, not past month	1	N/A
		Lifetime, not past year	N/A	8	N/A
	Never used	N/A	N/A	131	N/A
Total Cases Requiring Imputation				147	

N/A = not applicable.

NOTE: Other hallucinogens for recency directly include peyote, mescaline, and psilocybin and indirectly include any other hallucinogens not covered by LSD, PCP, Ecstasy, ketamine, DMT/AMT/Foxy, or *Salvia divinorum*.

¹ Respondent imputed to IRHALLUCREC, if respondent did not use any "other" hallucinogen child drugs and

1. Respondent only used *Salvia divinorum* during lifetime and *Salvia divinorum* recency is missing or
2. Respondent used multiple hallucinogen child drugs during lifetime, only *Salvia divinorum* recency is missing, and the minimum of the nonmissing recencies is greater than IRHALLUCREC.

² The predictive mean vector components are defined by the following:

1. RU1 = P(Past month use of *Salvia divinorum* | lifetime use of *Salvia divinorum*)
2. RU2 = P(Past year but not past month use of *Salvia divinorum* | lifetime use of *Salvia divinorum*)
3. RU3 = P(Lifetime but not past year use of *Salvia divinorum* | lifetime use of *Salvia divinorum*)

³ Respondents who were logically assigned lifetime edited recency of use values are included in the completely missing row.

Table E.76 Imputation Details for Sedatives Recency of Misuse

Edited Recency of Misuse	Restrictions Based on Imputation-Revised Recency of Any Use	Number of Cases	Predictive Mean Vector ¹
Past year	None	1	1. $RM1/(RM1+RM2)$ 2. $RM2/(RM1+RM2)$
Lifetime	None	3	1. $RM1/(RM1+RM2+RM3)$ 2. $RM2/(RM1+RM2+RM3)$ 3. $RM3/(RM1+RM2+RM3)$
	Lifetime not past year misuse	2	N/A
Not Past Year	None	30	1. $RM3/(RM3+RM4)$ 2. $RM4/(RM3+RM4)$
	Never misused	188	N/A
Completely missing	None	15	1. RM1 2. RM2 3. RM3 4. RM4
	Did not misuse in past year (lifetime misuse unknown)	5	1. $RM3/(RM3+RM4)$ 2. $RM4/(RM3+RM4)$
	Never misused	149	N/A
Total Cases Requiring Imputation		393	

N/A = not applicable.

¹ The predictive mean vector components are defined by the following:

1. $RM1 = P(\text{Past month misuse of sedatives} \mid \text{lifetime use of sedatives})$
2. $RM2 = P(\text{Past year but not past month misuse of sedatives} \mid \text{lifetime use of sedatives})$
3. $RM3 = P(\text{Lifetime but not past year misuse of sedatives} \mid \text{lifetime use of sedatives})$
4. $RM4 = P(\text{Never misused sedatives} \mid \text{lifetime use of sedatives})$

Table E.77 Imputation Details for Stimulants Recency of Misuse

Edited Recency of Misuse	Restrictions Based on Imputation-Revised Recency of Any Use	Number of Cases	Predictive Mean Vector ¹
Past year	None	5	1. $RM1/(RM1+RM2)$ 2. $RM2/(RM1+RM2)$
Lifetime	None	9	1. $RM1/(RM1+RM2+RM3)$ 2. $RM2/(RM1+RM2+RM3)$ 3. $RM3/(RM1+RM2+RM3)$
	Lifetime not past year misuse	0	N/A
Not Past Year	None	24	1. $RM3/(RM3+RM4)$ 2. $RM4/(RM3+RM4)$
	Never misused	140	N/A
Completely missing	None	21	1. RM1 2. RM2 3. RM3 4. RM4
	Did not misuse in past year (lifetime misuse unknown)	3	1. $RM3/(RM3+RM4)$ 2. $RM4/(RM3+RM4)$
	Never misused	138	N/A
Total Cases Requiring Imputation		340	

N/A = not applicable.

¹ The predictive mean vector components are defined by the following:

1. $RM1 = P(\text{Past month misuse of stimulants} \mid \text{lifetime use of stimulants})$
2. $RM2 = P(\text{Past year but not past month misuse of stimulants} \mid \text{lifetime use of stimulants})$
3. $RM3 = P(\text{Lifetime but not past year misuse of stimulants} \mid \text{lifetime use of stimulants})$
4. $RM4 = P(\text{Never misused stimulants} \mid \text{lifetime use of stimulants})$

Table E.78 Imputation Details for Tranquilizers Recency of Misuse

Edited Recency of Misuse	Restrictions Based on Imputation-Revised Recency of Any Use	Number of Cases	Predictive Mean Vector ¹
Past year	None	10	1. $RM1/(RM1+RM2)$ 2. $RM2/(RM1+RM2)$
Lifetime	None	4	1. $RM1/(RM1+RM2+RM3)$ 2. $RM2/(RM1+RM2+RM3)$ 3. $RM3/(RM1+RM2+RM3)$
	Lifetime not past year misuse	0	N/A
Not Past Year	None	28	1. $RM3/(RM3+RM4)$ 2. $RM4/(RM3+RM4)$
	Never misused	288	N/A
Completely missing	None	30	1. RM1 2. RM2 3. RM3 4. RM4
	Did not misuse in past year (lifetime misuse unknown)	4	1. $RM3/(RM3+RM4)$ 2. $RM4/(RM3+RM4)$
	Never misused	166	N/A
Total Cases Requiring Imputation		530	

N/A = not applicable.

¹ The predictive mean vector components are defined by the following:

1. $RM1 = P(\text{Past month misuse of tranquilizers} \mid \text{lifetime use of tranquilizers})$
2. $RM2 = P(\text{Past year but not past month misuse of tranquilizers} \mid \text{lifetime use of tranquilizers})$
3. $RM3 = P(\text{Lifetime but not past year misuse of tranquilizers} \mid \text{lifetime use of tranquilizers})$
4. $RM4 = P(\text{Never misused tranquilizers} \mid \text{lifetime use of tranquilizers})$

Table E.79 Imputation Details for Pain Relievers Recency of Misuse

Edited Recency of Misuse	Restrictions Based on Imputation-Revised Recency of Any Use	Number of Cases	Predictive Mean Vector ¹
Past year	None	14	1. RM1/(RM1+RM2) 2. RM2/(RM1+RM2)
Lifetime	None	48	1. RM1/(RM1+RM2+RM3) 2. RM2/(RM1+RM2+RM3) 3. RM3/(RM1+RM2+RM3)
	Lifetime not past year misuse	4	N/A
Not Past Year	None	126	1. RM3/(RM3+RM4) 2. RM4/(RM3+RM4)
	Never misused	302	N/A
Completely missing	None	77	1. RM1 2. RM2 3. RM3 4. RM4
	Did not misuse in past year (lifetime misuse unknown)	21	1. RM3/(RM3+RM4) 2. RM4/(RM3+RM4)
	Never misused	131	N/A
Total Cases Requiring Imputation		723	

N/A = not applicable.

¹ The predictive mean vector components are defined by the following:

1. RM1 = P(Past month misuse of pain relievers | lifetime use of pain relievers)
2. RM2 = P(Past year but not past month misuse of pain relievers | lifetime use of pain relievers)
3. RM3 = P(Lifetime but not past year misuse of pain relievers | lifetime use of pain relievers)
4. RM4 = P(Never misused pain relievers | lifetime use of pain relievers)

Table E.80 Imputation Details for OxyContin Past Year Recency of Misuse

Edited Recency of Misuse	Restrictions Based on Imputation-Revised Recency of Any Use	Restrictions Based on Imputation-Revised Pain Relievers Recency of Misuse	Number of Cases	Predictive Mean Vector ¹
Completely missing	None	None	5	1. RM1 2. RM2
		No past year misuse	6	N/A
	No past year misuse	N/A	106	N/A
Total Cases Requiring Imputation			117	

N/A = not applicable.

¹ The predictive mean vector components are defined by the following:

1. RM1 = P(Past year misuse of OxyContin | past year misuse of pain relievers and past year use of OxyContin)
2. RM2 = P(No past year misuse of OxyContin | past year misuse of pain relievers and past year use of OxyContin)

E.3.2.2.3 12-Month Frequency of Use

Table E.81 Logical Constraints for Various Drugs 12-Month Frequency

Constraint #	Logical Constraint
LogC1	Donor's proportion of past year drug use * recipient's maximum number of days could have used drug in the past year must be less than or equal to recipient's maximum possible past year drug frequency of use Recipient's maximum possible drug frequency of use in the past year is limited by the following factors: (1) it must be less than or equal to the maximum period recipient could have used drug, as determined by the month of first drug use (2) if the maximum period recipient could have used drug is greater than 30, but recipient is a past month drug user with a nonmissing 30-day drug frequency, then the past year drug frequency must be less than or equal to the maximum period (minus the number of days recipient did not use drug in the past month) (3) if recipient is not a past month drug user, then the past year drug frequency must be less than or equal to the maximum period (minus 30)
LogC2	Donor's proportion of past year drug use * recipient's maximum number of days could have used drug in the past year must be greater than or equal to recipient's minimum possible past year drug frequency of use Recipient's minimum possible drug frequency of use in the past year is limited by the following factors: (1) if recipient is a past month drug user, it must be at least as much as the 30-day drug frequency (2) if recipient is not a past month drug user but is a past year drug user, it must be at least 1
LogC3	If recipient's age at first use equals his or her current age, and (1) donor is a past month hallucinogen user, then donor's proportion of past year use * recipient's maximum number of days could have used in the past year must be less than or equal to the number of days between recipient's interview date and birthday (inclusive) (2) donor is a past year but not past month user, then donor's proportion of past year use * recipient's maximum number of days could have used in the past year must be less than or equal to the number of days between recipient's interview date and birthday (minus 30)

Table E.82 Likeness Constraints for Various Drugs 12-Month Frequency

Constraint #	Likeness Constraint
LikC1	Donor's drug recency must match recipient's drug recency
LikC2	State rank of donor = state rank of recipient
LikC3	Donor's predicted means each must be within 5 percent of recipient's predicted means
LikC4	Donor and recipient are in the same response propensity bin

Table E.83 Constraints and Portion of the Predictive Mean Vector for Inhalants 12-Month Frequency

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Response Propensity Bin				
					Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
Cycle 1									
1	Missing	98	1,2,3	1. PY	1-4: 1 1,2,4: 1	1-4: 2 1,2,4: 5	1-4: 4 1,2,4: 17	1-4: 12 1,2,4: 15	1-4: 17 1,2,4: 24
Cycle 2									
1	Missing	98	1,2,3	1. PY	1-4: 2 1,2,4: 2	1-4: 1 1,2,4: 5	1-4: 5 1,2,4: 14	1-4: 17 1,2,4: 14	1-4: 20 1,2,4: 18

¹ The predictive mean vector components are defined by the following:

1. PY = P(Inhalant use on a given day in the past year | past year inhalant use)

Table E.84 Constraints and Portion of the Predictive Mean Vector for Methamphetamine 12-Month Frequency

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Response Propensity Bin				
					Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
Cycle 1									
1	Missing	27	1,2,3	1. PY	1-4: 0 1,2,4: 2	1-4: 1 1,2,4: 3	1-4: 0 1,2,4: 5	1-4: 2 1,2,4: 4	1-4: 0 1,2,4: 10
Cycle 2									
1	Missing	27	1,2,3	1. PY	1-4: 0 1,2,4: 2	1-4: 0 1,2,4: 1	1-4: 0 1,2,4: 5	1-4: 1 1,2,4: 5	1-4: 1 1,2,4: 12

¹ The predictive mean vector components are defined by the following:

1. PY = P(Methamphetamine use on a given day in the past year | past year methamphetamine use)

Table E.85 Constraints and Portion of the Predictive Mean Vector for Hallucinogens 12-Month Frequency

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Response Propensity Bin				
					Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
Cycle 1									
1	Missing	140	1,2,3	1. PY	1-4: 6 1,2,4: 3	1-4: 7 1,2,4: 5	1-4: 16 1,2,4: 6	1-4: 25 1,2,4: 4	1-4: 32 1,2,4: 36
Cycle 2									
1	Missing	140	1,2,3	1. PY	1-4: 5 1,2,4: 3	1-4: 10 1,2,4: 2	1-4: 18 1,2,4: 5	1-4: 24 1,2,4: 12	1-4: 37 1,2,4: 24

¹ The predictive mean vector components are defined by the following:

1. PY = P(Hallucinogen use on a given day in the past year | past year hallucinogen use)

E.3.2.2.4 30-Day Frequency of Use

Table E.86 Logical Constraints for Various Drugs 30-Day Frequency

Constraint #	Logical Constraint
LogC1	Donor's 30-day drug use must be less than or equal to recipient's maximum number of days could have used drug in past 30 days and greater than or equal to recipient's minimum number of days could have used drug in past 30 days

Table E.87 Likeness Constraints for Various Drugs 30-Day Frequency

Constraint #	Likeness Constraint
LikC1	State rank of donor = state rank of recipient
LikC2	Donor's predicted means each must be within 5 percent of recipient's predicted means

Table E.88 Constraints and Portion of the Predictive Mean Vector for Inhalants 30-Day Frequency

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Response Propensity Bin				
					Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
Cycle 1									
1	Missing	28	1	1. PM	No Cases	1,2: 5	1,2: 5 1: 1	1,2: 9	1,2: 8
Cycle 2									
1	Missing	28	1	1. PM	1,2: 2 1: 1	1,2: 3 1: 1	1,2: 3 1: 4	1,2: 3 1: 2	1,2: 4 1: 5

¹ The predictive mean vector components are defined by the following:

1. PM = P(Inhalant use on a given day in the past month | past month inhalant use)

Table E.89 Constraints and Portion of the Predictive Mean Vector for Methamphetamine 30-Day Frequency

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Response Propensity Bin				
					Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
Cycle 1									
1	Missing	7	1	1. PM	No Cases	1,2: 1	1,2: 1	1,2: 2	1,2: 3
Cycle 2									
1	Missing	7	1	1. PM	1,2: 1 1: 1	1,2: 0 1: 1	No Cases	No Cases	1,2: 2 1: 2

¹ The predictive mean vector components are defined by the following:

1. PM = P(Methamphetamine use on a given day in the past month | past month methamphetamine use)

Table E.90 Constraints and Portion of the Predictive Mean Vector for Hallucinogens 30-Day Frequency

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Response Propensity Bin				
					Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
Cycle 1									
1	Missing	55	1	1. PM	1,2: 2 1: 1	1,2: 5 1: 1	1,2: 5 1: 3	1,2: 6 1: 7	1,2: 19 1: 6
Cycle 2									
1	Missing	55	1	1. PM	1,2: 2	1,2: 5	1,2: 4 1: 3	1,2: 11 1: 2	1,2: 13 1: 15

¹ The predictive mean vector components are defined by the following:

1. PM = P(Hallucinogens use on a given day in the past month | past month hallucinogens use)

Table E.91 Constraints and Portion of the Predictive Mean Vector for Sedatives 30-Day Frequency

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Response Propensity Bin				
					Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
Cycle 1									
1	Missing	1	1	1. PM	No Cases	No Cases	No Cases	No Cases	1,2: 1
Cycle 2									
1	Missing	1	1	1. PM	No Cases	No Cases	No Cases	1,2: 1	No Cases

¹ The predictive mean vector components are defined by the following:

1. PM = P(Sedative use on a given day in the past month | past month sedative use)

Table E.92 Constraints and Portion of the Predictive Mean Vector for Stimulants 30-Day Frequency

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Response Propensity Bin				
					Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
Cycle 1									
1	Missing	6	1	1. PM	No Cases	No Cases	1,2: 1	1,2: 2	1,2: 3
Cycle 2									
1	Missing	6	1	1. PM	No Cases	No Cases	1,2: 1	1,2: 1	1,2: 4

¹ The predictive mean vector components are defined by the following:

1. PM = P(Stimulant use on a given day in the past month | past month stimulant use)

Table E.93 Constraints and Portion of the Predictive Mean Vector for Tranquilizers 30-Day Frequency

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Response Propensity Bin				
					Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
Cycle 1									
1	Missing	9	1	1. PM	No Cases	1,2: 3	No Cases	1,2: 3	1,2: 3
Cycle 2									
1	Missing	9	1	1. PM	No Cases	No Cases	1,2: 3	1,2: 3	1,2: 3

¹ The predictive mean vector components are defined by the following:

1. PM = P(Tranquilizer use on a given day in the past month | past month tranquilizer use)

Table E.94 Constraints and Portion of the Predictive Mean Vector for Pain Relievers 30-Day Frequency

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Response Propensity Bin				
					Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
Cycle 1									
1	Missing	15	1	1. PM	No Cases	1,2: 1	1,2: 1	1,2: 3	1,2: 10
Cycle 2									
1	Missing	15	1	1. PM	No Cases	No Cases	1,2: 1	1,2: 2	1,2: 12

¹ The predictive mean vector components are defined by the following:

1. PM = P(Pain reliever use on a given day in the past month | past month pain reliever use)

E.3.2.2.5 Past Year Initiation of Misuse

Table E.95 Imputation Details for Past Year Initiation of Misuse for Sedatives

Edited Past Year Initiation of Misuse Indicator	Restrictions Based on Imputation-Revised Recency of Misuse	Number of Cases	Predictive Mean Vector ¹
Completely missing	None	4	1. INIT1 2. INIT2
	Not a past year initiate	172	N/A
Total Cases Requiring Imputation		176	

N/A = not applicable.

¹ The predictive mean vector components are defined by the following:

1. INIT1 = P(Past year initiation of sedatives | past year misuse of sedatives)
2. INIT2 = P(No past year initiation of sedatives | past year misuse of sedatives)

Table E.96 Imputation Details for Past Year Initiation of Misuse for Stimulants

Edited Past Year Initiation of Misuse Indicator	Restrictions Based on Imputation-Revised Recency of Misuse	Number of Cases	Predictive Mean Vector ¹
Completely missing	None	20	1. INIT1 2. INIT2
	Not a past year initiate	161	N/A
Total Cases Requiring Imputation		181	

N/A = not applicable.

¹ The predictive mean vector components are defined by the following:

1. INIT1 = P(Past year initiation of stimulants | past year misuse of stimulants)
2. INIT2 = P(No past year initiation of stimulants | past year misuse of stimulants)

Table E.97 Imputation Details for Past Year Initiation of Misuse for Tranquilizers

Edited Past Year Initiation of Misuse Indicator	Restrictions Based on Imputation-Revised Recency of Misuse	Number of Cases	Predictive Mean Vector ¹
Completely missing	None	20	1. INIT1 2. INIT2
	Not a past year initiate	194	N/A
Total Cases Requiring Imputation		214	

N/A = not applicable.

¹ The predictive mean vector components are defined by the following:

1. INIT1 = P(Past year initiation of tranquilizers | past year misuse of tranquilizers)
2. INIT2 = P(No past year initiation of tranquilizers | past year misuse of tranquilizers)

Table E.98 Imputation Details for Past Year Initiation of Misuse for Pain Relievers

Edited Past Year Initiation of Misuse Indicator	Restrictions Based on Imputation-Revised Recency of Misuse	Number of Cases	Predictive Mean Vector ¹
Completely missing	None	42	1. INIT1 2. INIT2
	Not a past year initiate	256	N/A
Total Cases Requiring Imputation		298	

N/A = not applicable.

¹ The predictive mean vector components are defined by the following:

1. INIT1 = P(Past year initiation of pain relievers | past year misuse of pain relievers)
2. INIT2 = P(No past year initiation of pain relievers | past year misuse of pain relievers)

E.3.2.3 Age at First Use

Table E.99 Logical Constraints for modPMN Age at First Use, Univariate

Constraint #	Logical Constraint
LogC1	Donor's age at first drug use must be less than or equal to recipient's age
LogC2	If recipient was not a past year drug user, then donor cannot have age at first drug use equal to recipient's age
LogC3	If recipient's 12-month drug frequency was greater than the number of days since recipient's most recent birthday, then donor cannot have age at first drug use equal to recipient's age
LogC4	If recipient's most recent birthday was within the past 30 days, and recipient's drug recency was past year but not past month, then donor cannot have age at first drug use equal to recipient's age
LogC5	If recipient's drug recency was past year but not past month, and the difference between the number of days since recipient's most recent birthday and recipient's 12-month drug frequency was less than 30, then donor cannot have age at first drug use equal to recipient's age. (These recipients are missed by LogC3, but the idea is the same: the date of first drug use must be earlier than recipient's most recent birthday.)
LogC6	If recipient was not a past year user of hallucinogens or was not a past year user of any hallucinogen child drugs (LSD, PCP, Ecstasy, ketamine, <i>Salvia divinorum</i> , or DMT/AMT/Foxy), then donor cannot have hallucinogen age at first drug use equal to recipient's age
LogC7	If recipient's most recent birthday was within the past 30 days, and recipient's hallucinogen recency was past year but not past month or any hallucinogen child drugs recency (LSD, PCP, Ecstasy, ketamine, <i>Salvia divinorum</i> , or DMT/AMT/Foxy) was past year but not past month, then donor cannot have hallucinogen age at first drug use equal to recipient's age
LogC8	If recipient was missing the age at first hallucinogen use, then donor must have hallucinogen age at first use less than or equal to the recipient's child drug(s) age(s) at first use (if existing; LSD, PCP, or Ecstasy)
LogC9	If recipient was not a lifetime user of "other" ¹ hallucinogens, then donor must not have values that would cause the minimum of the child drug(s) age(s) at first use to be greater than the overall hallucinogens age at first use for recipient. For example, if donor was not a lifetime user of "other" hallucinogens and has HALLUCAGE = 17, LSDAGE = missing, PCPAGE = 19, and ECSTMOAGE = missing, then donor must have LSDAGE = 17 and/or ECSTMOAGE = 17.
LogC10	Donor's age at first use must be greater than or equal to recipient's hallucinogen age at first use

¹ Other hallucinogens for LogC12 directly include peyote, mescaline, psilocybin, ketamine, DMT/AMT/Foxy, *Salvia divinorum* and indirectly include any other hallucinogens indicated by write-in responses.

Table E.100 Likeness Constraints for modPMN Age at First Use, Univariate

Constraint #	Likeness Constraint
LikC1	State rank of donor = state rank of recipient
LikC2	Donor's predicted mean must be within 5 percent of recipient's predicted mean
LikC3	Age of donor = age of recipient
LikC4	Age of donor must be greater than or equal to age of recipient
LikC5	If recipient was a past year drug user, then donor was a past year drug user; if recipient was not a past year drug user, then donor was the same
LikC6	If recipient was not a past year drug user, then donor was the same
LikC7	If recipient was a past year user, then donor was a past year user
LikC8	For each child drug (LSD, PCP, Ecstasy), if recipient was a past year user, then donor was a past year user
LikC9	For each child drug (LSD, PCP, Ecstasy), if recipient was a lifetime but not past year user, then donor was not a past year user
LikC10	Donor agrees with recipient with respect to lifetime use for child drug(s) (LSD, PCP, Ecstasy)
LikC11	Donor was at least as old as recipient, but no more than 20 years older than recipient
LikC12	Donor was no more than 20 years older than recipient
LikC13	Donor and recipient are in the same response propensity bin

Table E.101 Constraints and Portion of the Predictive Mean Vector for Inhalants Age at First Use

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Response Propensity Bin				
					Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
Cycle 1									
1	Missing	504	1,2,3,4,5	1. PrAFU	1-3,5,13: 14 2,3,5,13: 0 3,5,13: 2	1-3,5,13: 19 2,3,5,13: 3 3,5,13: 5	1-3,5,13: 39 2,3,5,13: 6 3,5,13: 4 3,6,13: 1	1-3,5,13: 85 2,3,5,13: 10 3,5,13: 12 3,6,13: 2	1-3,5,13: 266 2,3,5,13: 13 3,5,13: 15 3,6,13: 1 4,6,13: 6 6,13: 1
Cycle 2									
1	Missing	504	1,2,3,4,5	1. PrAFU	1-3,5,13: 13 2,3,5,13: 1 3,5,13: 2	1-3,5,13: 20 2,3,5,13: 3 3,5,13: 4	1-3,5,13: 37 2,3,5,13: 6 3,5,13: 6 3,6,13: 1	1-3,5,13: 85 2,3,5,13: 10 3,5,13: 12 3,6,13: 2	1-3,5,13: 262 2,3,5,13: 12 3,5,13: 20 3,6,13: 1 4,6,13: 6 6,13: 1

¹ The predictive mean vector components are defined by the following:

1. PrAFU = Predicted inhalant age at first use | used inhalants in lifetime

Table E.102 Constraints and Portion of the Predictive Mean Vector for Methamphetamine Age at First Use

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Response Propensity Bin				
					Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
Cycle 1									
1	Missing	34	1,2,3,4,5	1. PrAFU	1-3,5,13: 1 2,3,5,13: 0 3,5,13: 0 3,6,13: 1	1-3,5,13: 1 2,3,5,13: 2	1-3,5,13: 0 2,3,5,13: 3	1-3,5,13: 2 2,3,5,13: 1 3,5,13: 1	1-3,5,13: 8 2,3,5,13: 7 3,5,13: 4 3,6,13: 3
Cycle 2									
1	Missing	34	1,2,3,4,5	1. PrAFU	1-3,5,13: 2 2,3,5,13: 0 3,5,13: 1	1-3,5,13: 2	1-3,5,13: 1 2,3,5,13: 2 3,5,13: 2 3,6,13: 0 4,6,13: 1	1-3,5,13: 1	1-3,5,13: 6 2,3,5,13: 12 3,5,13: 2 3,6,13: 2

¹ The predictive mean vector components are defined by the following:

1. PrAFU = Predicted methamphetamine age at first use | used methamphetamine in lifetime

Table E.103 Constraints and Portion of the Predictive Mean Vector for Hallucinogens Age at First Use

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Response Propensity Bin				
					Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
Cycle 1									
1	Missing	146	1,3,5-9	1. PrHALAFU	1-3,6-10,13: 3 2,3,6-10,13: 0 2,3,6-9,13: 1 3,6-9,13: 2	1-3,6-10,13: 5 2,3,6-10,13: 3 2,3,6-9,13: 4 3,6-9,13: 1	1-3,6-10,13: 4 2,3,6-10,13: 2 2,3,6-9,13: 3 3,6-9,13: 3 3,6,9,13: 0 6,9,11,13: 1	1-3,6-10,13: 6 2,3,6-10,13: 4 2,3,6-9,13: 11 3,6-9,13: 6 3,6,9,13: 1	1-3,6-10,13: 22 2,3,6-10,13: 28 2,3,6-9,13: 20 3,6-9,13: 14 3,6,9,13: 1 6,9,11,13: 1
Cycle 2									
1	Missing	146	1,3,5-9	1. PrHALAFU	1-3,6-10,13: 3 2,3,6-10,13: 0 2,3,6-9,13: 0 3,6-9,13: 3	1-3,6-10,13: 4 2,3,6-10,13: 6 2,3,6-9,13: 3	1-3,6-10,13: 4 2,3,6-10,13: 3 2,3,6-9,13: 2 3,6-9,13: 3 3,6,9,13: 0 6,9,11,13: 1	1-3,6-10,13: 6 2,3,6-10,13: 5 2,3,6-9,13: 10 3,6-9,13: 6 3,6,9,13: 1	1-3,6-10,13: 33 2,3,6-10,13: 23 2,3,6-9,13: 18 3,6-9,13: 10 3,6,9,13: 1 6,9,11,13: 1

¹ The predictive mean vector components are defined by the following:

1. PrHALAFU = Predicted hallucinogen age at first use | used hallucinogens in lifetime

Table E.104 Constraints and Portion of the Predictive Mean Vector for LSD Age at First Use

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Response Propensity Bin				
					Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
Cycle 1									
1	Missing	82	1,2,4,10	1. PrLSDAFU	1-3,6,7,13: 5 2,3,6,7,13: 0 3,6,7,13: 1	1-3,6,7,13: 3 2,3,6,7,13: 2 3,6,7,13: 3	1-3,6,7,13: 5 2,3,6,7,13: 1 3,6,7,13: 4 3,6,13: 1 6,11,13: 2	1-3,6,7,13: 12 2,3,6,7,13: 3 3,6,7,13: 3 3,6,13: 0 6,11,13: 1	1-3,6,7,13: 13 2,3,6,7,13: 10 3,6,7,13: 12 3,6,13: 0 6,11,13: 1
Cycle 2									
1	Missing	82	1,2,4,10	1. PrLSDAFU	1-3,6,7,13: 5 2,3,6,7,13: 0 3,6,7,13: 1	1-3,6,7,13: 6 2,3,6,7,13: 0 3,6,7,13: 2	1-3,6,7,13: 5 2,3,6,7,13: 2 3,6,7,13: 2 3,6,13: 1 6,11,13: 2	1-3,6,7,13: 11 2,3,6,7,13: 6 3,6,7,13: 4 3,6,13: 0 6,11,13: 2	1-3,6,7,13: 13 2,3,6,7,13: 10 3,6,7,13: 9 3,6,13: 0 6,11,13: 1

¹ The predictive mean vector components are defined by the following:

1. PrLSDAFU = Predicted LSD age at first use | used LSD in lifetime

Table E.105 Constraints and Portion of the Predictive Mean Vector for PCP Age at First Use

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Response Propensity Bin				
					Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
Cycle 1									
1	Missing	37	1,2,4,10	1. PrPCPAFU	No Cases	1-3,6,7,13: 0 2,3,6,7,13: 0 3,6,7,13: 1	1-3,6,7,13: 3 2,3,6,7,13: 0 3,6,7,13: 3 3,6,13: 0 6,11,13: 4	1-3,6,7,13: 1 2,3,6,7,13: 3 3,6,7,13: 2 3,6,13: 0 6,11,13: 3 6,12,13: 0 12,13: 0 13: 1	1-3,6,7,13: 1 2,3,6,7,13: 1 3,6,7,13: 11 3,6,13: 0 6,11,13: 3
Cycle 2									
1	Missing	37	1,2,4,10	1. PrPCPAFU	No Cases	1-3,6,7,13: 0 2,3,6,7,13: 1 3,6,7,13: 1	1-3,6,7,13: 1 2,3,6,7,13: 2 3,6,7,13: 1 3,6,13: 0 6,11,13: 4	1-3,6,7,13: 2 2,3,6,7,13: 1 3,6,7,13: 5 3,6,13: 0 6,11,13: 1	1-3,6,7,13: 1 2,3,6,7,13: 2 3,6,7,13: 11 3,6,13: 0 6,11,13: 4

¹ The predictive mean vector components are defined by the following:

1. PrPCPAFU = Predicted PCP age at first use | used PCP in lifetime

Table E.106 Constraints and Portion of the Predictive Mean Vector for Ecstasy Age at First Use

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Response Propensity Bin				
					Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
Cycle 1									
1	Missing	50	1,2,4,10	1. PrECSAFU	1-3,6,7,13: 2 2,3,6,7,13: 2	1-3,6,7,13: 0 2,3,6,7,13: 1 3,6,7,13: 0 3,6,13: 0 6,11,13: 1	1-3,6,7,13: 3 2,3,6,7,13: 2	1-3,6,7,13: 7 2,3,6,7,13: 2 3,6,7,13: 1	1-3,6,7,13: 11 2,3,6,7,13: 4 3,6,7,13: 12 3,6,13: 0 6,11,13: 2
Cycle 2									
1	Missing	50	1,2,4,10	1. PrECSAFU	1-3,6,7,13: 5 2,3,6,7,13: 1	1-3,6,7,13: 1	1-3,6,7,13: 3 2,3,6,7,13: 2	1-3,6,7,13: 8 2,3,6,7,13: 3 3,6,7,13: 1	1-3,6,7,13: 10 2,3,6,7,13: 6 3,6,7,13: 8 3,6,13: 0 6,11,13: 2

¹ The predictive mean vector components are defined by the following:
1. PrAFU = Predicted Ecstasy age at first use | used Ecstasy in lifetime

Table E.107 Imputation Details for Age of First Misuse for Sedatives

Edited Age of First Misuse	Restrictions Based on Imputation-Revised Past Year Initiation Indicator	Number of Cases	Predictive Mean Vector ¹
Completely missing	None	7	1. AFM1 2. AFM2
	Legitimate skip - not a past year initiate	174	N/A
Total Cases Requiring Imputation		181	

N/A = not applicable.

¹ The predictive mean vector components are defined by the following:

1. AFM1 = P(First misuse of sedatives at current age | past year initiation of sedatives)
2. AFM2 = P(First misuse of sedatives at age – 1 | past year initiation of sedatives)

Table E.108 Imputation Details for Age of First Misuse for Stimulants

Edited Age of First Misuse	Restrictions Based on Imputation-Revised Past Year Initiation Indicator	Number of Cases	Predictive Mean Vector ¹
Completely missing	None	16	1. AFM1 2. AFM2
	Legitimate skip - not a past year initiate	177	N/A
Total Cases Requiring Imputation		193	

N/A = not applicable.

¹ The predictive mean vector components are defined by the following:

1. AFM1 = P(First misuse of stimulants at current age | past year initiation of stimulants)
2. AFM2 = P(First misuse of stimulants at age – 1 | past year initiation of stimulants)

Table E.109 Imputation Details for Age of First Misuse for Tranquilizers

Edited Age of First Misuse	Restrictions Based on Imputation-Revised Past Year Initiation Indicator	Number of Cases	Predictive Mean Vector ¹
Completely missing	None	21	1. AFM1 2. AFM2
	Legitimate skip - not a past year initiate	207	N/A
Total Cases Requiring Imputation		228	

N/A = not applicable.

¹ The predictive mean vector components are defined by the following:

1. AFM1 = P(First misuse of tranquilizers at current age | past year initiation of tranquilizers)
2. AFM2 = P(First misuse of tranquilizers at age – 1 | past year initiation of tranquilizers)

Table E.110 Imputation Details for Age of First Misuse for Pain Relievers

Edited Age of First Misuse	Restrictions Based on Imputation-Revised Past Year Initiation Indicator	Number of Cases	Predictive Mean Vector ¹
Completely missing	None	120	1. AFM1 2. AFM2
	Legitimate skip - not a past year initiate	282	N/A
Total Cases Requiring Imputation		402	

N/A = not applicable.

¹ The predictive mean vector components are defined by the following:

1. AFM1 = P(First misuse of pain relievers at current age | past year initiation of pain relievers)
2. AFM2 = P(First misuse of pain relievers at age – 1 | past year initiation of pain relievers)

E.3.2.4 Substance Dependence and Abuse

Table E.111 Imputation Details for Past Year Dependence of Hallucinogens

Edited Past Year Dependence	Number of Cases	Predictive Mean Vector ¹
Completely missing	87	1. SUD1 2. SUD2
Total Cases Requiring Imputation		87

¹ The predictive mean vector components are defined by the following:

1. SUD1 = P(Past year dependence of hallucinogens | past year use of hallucinogens)
2. SUD2 = P(No past year dependence of hallucinogens | past year use of hallucinogens)

Table E.112 Imputation Details for Past Year Abuse of Hallucinogens

Edited Past Year Abuse	Number of Cases	Predictive Mean Vector ¹
Completely missing	82	1. SUD1 2. SUD2
Total Cases Requiring Imputation		82

¹ The predictive mean vector components are defined by the following:

1. SUD1 = P(Past year abuse of hallucinogens | past year use of hallucinogens)
2. SUD2 = P(No past year abuse of hallucinogens | past year use of hallucinogens)

Table E.113 Imputation Details for Past Year Dependence of Inhalants

Edited Past Year Dependence	Number of Cases	Predictive Mean Vector ¹
Completely missing	68	1. SUD1 2. SUD2
Total Cases Requiring Imputation		68

¹ The predictive mean vector components are defined by the following:

1. SUD1 = P(Past year dependence of inhalants | past year use of inhalants)
2. SUD2 = P(No past year dependence of inhalants | past year use of inhalants)

Table E.114 Imputation Details for Past Year Abuse of Inhalants

Edited Past Year Abuse	Number of Cases	Predictive Mean Vector ¹
Completely missing	65	1. SUD1 2. SUD2
Total Cases Requiring Imputation	65	

¹ The predictive mean vector components are defined by the following:

1. SUD1 = P(Past year abuse of inhalants | past year use of inhalants)
2. SUD2 = P(No past year abuse of inhalants | past year use of inhalants)

Table E.115 Imputation Details for Past Year Dependence of Methamphetamine

Edited Past Year Dependence	Number of Cases	Predictive Mean Vector ¹
Completely missing	12	1. SUD1 2. SUD2
Total Cases Requiring Imputation	12	

¹ The predictive mean vector components are defined by the following:

1. SUD1 = P(Past year dependence of methamphetamine | past year use of methamphetamine or needle use to inject methamphetamine)
2. SUD2 = P(No past year dependence of methamphetamine | past year use of methamphetamine or needle use to inject methamphetamine)

Table E.116 Imputation Details for Past Year Abuse of Methamphetamine

Edited Past Year Abuse	Number of Cases	Predictive Mean Vector ¹
Completely missing	12	1. SUD1 2. SUD2
Total Cases Requiring Imputation	12	

¹ The predictive mean vector components are defined by the following:

1. SUD1 = P(Past year abuse of methamphetamine | past year use of methamphetamine or needle use to inject methamphetamine)
2. SUD2 = P(No past year abuse of methamphetamine | past year use of methamphetamine or needle use to inject methamphetamine)

Table E.117 Imputation Details for Past Year Dependence of Stimulants

Edited Past Year Dependence	Number of Cases	Predictive Mean Vector ¹
Completely missing	18	1. SUD1 2. SUD2
Total Cases Requiring Imputation	18	

¹ The predictive mean vector components are defined by the following:

1. SUD1 = P(Past year dependence of stimulants | past year misuse of stimulants)
2. SUD2 = P(No past year dependence of stimulants | past year misuse of stimulants)

Table E.118 Imputation Details for Past Year Abuse of Stimulants

Edited Past Year Abuse	Number of Cases	Predictive Mean Vector ¹
Completely missing	17	1. SUD1 2. SUD2
Total Cases Requiring Imputation	17	

¹ The predictive mean vector components are defined by the following:

1. SUD1 = P(Past year abuse of stimulants | past year misuse of stimulants)
2. SUD2 = P(No past year abuse of stimulants | past year misuse of stimulants)

Table E.119 Imputation Details for Past Year Dependence of Pain Relievers

Edited Past Year Dependence	Number of Cases	Predictive Mean Vector ¹
Completely missing	53	1. SUD1 2. SUD2
Total Cases Requiring Imputation	53	

¹ The predictive mean vector components are defined by the following:

1. SUD1 = P(Past year dependence of pain relievers | past year misuse of pain relievers)
2. SUD2 = P(No past year dependence of pain relievers | past year misuse of pain relievers)

Table E.120 Imputation Details for Past Year Abuse of Pain Relievers

Edited Past Year Abuse	Number of Cases	Predictive Mean Vector ¹
Completely missing	41	1. SUD1 2. SUD2
Total Cases Requiring Imputation	41	

¹ The predictive mean vector components are defined by the following:

1. SUD1 = P(Past year abuse of pain relievers | past year misuse of pain relievers)
2. SUD2 = P(No past year abuse of pain relievers | past year misuse of pain relievers)

Table E.121 Imputation Details for Past Year Dependence of Tranquilizers

Edited Past Year Dependence	Number of Cases	Predictive Mean Vector ¹
Completely missing	26	1. SUD1 2. SUD2
Total Cases Requiring Imputation	26	

¹ The predictive mean vector components are defined by the following:

1. SUD1 = P(Past year dependence of tranquilizers | past year misuse of tranquilizers)
2. SUD2 = P(No past year dependence of tranquilizers | past year misuse of tranquilizers)

Table E.122 Imputation Details for Past Year Abuse of Tranquilizers

Edited Past Year Abuse	Number of Cases	Predictive Mean Vector ¹
Completely missing	20	1. SUD1 2. SUD2
Total Cases Requiring Imputation	20	

¹ The predictive mean vector components are defined by the following:

1. SUD1 = P(Past year abuse of tranquilizers | past year misuse of tranquilizers)
2. SUD2 = P(No past year abuse of tranquilizers | past year misuse of tranquilizers)

Table E.123 Imputation Details for Past Year Dependence of Sedatives

Edited Past Year Dependence	Number of Cases	Predictive Mean Vector ¹
Completely missing	4	1. SUD1 2. SUD2
Total Cases Requiring Imputation	4	

¹ The predictive mean vector components are defined by the following:

1. SUD1 = P(Past year dependence of sedatives | past year misuse of sedatives)
2. SUD2 = P(No past year dependence of sedatives | past year misuse of sedatives)

Table E.124 Imputation Details for Past Year Abuse of Sedatives

Edited Past Year Abuse	Number of Cases	Predictive Mean Vector ¹
Completely missing	4	1. SUD1 2. SUD2
Total Cases Requiring Imputation	4	

¹ The predictive mean vector components are defined by the following:

1. SUD1 = P(Past year abuse of sedatives | past year misuse of sedatives)
2. SUD2 = P(No past year abuse of sedatives | past year misuse of sedatives)

E.4 Household Composition (Roster) Variables

Tables E.125 through E.140 present information on the missingness patterns, constraints, and predictive mean vectors applied during the imputation procedures for the six household composition (roster) variables: number of rostered individuals, number of children younger than 18, number of individuals aged 65 or older, indicator of whether the respondent has family members in household, number of respondent's family members in the household (including foster relationships), and number of respondent's family members younger than 18 in the household (including foster relationships).

Table E.125 Likeness Constraints for Number of Rostered People

Constraint #	Likeness Constraint
LikC1	Donor's predicted means must be within 5 percent of recipient's predicted means

Table E.126 Constraints and Portion of the Predictive Mean Vector for Number of Rostered People

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group			
					12-17	18-25	26-64	65 or Older
1	Completely missing	50	None	1. C1	1: 3	1: 11	1: 28	1: 8

¹ The predictive mean vector components are defined by the following:

1. C1 = Predicted count from Poisson regression model

Table E.127 Logical Constraints for Number of Children Younger than 18

Constraint #	Logical Constraint
LogC1	Lower and upper bounds were restricted based on IRHHSIZE and nonmissing ages in the roster

Table E.128 Likeness Constraints for Number of Children Younger than 18

Constraint #	Likeness Constraint
LikC1	Donor's predicted means must be within 5 percent of recipient's predicted means
LikC2	IRHHSIZE of donor = IRHHSIZE of recipient

Table E.129 Constraints and Portion of the Predictive Mean Vector for Number of Children Younger than 18

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group			
					12-17	18-25	26-64	65 or Older
1	Completely missing	222	1	1. C1	1,2: 52	1,2: 53 2: 2	1,2: 102 2: 2	1,2: 11

¹ The predictive mean vector components are defined by the following:

1. C1 = Predicted count from Poisson regression model

Table E.130 Logical Constraints for Number of People Aged 65 or Older

Constraint #	Logical Constraint
LogC1	Lower and upper bounds were restricted based on IRHHSIZE, IRKID17, and nonmissing ages in the roster

Table E.131 Likeness Constraints for Number of People Aged 65 or Older

Constraint #	Likeness Constraint
LikC1	Donor's predicted means must be within 5 percent of recipient's predicted means
LikC2	IRHHSIZE of donor = IRHHSIZE of recipient

Table E.132 Constraints and Portion of the Predictive Mean Vector for Number of People Aged 65 or Older

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group			
					12-17	18-25	26-64	65 or Older
1	Completely missing	282	1	1. C1	1,2: 119 1: 1	1,2: 74	1,2: 76 1: 1	1,2: 10 1: 1

¹ The predictive mean vector components are defined by the following:

1. C1 = Predicted count from Poisson regression model

Table E.133 Likeness Constraints for Indicator of Whether the Respondent Has Family Members in Household

Constraint #	Likeness Constraint
LikC1	Donor's predicted means must be within 5 percent of recipient's predicted means
LikC2	Donor's predicted means must be within 20 percent of recipient's predicted means
LikC3	IRKID17 of donor = IRKID17 of recipient
LikC4	If recipient was married, then donor was married; otherwise, if recipient was not currently married, then donor was not currently married

Table E.134 Constraints and Portion of the Predictive Mean Vector for Indicator of Whether the Respondent Has Family Members in Household

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group			
					12-17	18-25	26-64	65 or Older
1	Completely missing	54	None	1. F1	1-4: 3	1-4: 12	1-4: 34	1-4: 5

¹ The predictive mean vector components are defined by the following:

1. F1 = P(No other family members in the household)

Table E.135 Logical Constraints for Number of Respondent's Family Members in Household (Including Foster Relationships)

Constraint #	Logical Constraint
LogC1	Lower and upper bounds were restricted based on IRHHSIZE and nonmissing ages in the roster

Table E.136 Likeness Constraints for Number of Respondent's Family Members in Household (Including Foster Relationships)

Constraint #	Likeness Constraint
LikC1	Donor's predicted means must be within 5 percent of recipient's predicted means
LikC2	IRKID17 of donor = IRKID17 of recipient
LikC3	IRHHSIZE of donor = IRHHSIZE of recipient

Table E.137 Constraints and Portion of the Predictive Mean Vector for Number of Respondent's Family Members in Household (Including Foster Relationships)

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group			
					12-17	18-25	26-64	65 or Older
1	Completely missing	72	1	1. C1	1-3: 6	1-3: 15	1-3: 45 2,3: 1	1-3: 5

¹ The predictive mean vector components are defined by the following:

1. C1 = Predicted count from Poisson regression model

Table E.138 Logical Constraints for Number of Respondent's Family Members in Household Younger than 18 (Including Foster Relationships)

Constraint #	Logical Constraint
LogC1	Lower and upper bounds were restricted based on IRFMLYSZ and nonmissing ages in the roster

Table E.139 Likeness Constraints for Number of Respondent's Family Members in Household Younger than 18 (Including Foster Relationships)

Constraint #	Likeness Constraint
LikC1	Donor's predicted means must be within 5 percent of recipient's predicted means
LikC2	IRKID17 of donor = IRKID17 of recipient
LikC3	IRHHSIZE of donor = IRHHSIZE of recipient
LikC4	IRFMLYSZ of donor = IRFMLYSZ of recipient

Table E.140 Constraints and Portion of the Predictive Mean Vector for Number of Respondent's Family Members in Household Younger than 18 (Including Foster Relationships)

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group			
					12-17	18-25	26-64	65 or Older
1	Completely missing	73	1	1. C1	1-4: 18 2-4: 0 3,4: 1	1-4: 11 2-4: 1 3,4: 0 4: 1	1-4: 37 2-4: 4	No Cases

¹ The predictive mean vector components are defined by the following:

1. C1 = Predicted count from Poisson regression model

E.5 Income Variables

Tables E.141 through E.146 present information on the missingness patterns, constraints, and predictive mean vectors applied during the imputation procedures for the income variables.

E.5.1 Binary Variable Phase

Five of the binary income variables were directly related to a respondent's socioeconomic status. Hence, if a recipient required imputation for one or more of these five variables (i.e., welfare payments, welfare services, Supplemental Nutrition Assistance Program [SNAP] benefits, binary income, and months on welfare), but had information on at least one of these variables, then the donors were restricted so that donors and recipients had the same values for these nonmissing variables. These five variables are referred to as "welfare-correlated variables."

There were a large number of missingness patterns for the source-of-income variables because they are imputed simultaneously in a set. A respondent could be missing any combination of the nine source-of-income variables. Because the constraints and predictive mean vectors can be described easily, only one row in Table E.143 is used to summarize the many multiple variable missingness patterns.

Table E.141 Logical Constraints for Source of Income

Constraint #	Logical Constraint
LogC1	IRFAMSKP of donor = IRFAMSKP of recipient ¹
LogC2	Recipient is missing months on welfare. If recipient is also known to have received either welfare payments or welfare services, then donor must also have received welfare payments or welfare services. (This prevents donor from giving a skip code to a recipient who should have a value from 1 to 12.)

¹ This is only a logical constraint when family binary total income is missing.

Table E.142 Likeness Constraints for Source of Income

Constraint #	Likeness Constraint
LikC1	Age of donor = age of recipient
LikC2	If any of the welfare-correlated edited binary income variables (welfare payments, welfare services, SNAP benefits, binary total income, and months on welfare) were missing, then donor must match recipient on all nonmissing welfare-correlated edited binary income variables
LikC3	Donor's predicted means must be within 5 percent of recipient's predicted means for all missing family variables
LikC4	Donor must match recipient with respect to personal welfare payments (if nonmissing) and welfare services (if nonmissing)
LikC5	If recipient is missing social security, then donor must match recipient with respect to whether there are adults aged 65 or older in the household
LikC6	If recipient is missing welfare payments and/or welfare services, then donor must match recipient with respect to whether there are children younger than 18 in the household
LikC7	Age of donor must be within 5 years of age of recipient
LikC8	If recipient is missing binary income at the family level but not at the personal level, then donor must match recipient with respect to binary income at the personal level
LikC9	If recipient is not missing binary income at the personal level, then donor must match recipient with respect to binary income at the personal level

SNAP = Supplemental Nutrition Assistance Program.

Table E.143 Constraints and Portion of the Predictive Mean Vector for Source of Income

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group			
					12-17	18-25	26-64	65 or Older
1	Missing welfare months, receiving family payment and/or family service	110	1,2	1. WMS, and probabilities associated with other missing elements	1-6: 11 1,2,4-6: 20	1-6: 8 1,2,4-6: 33	1-6: 4 1,2,4-6: 20 2,4-7: 8	1-6: 0 1,2,4-6: 0 2,4-7: 6
2	Missing welfare months, not receiving welfare payments, missing welfare services	94	1,2	1. SVC*WMS, SVC, and probabilities associated with other missing elements	1-6: 15 1,2,4-6: 16 2,4-7: 2	1-6: 5 1,2,4-6: 27	1-6: 3 1,2,4-6: 22	1-6: 2 1,2,4-6: 2
3	Missing welfare months, missing welfare payments, not receiving welfare services	188	1,2	1. PMT*WMS, PMT, and probabilities associated with other missing elements	1-6: 35 1,2,4-6: 23	1-6: 15 1,2,4-6: 61	1-6: 5 1,2,4-6: 41	1-6: 2 1,2,4-6: 5 2,4-7: 1
4	Missing welfare months, missing welfare payments, missing welfare services	210	1,2	1. [1-(1-PMT)(1-SVC)]*WMS, PMT, SVC, and probabilities associated with other missing elements	1-6: 21 1,2,4-6: 66	1-6: 1 1,2,4-6: 43	1-6: 1 1,2,4-6: 64	1-6: 0 1,2,4-6: 13 2,4-7: 1
5	All other missingness patterns with number of months on welfare nonmissing ²	2,997	1	1. Probabilities associated with missing elements	1,2,3,5,6: 618 1,2,5,6: 44 2,5-7: 1	1,2,3,5,6: 908 1,2,5,6: 112 2,5-7: 3	1,2,3,5,6: 849 1,2,5,6: 195 2,5-7: 4	1,2,3,5,6: 170 1,2,5,6: 90 2,5-7: 3

¹ The predictive mean vector components are defined by the following:

1. PMT = P(Family in household received income from welfare payments)
2. SVC = P(Family in household received income from other welfare services)
3. WMS = P(Family in household received any welfare on a given month in the past year | family received any welfare in the past year)

² There were many other missingness patterns for source of income because a respondent could be missing any combination of the nine source-of-income variables. Because the constraints and predictive mean vectors can be described easily, all of these other missingness patterns are represented by a single row in the table.

E.5.2 Specific Category Phase

Table E.144 Logical Constraints for Finer Categories of Income

Constraint #	Logical Constraint
LogC1	IRFAMSKP of donor = IRFAMSKP of recipient
LogC2	Family income (FMINC3) of donor must be greater than personal income (PRINC3) of recipient
LogC3	Personal income (PRINC3) of donor must be less than family income (FMINC3) of recipient

Table E.145 Likeness Constraints for Finer Categories of Income

Constraint #	Likeness Constraint
LikC1	Donor's predicted mean must be within 5 percent of recipient's predicted mean
LikC2	PRINC3 of donor = PRINC3 of recipient
LikC3	FMINC3 of donor = FMINC3 of recipient
LikC4	IRPINC1 of donor = IRPINC1 of recipient
LikC5	IRFINC1 of donor = IRFINC1 of recipient

Table E.146 Constraints and Portion of the Predictive Mean Vector for Finer Categories of Income

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group			
					12-17	18-25	26-64	65 or Older
1	Missing personal income only	580	13	1. $X\beta$	1,3-5: 60 3-5: 1 4,5: 1	1,3-5: 129	1,3-5: 209	1,3-5: 177 3-5: 3
2	Missing family income only	4,557	1,2	1. $X\beta$	1,2,4,5: 1,182 2,4,5: 2 4,5: 1	1,2,4,5: 1,674 2,4,5: 2	1,2,4,5: 1,487 2,4,5: 1	1,2,4,5: 205 2,4,5: 3
3	Completely missing	1,460	1	1. $X\beta$	1,4,5: 186 4,5: 1	1,4,5: 241	1,4,5: 787	1,4,5: 245

¹ For more details, see Section 12.3.2.2.

E.6 Health Insurance Variables

Tables E.147 through E.150 present information on the missingness patterns, constraints, and predictive mean vectors applied during the imputation procedures for the health insurance variables. Prior to 2013, two methods were presented for health insurance variables: the "constituent variables method" and the "old method." See Chapter 13 for an explanation of the "constituent variables method" and why the "old method" is no longer presented.

In several instances, variable names are used without description for the purposes of brevity (see Chapter 13 for details). For the health insurance imputations, matches between donors and recipients were attempted on the nonmissing values of the variables CAIDCHIP, MEDICARE, CHAMPUS, and PRVHLTIN. These variables are the edited indicators of whether the respondent received health insurance from Medicaid/state health insurance programs for children, Medicare, CHAMPUS, or private health insurance, respectively. These were the base

variables used in the creation of the imputation-revised variables (IRMCDCHP, IRMEDICR, IRCHMPUS, IRPRVHLT, and IROTHHLT). In addition to the edited health insurance variables, other variables, which were used as likeness constraints, are identified in the tables only by their variable names. These include SERVICE (an indicator of whether the respondent had ever been in the military), GOVTPROG (an indicator of whether the respondent's family participated in government public assistance programs), INCOME (a four-level categorical family-income variable, with levels of less than \$20,000, \$20,000 to less than \$50,000, \$50,000 to less than \$75,000, and \$75,000 or more), IRFAMIN1 (a two-level family income variable, with levels of less than \$20,000 and \$20,000 or more), and IRFAMSOC (an indicator of whether the respondent's family in the household received income from social security).

For the multivariate predictive mean neighborhood (MPMN) method, the likeness constraints, which were applied to the variables, differed between missingness patterns, and sometimes the constraints differed between age groups within the same missingness pattern.

Table E.147 Likeness Constraints for Health Insurance (MPMN)

Constraint #	Likeness Constraint
LikC1	Donor's predicted means must be within 5 percent of recipient's predicted means for all missing insurance variables (CAIDCHIP, MEDICARE, CHAMPUS, and PRVHLTIN)
LikC2	GOVTPROG of donor = GOVTPROG of recipient
LikC3	IRFAMSOC of donor = IRFAMSOC of recipient
LikC4	SERVICE of donor = SERVICE of recipient (if nonmissing)
LikC5	INCOME of donor = INCOME of recipient
LikC6	IRFAMIN1 of donor = IRFAMIN1 of recipient
LikC7	Donor must match recipient for all nonmissing insurance variables
LikC8	If recipient is between 18 and 64 years old and has a nonmissing value for edited work status (JBSTATR), and (1) recipient has no job due to disability (JBSTATR = 14), then donor must have no job due to disability (2) recipient has JBSTATR not equal to 14, then donor must have JBSTATR not equal to 14

MPMN = multivariate predictive mean neighborhood.

Table E.148 Constraints and Portion of the Predictive Mean Vector for Health Insurance (MPMN)

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group			
					12-17	18-25	26-64	65 or Older
1	Only missing CAIDCHIP	247	None	1. HI1	1,2,7: 79 1,2: 1	1,2,7: 93 1,2: 3	1,2,7: 49 1,2: 2	1,2,7: 20
2	Only missing MEDICARE	83	None	1. HI2	1,3,7: 23	1,3,7,8: 37 1,3,8: 1	1,3,7,8: 19	1,3,7: 2 1: 1
3	Missing CAIDCHIP and MEDICARE	62	None	1. HI1 2. HI2	1-3,7: 12 1-3: 4 1: 1 None: 1	1,2,7,8: 23 1,2,8: 4 1,8: 0 8: 2	1,2,7,8: 10 1,2,8: 2 1,8: 0 8: 1	1,2,7: 1 1,2: 0 1: 0 None: 1
4	Only missing CHAMPUS	35	None	1. HI3	1,4,7: 11	1,4,7: 15 1,4: 0 4: 1	1,4,7: 4	1,4,7: 4
5	Missing CAIDCHIP and CHAMPUS	5	None	1. HI1 2. HI3	1,2,4,7: 0 1,2,4: 1	1,2,4,7: 3	1,2,4,7: 1	No Cases
6	Missing MEDICARE and CHAMPUS	2	None	1. HI2 2. HI3	1,3,4,7: 1	1,4,7,8: 0 1,4,8: 0 4,8: 1	No Cases	No Cases
7	Missing CAIDCHIP, MEDICARE, and CHAMPUS	8	None	1. HI1 2. HI2 3. HI3	1-4,7: 2 1-4: 0 1,4: 0 4: 1	1,2,4,7,8: 1 1,2,4,8: 0 1,4,8: 0 4,8: 4	No Cases	No Cases
8	Only missing PRVHLTIN	238	None	1. HI4	1,5,7: 73	1,5,7: 98 1,5: 1	1,5,7: 43	1,6,7: 21 1,6: 2
9	Missing CAIDCHIP and PRVHLTIN	59	None	1. HI1 2. HI4	1,2,5,7: 30 1,2,5: 0 1,6: 0 1: 0 None: 2	1,2,5,7: 14 1,2,5: 3 1,6: 1 1: 0 None: 1	1,2,5,7: 5 1,2,5: 0 1,6: 1 1: 0 None: 1	1,2,6,7: 1
10	Missing MEDICARE and PRVHLTIN	15	None	1. HI2 2. HI4	1,3,5,7: 6	1,5,7,8: 6 1,5,8: 0 1,6,8: 1	1,5,7,8: 1 1,5,8: 0 1,6,8: 0 1,8: 0 8: 1	No Cases
11	Missing CAIDCHIP, MEDICARE, and PRVHLTIN	36	None	1. HI1 2. HI2 3. HI4	1-3,5,7: 10 1-3,5: 0 1,6: 0 1: 0 None: 5	1,2,5,7,8: 4 1,2,5,8: 0 1,6,8: 2 1,8: 1 8: 7	1,2,5,7,8: 1 1,2,5,8: 0 1,6,8: 0 1,8: 0 8: 4	1,2,6,7: 0 1,2,6: 0 1: 0 None: 2
12	Missing CHAMPUS and PRVHLTIN	8	None	1. HI3 2. HI4	1,4,5,7: 2	1,4,5,7: 4	1,4,5,7: 2	No Cases

Table E.148 Constraints and Portion of the Predictive Mean Vector for Health Insurance (MPMN) (continued)

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group			
					12-17	18-25	26-64	65 or Older
13	Missing CAIDCHIP, CHAMPUS, and PRVHLTIN	17	None	1. HI1 2. HI3 3. HI4	1,2,4,5,7: 4 1,2,4,5: 0 1,4,6: 0 1,4: 0 4: 6	1,2,4,5,7: 1 1,2,4,5: 0 1,4,6: 1 1,4: 0 4: 3	1,2,4,5,7: 0 1,2,4,5: 0 1,4,6: 0 1,4: 1 4: 1	No Cases
14	Missing MEDICARE, CHAMPUS, and PRVHLTIN	1	None	1. HI2 2. HI3 3. HI4	No Cases	1,4,5,7,8: 1	No Cases	No Cases
15	Missing CAIDCHIP, MEDICARE, CHAMPUS, and PRVHLTIN	90	None	1. HI1 2. HI2 3. HI3 4. HI4	1-5: 21 1,4,6: 0 1,4: 0 4: 14	1,2,4,5,8: 2 1,4,6,8: 1 1,4,8: 1 4,8: 10	1,2,4,5,8: 5 1,4,6,8: 0 1,4,8: 0 4,8: 27	1,2,4,6: 1 1,4: 0 4: 8

MPMN = multivariate predictive mean neighborhood.

¹ The predictive mean vector components are defined by the following:

1. HI1 = P(Respondent has CAIDCHIP)
2. HI2 = P(Respondent has Medicare)
3. HI3 = P(Respondent has CHAMPUS)
4. HI4 = P(Respondent has private health insurance)

Table E.149 Constraints for Health Insurance (UPMN), Any Other Health Insurance

Constraint #	Likeness Constraint
LikC1	Donor's predicted means each must be within 5 percent of recipient's predicted means
LikC2	If recipient is between 26 and 64 years old, then donor must also be between 26 and 64 years old; if recipient is aged 65 years or older, then donor must also be aged 65 years or older

UPMN = univariate predictive mean neighborhood.

Table E.150 Constraints and Portion of the Predictive Mean Vector for Health Insurance (UPMN), Any Other Health Insurance

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases, by Age Group		
					12-17	18-25	26 or Older
1	Missing ANYOTHER	190	None	1. HI1	1: 40	1: 106	1,2: 44

UPMN = univariate predictive mean neighborhood.

¹ The predictive mean vector components are defined by the following:

1. HI1 = P(Respondent has any other health insurance)

E.7 Roster Pair Variables

Tables E.151 through E.170 present information on the missingness patterns, constraints, and predictive mean vectors applied during the imputation procedures for the pair variables. Tables E.151 through E.154 correspond to the first stage of imputation, where the pair relationship variable (IRPRREL) is created. Tables E.155 through E.162 are for the second stage, imputation of the multiplicity variables. Finally, Tables E.163 through E.170 correspond to the third stage, where the household counts are imputed.

There are a few instances where variable names are used without description for the purposes of brevity. In these cases, the variables are defined in a table at the beginning of the section in which they are used.

E.7.1 Stage One: Pair Relationship

Table E.151 Variables Used in Constraints for Stage One, Pair Relationship

Variable	Value	Constraint
MARIT1	1	At least one pair member has a marital status of "Married," "Widowed," or "Divorced or Separated" that was not imputed or logically assigned
	0	Neither pair member has a marital status of "Married," "Widowed," or "Divorced or Separated" that was not imputed or logically assigned
MARIT2	2	Both pair members have a marital status of "Married" that was not imputed or logically assigned
	1	One pair member has a marital status of "Married" that was not imputed or logically assigned
	0	Neither pair member has a marital status of "Married," "Widowed," or "Divorced or Separated" that was not imputed or logically assigned
MARIT3	5	Both pair members have a marital status of "Married" that was not imputed or logically assigned
	4	One pair member has a marital status of "Married," and the other has a marital status of "Widowed" or "Divorced or Separated," neither of which was imputed or logically assigned
	3	One pair member has a marital status of "Married," and the other either has a marital status of "Never Married" or was younger than 15 and therefore had a legitimate skip for marital status, neither of which was imputed or logically assigned
	2	Both pair members have a marital status of "Widowed" or "Divorced or Separated" that was not imputed or logically assigned
	1	One pair member has a marital status of "Widowed" or "Divorced or Separated," and the other either has a marital status of "Never Married" or was younger than 15 and therefore had a legitimate skip for marital status, neither of which was imputed or logically assigned
	0	Both pair members either have a marital status of "Never Married" or were younger than 15 and therefore had a legitimate skip for marital status, neither of which was imputed or logically assigned

Table E.152 Logical Constraints for Stage One, Pair Relationship

Constraint #	Logical Constraint
LogC1	If recipient is a spouse-spouse pair, with or without children, then donor must be a spouse-spouse pair either with or without children
LogC2	If recipient could be either a parent-child pair where the child is aged 12 to 14 or some other clearly identifiable pair that is not of interest, then donor must be either a parent-child pair where the child is aged 12 to 14 or some other clearly identifiable pair that is not of interest
LogC3	If recipient could be either a parent-child pair where the child is aged 15 to 17 or some other clearly identifiable pair that is not of interest, then donor must be either a parent-child pair where the child is aged 15 to 17 or some other clearly identifiable pair that is not of interest
LogC4	If recipient could be either a parent-child pair where the child is aged 18 to 20 or some other clearly identifiable pair that is not of interest, then donor must be either a parent-child pair where the child is aged 18 to 20 or some other clearly identifiable pair that is not of interest
LogC5	If recipient could be either a parent-child pair where the child is aged 21 or older or some other clearly identifiable pair that is not of interest, then donor must be either a parent-child pair where the child is aged 21 or older or some other clearly identifiable pair that is not of interest
LogC6	If recipient could be either a sibling-sibling pair where one sibling is aged 12 to 17 and the other is aged 18 to 25 or some other clearly identifiable pair that is not of interest, then donor must be either a sibling-sibling (12-17/18-25) pair or some other clearly identifiable pair that is not of interest
LogC7	If recipient could be either a sibling-sibling pair where the siblings are not in either age range of interest or some other clearly identifiable pair that is not of interest, then donor must be a sibling-sibling pair where the siblings are not in an age range of interest or some other clearly identifiable pair that is not of interest
LogC8	If recipient could be either a spouse-spouse pair with children or another clearly identifiable pair that is not of interest, then donor must be a spouse-spouse pair with children or some other clearly identifiable pair that is not of interest
LogC9	If recipient could be either a spouse-spouse pair without children or another clearly identifiable pair that is not of interest, then donor must be a spouse-spouse pair without children or some other clearly identifiable pair that is not of interest
LogC10	If recipient could be a spouse-spouse pair with or without children or another clearly identifiable pair that is not of interest, then donor must be a spouse-spouse pair with or without children or some other clearly identifiable pair that is not of interest
LogC11	If recipient could be either a grandparent-grandchild pair or another clearly identifiable pair that is not of interest, then donor must be a grandparent-grandchild pair or some other clearly identifiable pair that is not of interest

Table E.153 Likeness Constraints for Stage One, Pair Relationship

Constraint #	Likeness Constraint
LikC1	Donor's predicted means each must be within 5 percent of recipient's predicted means
LikC2	If recipient is not a possible parent-child pair, and recipient has at least one parent in the household, then donor must also have at least one parent in the household
LikC3	If recipient is not a possible parent-child pair, and recipient has no parents in the household, then donor must also not have any parents in the household
LikC4	If the number of children aged 0 to 18 ¹ is nonmissing for both recipient and donor, then the two values must be equal
LikC5	If the recipient pair members are of the same gender, then the donor pair members must be of the same gender. If the recipient pair members are of a different gender, then the donor pair members must also be of a different gender.
LikC6	The age of the donor's younger pair member must be the same age as the recipient's younger pair member
LikC7	If donor is a spouse-spouse pair, with or without children, then donor's younger pair member must be the same age as recipient's younger pair member
LikC8	If both donor and recipient have pair members of different ages, then donor's younger pair member must be the same age as recipient's younger pair member
LikC9	The younger pair member of both donor and recipient must fall within the same age group: 21-25, 26-34, 35-49, or 50 or older
LikC10	The older pair member of both donor and recipient must fall within the same age group: 26-34, 35-49, or 50 or older
LikC11	If recipient pair members are in the same age group (21-25, 26-34, 35-49, or 50 or older), then donor cannot be a parent-child pair where the child is aged 21 or older
LikC12	Neither of the donor's pair members can have had an imputed marital status
LikC13	If donor is a sibling-sibling (12-14/15-17) pair, then neither of donor's pair members can have had an imputed marital status
LikC14	If neither of the recipient's pair members' marital status was imputed or if recipient has MARIT1 = 1, then donor must have the same value for MARIT1 as recipient
LikC15	If neither of the recipient's pair members had an imputed marital status, then donor must have the same value for MARIT2 as recipient. However, if one of the recipient's pair members had an imputed marital status, and recipient had MARIT2 = 1, then donor must have had MARIT2 = 1 or 2.
LikC16	If donor is a sibling-sibling (12-14/15-17) pair, then apply Likeness Constraint 12
LikC17	If neither of the recipient's pair members had an imputed marital status, then donor must have the same value for MARIT3 as recipient. However, if one of the recipient's pair members had an imputed marital status, and recipient had MARIT2 = 1, then donor must have had MARIT2 = 1 or 2.
LikC18	If neither of the recipient's pair members had an imputed marital status, then donor must have the same value of MARIT3 as the recipient. However, if one of the recipient's pair members had an imputed marital status, and recipient had MARIT1 = 1, then donor must also have had MARIT1 = 1.

¹ For age group pairs 3 and 4, this constraint is on the number of children aged 0 to 11.

Table E.154 Constraints and Portion of the Predictive Mean Vector for Stage One, Pair Relationship

#	Age Group Pair ¹	Total Number of Cases	Logical Constraints	Predictive Mean Vector ²	Likeness Constraints: Number of Cases
1	One pair member is aged 12 to 14, the other is aged 15 to 17	16	None	1. SIB	1-3: 15 2: 1
2	One pair member is aged 12 to 14, the other is aged 18 to 25	12	2,6	1. SIB	1-3,12,14: 11 2,3,12,14: 1
3	Both pair members are aged 15 to 17	11	1,3,7-11	1. SIB	1-5,8,13,16: 9 2-5,8,12,15: 2
4	One pair member is aged 15 to 17, the other is aged 18 to 25	14	1,3,6-11	1. SIB	1-6,12,17: 11 2-6,12,17: 3
5	One pair member is aged 18 to 20, the other is aged 18 to 25	34	1,4,7-10	1. SPOUSE1 2. SPOUSE2	1-5,12,17: 8 2-5,12,17: 26
6	Both pair members are aged 21 to 25	39	1,5,7-10	1. SPOUSE1 2. SPOUSE2	1-5,12,15: 9 2-5,12,15: 30
7	One pair member is aged 12 to 14, the other is aged 26 or older	7	2,7,11	1. PC	1-3,10,12,14: 4 2,3,10,12,14: 3
8	One pair member is aged 15 to 17, the other is aged 26 or older	5	3,7-11	1. PC	1-5,7,10,12,18: 3 2-5,7,10,12,18: 2
9	One pair member is aged 18 to 20, the other is aged 26 or older	18	1,4,7-11	1. PC	1-5,10,12,18: 2 2-5,10,12,18: 13 2-5,10,12,14: 1 2-5,12,14: 2
10	One pair member is aged 21 or older, the other is aged 26 or older	110	1,5,7-11	1. SPOUSE1 2. SPOUSE2	1-5,9-12,15: 9 2-5,9-12,15: 98 2-5,11,12,15: 3
11	Both pair members are aged 12 to 17	5	None	1. SIB	1-3: 5

¹ Because there was only one variable to be imputed, PAIRREL, there was only one missingness pattern. However, as the predictive mean vector and constraints vary by age group, the age groups are presented as missingness patterns in this table, and the column heading has been changed accordingly.

² The predictive mean vector components are defined by the following:

1. SIB = P(the pair relationship is sibling-sibling)
2. SPOUSE1 = P(the pair relationship is spouse-spouse, with children)
3. SPOUSE2 = P(the pair relationship is spouse-spouse, without children)
4. PC = P(the pair relationship is parent-child)

E.7.2 Stage Two: Multiplicity Counts

Table E.155 Logical Constraints for Stage Two, Multiplicity Counts

Constraint #	Logical Constraint
LogC2	If recipient is a parent-child pair and if recipient's household count(s) in the appropriate age group(s) (12-14, 15-17, and/or 18-20) is nonmissing, then donor's and recipient's count(s) must match. Also, donor's parent-focus multiplicity count(s) must fall within recipient's bounds.
LogC3	Donor's sibling-sibling (12-14/15-17) older focus multiplicity count must fall within recipient's bounds
LogC4	Donor's sibling-sibling (12-14/15-17) younger focus multiplicity count must fall within recipient's bounds
LogC5	Donor's sibling-sibling (12-17/18-25) older focus multiplicity count must fall within recipient's bounds
LogC6	Donor's sibling-sibling (12-17/18-25) younger focus multiplicity count must fall within recipient's bounds

Table E.156 Likeness Constraints for Stage Two, Multiplicity Counts

Constraint #	Likeness Constraint
LikC1	Donor's predicted means each must be within 5 percent of recipient's predicted means
LikC2	Donor's pair relationship must match recipient's pair relationship
LikC3	Donor and recipient must have the same household size
LikC4	If recipient's count of household members aged 0 to 11 is nonmissing, then donor's value must match
LikC5	If recipient's count of household members aged 12 to 14 is nonmissing, then donor's value must match
LikC6	If recipient's count of household members aged 12 to 17 is nonmissing, then donor's value must match
LikC7	If recipient's count of household members aged 15 to 17 is nonmissing, then donor's value must match
LikC8	If recipient's count of household members aged 18 to 25 is nonmissing, then donor's value must match
LikC9	If recipient's count of household members aged 26 to 34 is nonmissing, then donor's value must match
LikC10	If recipient's count of household members aged 26 or older is nonmissing, then donor's value must match
LikC11	If recipient's count of household members aged 35 to 49 is nonmissing, then donor's value must match
LikC12	If recipient's count of household members aged 50 or older is nonmissing, then donor's value must match

Table E.157 Constraints and Portion of the Predictive Mean Vector for Parent-Child, Child Focus Multiplicity Counts

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases
1	Completely missing	125	None	1. PRED1	1-3,4,6,9,11,12: 62 2,3,4,6,9,11,12: 52 2,3,9,11,12: 6 2,10: 5

¹ PRED1 is the predicted mean for the number of parents of a child aged 12 to 20 who is a member of a parent-child pair.

Table E.158 Constraints and Portion of the Predictive Mean Vector for Parent-Child, Parent Focus Multiplicity Counts

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases
1	Completely missing	0	2	1. PRED1	No Cases

¹ PRED1 is the predicted mean for the number of children aged 12 to 20 belonging to a parent who is a member of a parent-child pair.

Table E.159 Constraints and Portion of the Predictive Mean Vector for Sibling-Sibling (12-14/15-17), Older Focus Multiplicity Counts

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases
1	Missing	3	3	1. PRED1	1,3-6: 3

¹ PRED1 is the predicted mean for the number of siblings aged 12 to 14 for a respondent aged 15 to 17 who is a member of a sibling-sibling pair.

Table E.160 Constraints and Portion of the Predictive Mean Vector for Sibling-Sibling (12-14/15-17), Younger Focus Multiplicity Counts

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases
1	Missing	2	4	1. PRED1	1,3-6: 1 3-6: 1

¹ PRED1 is the predicted mean for the number of siblings aged 15 to 17 for a respondent aged 12 to 14 who is a member of a sibling-sibling pair.

Table E.161 Constraints and Portion of the Predictive Mean Vector for Sibling-Sibling (12-17/18-25), Older Focus Multiplicity Counts

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases
1	Missing	1	5	1. PRED1	1,3,4,6,8: 0 3,4,6,8: 0 3,6: 1

¹ PRED1 is the predicted mean for the number of siblings aged 12 to 17 for a respondent aged 18 to 25 who is a member of a sibling-sibling pair.

Table E.162 Constraints and Portion of the Predictive Mean Vector for Sibling-Sibling (12-17/18-25), Younger Focus Multiplicity Counts

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases
1	Missing	3	6	1. PRED1	1,3,4,6,8: 1 3,4,6,8: 2

¹ PRED1 is the predicted mean for the number of siblings aged 18 to 25 for a respondent aged 12 to 17 who is a member of a sibling-sibling pair.

E.7.3 Stage Three: Household Counts

Table E.163 Variables Used in Constraints for Stage Three, Household Counts

Variable	Value	Constraint
Males XXXX1	0	No household members in the specified age group are male
	1	All household members in the specified age group are male
	2	Some household members in the specified age group are male
	3	There are no household members in the specified age group

¹ XXXX denotes the 4-digit age range (e.g., 1217 denotes the age range 12 to 17).

Table E.164 Logical Constraints for Stage Three, Household Counts

Constraint #	Logical Constraint
LogC1	Donor's count of older siblings aged 15 to 17 must fall within recipient's bounds
LogC2	Donor's count of older siblings aged 15 to 17 cannot be larger than recipient's nonmissing count of household members aged 15 to 17
LogC3	Donor's count of older siblings aged 18 to 25 must fall within recipient's bounds
LogC4	Donor's count of older siblings aged 18 to 25 cannot be larger than recipient's nonmissing count of household members aged 18 to 25
LogC5	Donor's count of spouse-spouse pairs must fall within recipient's bounds
LogC6	Donor's count of spouse-spouse pairs with children must fall within recipient's bounds
LogC7	Donor's parent-child (12-14) parent count must be greater than 0
LogC8	Donor's parent-child (12-17) parent count must be greater than 0
LogC9	Donor's parent-child (12-20) parent count must be greater than 0
LogC10	Donor's parent-child (12-14) child count must be greater than 0
LogC11	Donor's parent-child (12-17) child count must be greater than 0
LogC12	Donor's parent-child (12-20) child count must be greater than 0
LogC13	Donor's parent-child (12-14) child count must fall within recipient's bounds
LogC14	Donor's parent-child (12-17) child count must fall within recipient's bounds
LogC15	Donor's parent-child (12-20) child count must fall within recipient's bounds
LogC16	Donor's parent-child (12-14) parent count must fall within recipient's bounds
LogC17	Donor's parent-child (12-17) parent count must fall within recipient's bounds
LogC18	Donor's parent-child (12-20) parent count must fall within recipient's bounds

Table E.165 Likeness Constraints for Stage Three, Household Counts

Constraint #	Likeness Constraint
LikC1	Donor's predicted means each must be within 5 percent of recipient's predicted means
LikC2	If recipient lives in a multi-family home, then donor must also live in a multi-family home
LikC3	Donor's household size must be the same as recipient's household size
LikC4	If recipient's count of household members aged 0 to 11 is nonmissing, then donor's count must match
LikC5	If recipient's count of household members aged 12 to 14 is nonmissing, then donor's count must match
LikC6	If recipient's count of household members aged 12 to 17 is nonmissing, then donor's count must match
LikC7	If recipient's count of household members aged 12 to 20 is nonmissing, then donor's count must match
LikC8	If recipient's count of household members aged 15 to 17 is nonmissing, then donor's count must match
LikC9	If recipient's count of household members aged 15 or older is nonmissing, then donor's count must match
LikC10	If recipient's count of household members aged 18 to 25 is nonmissing, then donor's count must match
LikC11	If recipient's count of household members aged 26 to 34 is nonmissing, then donor's count must match
LikC12	If recipient's count of household members aged 26 to 49 is nonmissing, then donor's count must match
LikC13	If recipient's count of household members aged 35 to 49 is nonmissing, then donor's count must match
LikC14	If recipient's count of household members aged 50 or older is nonmissing, then donor's count must match
LikC15	Donor's and recipient's counts of household members aged 0 to 17 must both be 0, 1, or greater than 1
LikC16	Donor's and recipient's counts of household members aged 0 to 17 must either both be 0 or both be positive
LikC17	Donor's and recipient's counts of household members aged 12 to 14 must either both be 0 or both be positive
LikC18	Donor's and recipient's counts of household members aged 12 to 17 must either both be 0 or both be positive
LikC19	Donor's and recipient's counts of household members aged 15 to 17 must either both be 0 or both be positive
LikC20	Donor's and recipient's counts of household members aged 18 to 25 must either both be 0 or both be positive
LikC21	If recipient's value of MALES15P is nonmissing, then donor's value of MALES15P must match
LikC22	If recipient's count of household members aged 15 or older is positive, then donor and recipient must have the same count of household members aged 15 or older and the same values for MALES15P
LikC23	If recipient's value of MALES1825 is nonmissing, then donor's value of MALES1825 must match
LikC24	If recipient's value of MALES1834 is nonmissing, then donor's value of MALES1834 must match
LikC25	If recipient's value of MALES2634 is nonmissing, then donor's value of MALES2634 must match

Table E.165 Likeness Constraints for Stage Three, Household Counts (continued)

Constraint #	Likeness Constraint
LikC26	If recipient's value of MALES2649 is nonmissing, then donor's value of MALES2649 must match
LikC27	If recipient's value of MALES2649 is nonmissing and not equal to 3, then donor's value of MALES2649 must match
LikC28	If recipient's value of MALES3549 is nonmissing, then donor's value of MALES3549 must match
LikC29	If recipient's value of MALES50P is nonmissing, then donor's value of MALES50P must match
LikC30	If recipient could possibly have a parent-child (12-14) parent count of 0, then donor's and recipient's counts of household members aged 12 to 14 must match
LikC31	If recipient could possibly have a parent-child (12-17) parent count of 0, then donor's and recipient's counts of household members aged 12 to 17 must match
LikC32	If recipient could possibly have a parent-child (12-20) parent count of 0, then donor's and recipient's counts of household members aged 12 to 20 must match
LikC33	If recipient could possibly have a parent-child (12-14) child count of 0, then donor's and recipient's counts of household members aged 26 to 34, 35 to 49, and 50 or older must match
LikC34	If recipient could possibly have a parent-child (12-17) child count of 0, then donor's and recipient's counts of household members aged 26 to 34, 35 to 49, and 50 or older must match
LikC35	If recipient could possibly have a parent-child (12-20) child count of 0, then donor's and recipient's counts of household members aged 26 to 34, 35 to 49, and 50 or older must match
LikC36	If recipient could possibly have a parent-child (12-14) child count of 0, then donor's and recipient's counts of household members aged 26 to 49 must match
LikC37	If recipient could possibly have a parent-child (12-17) child count of 0, then donor's and recipient's counts of household members aged 26 to 49 must match
LikC38	If recipient could possibly have a parent-child (12-20) child count of 0, then donor's and recipient's counts of household members aged 26 to 49 must match
LikC39	If recipient's count of household members aged 12 to 14 is nonmissing, then donor and recipient must have the same screener count of household members aged 12 to 14
LikC40	If recipient's count of household members aged 12 to 17 is nonmissing, then donor and recipient must have the same screener count of household members aged 12 to 17
LikC41	If recipient's count of household members aged 15 to 17 is nonmissing, then donor and recipient must have the same screener count of household members aged 15 to 17
LikC42	If recipient's count of household members aged 18 to 25 is nonmissing, then donor and recipient must have the same screener count of household members aged 18 to 25
LikC43	Donor and recipient must have parent-child (12-14) parent counts that are both 0 or both positive
LikC44	Donor and recipient must have parent-child (12-17) parent counts that are both 0 or both positive
LikC45	If recipient's counts of household members aged 12 to 14 and 15 to 17 are both nonmissing, then donor and recipient must have parent-child (12-14 and 12-17) parent counts that are both 0 or both positive
LikC46	If recipient's counts of household members aged 12 to 17 and 18 to 25 are both nonmissing, then donor and recipient must have parent-child (12-17) parent counts that are both 0 or both positive

Table E.166 Constraints and Portion of the Predictive Mean Vector for Household Sibling Count (12-14/15-17), Older Focus

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases	
					Responding Pairs	Non-Pair Members
1	Missing	46	1,2	1. H1	1,3-5,8,39,41,43,44: 6 3-5,8,39,41,43,44: 7 5,8,39,41,43,44: 3 17,19,45: 2	1,3-5,8,39,41,43,44: 20 3-5,8,39,41,43,44: 3 5,8,39,41,43,44: 5

¹ H1 is the predicted mean count of older siblings aged 15 to 17.

Table E.167 Constraints and Portion of the Predictive Mean Vector for Household Sibling Count (12-17/18-25), Older Focus

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases	
					Responding Pairs	Non-Pair Members
1	Missing	87	3,4	1. H2	1,3,4,6,10,40,42,44: 30 3,4,6,10,40,42,44: 8 6,10,40,42,44: 6 8,20,46: 3	1,3,4,6,10,40,42,44: 24 3,4,6,10,40,42,44: 6 6,10,40,42,44: 7 18,20,46: 3

¹ H2 is the predicted mean count of older siblings aged 18 to 25.

Table E.168 Constraints and Portion of the Predictive Mean Vector for Household Spouse-Spouse Count

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases	
					Responding Pairs	Non-Pair Members
1	Missing	97	5	1. SPOUSE1 2. SPOUSE2	1-3,9-11,13-15,21,23,25,28,29: 8 2,3,9-11,13-15,21,23,25,28,29: 25 2,15,22,23,25,28,29: 14 2,22,24,26,29: 7	1-3,9-11,13-15,21,23,25,28,29: 17 2,3,9-11,13-15,21,23,25,28,29: 11 2,15,22,23,25,28,29: 3 2,22,24,26,29: 8 2,22,27: 4

¹ The predictive mean vector components are defined by the following:

1. SPOUSE1 = P(the number of spouse-spouse pairs in the household is 0)
2. SPOUSE2 = P(the number of spouse-spouse pairs in the household is 1)

Table E.169 Constraints and Portion of the Predictive Mean Vector for Household Spouse-Spouse with Children Count

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases	
					Responding Pairs	Non-Pair Members
1	Missing	46	6	1. SPOUSE3	1-3,9-11,13-15,21,23,25,28,29: 6 2,3,9-11,13-15,21,23,25,28,29: 5 2,15,22,23,25,28,29: 9 2,16,22,24,26,29: 3 2,16,22,27: 2	1-3,9-11,13-15,21,23,25,28,29: 11 2,3,9-11,13-15,21,23,25,28,29: 6 2,15,22,23,25,28,29: 0 2,16,22,24,26,29: 0 2,16,22,27: 3 2,16,22: 0 2,16,22*: 1

*A donor could not be found for these cases. Imputed values were randomly assigned within the bounds created during the editing process.

¹ The predictive mean vector components are defined by the following:

1. SPOUSE3 = P(there is at least one spouse-spouse pair with children in the household)

Table E.170 Constraints and Portion of the Predictive Mean Vector for Household Parent-Child Counts

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases	
					Responding Pairs	Non-Pair Members
1	Missing PPCP1220	12	9,18	1. PC1 2. PC2 3. H3	1-4,10,11,13,14,32: 1 2-4,10,11,13,14,32: 10	1-4,10,11,13,14,32: 0 2-4,10,11,13,14,32: 1
2	Missing PPCP1217	2	8,17	1. PC1 2. PC2 3. H3	1-4,10,11,13,14,31: 1 2-4,10,11,13,14,31: 1	No Cases
3	Missing PPCP1217 and PPCP1220	20	8,9,17,18	1. PC1 2. PC2 3. H3	1-4,10,11,13,14,31,32: 4 2-4,10,11,13,14,31,32: 14 2,4,10,11,13,14,31,32: 0 2,12,31,32: 1	1-4,10,11,13,14,31,32: 0 2-4,10,11,13,14,31,32: 0 2,4,10,11,13,14,31,32: 0 2,12,31,32: 1
4	Missing PPCP1214	2	7,16	1. PC1 2. PC2 3. H3	1-4,10,11,13,14,30: 0 2-4,10,11,13,14,30: 2	No Cases
5	Missing PPCP1214 and PPCP1220	0	7,9,16,18	1. PC1 2. PC2 3. H3	No Cases	No Cases

Table E.170 Constraints and Portion of the Predictive Mean Vector for Household Parent-Child Counts (continued)

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases	
					Responding Pairs	Non-Pair Members
6	Missing PPCP1214 and PPCP1217	6	7,8,16,17	1. PC1 2. PC2 3. H3	1-4,10,11,13,14,30,31: 0 2-4,10,11,13,14,30,31: 6	No Cases
7	Missing PPCP1214, PPCP1217, and PPCP1220	40	7-9,16-18	1. PC1 2. PC2 3. H3	1-4,10,11,13,14,30-32: 1 2-4,10,11,13,14,30-32: 33 2,4,10,11,13,14,30-32: 3 2,12,30-32: 2	1-4,10,11,13,14,30-32: 0 2-4,10,11,13,14,30-32: 0 2,4,10,11,13,14,30-32: 0 2,12,30-32: 1
8	Missing PPCC1220	14	12,15	1. PC1 2. PC2 3. H3	1-4,7,10,35: 1 2-4,7,10,35: 7 2,4,7,10,35: 4 2,7,38: 1	1-4,7,10,35: 0 2-4,7,10,35: 1
9	Missing PPCC1220 and PPCP1220	16	15,18	1. PC1 2. PC2 3. H3	1-4,7,10,11,13,14,32,35: 0 2-4,7,10,11,13,14,32,35: 10 2,4,7,10,11,13,14,32,35: 1 2,7,12,32,38: 2	1-4,7,10,11,13,14,32,35: 1 2-4,7,10,11,13,14,32,35: 2
10	Missing PPCC1220 and PPCP1217	0	8,12,15,17	1. PC1 2. PC2 3. H3	No Cases	No Cases
11	Missing PPCC1220, PPCP1217, and PPCP1220	1	8,15,17,18	1. PC1 2. PC2 3. H3	1-4,7,10,11,13,14,31,32,35: 0 2-4,7,10,11,13,14,31,32,35: 1	No Cases
12	Missing PPCC1220 and PPCP1214	0	7,12,15,16	1. PC1 2. PC2 3. H3	No Cases	No Cases
13	Missing PPCC1220, PPCP1214, and PPCP1220	0	7,15,16,18	1. PC1 2. PC2 3. H3	No Cases	No Cases
14	Missing PPCC1220, PPCP1214, and PPCP1217	0	7,8,12,15-17	1. PC1 2. PC2 3. H3	No Cases	No Cases
15	Missing PPCC1220, PPCP1214, PPCP1217, and PPCP1220	1	7,8,15-18	1. PC1 2. PC2 3. H3	No Cases	1-4,7,10,11,13,14,30-32,35: 0 2-4,7,10,11,13,14,30-32,35: 1

Table E.170 Constraints and Portion of the Predictive Mean Vector for Household Parent-Child Counts (continued)

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases	
					Responding Pairs	Non-Pair Members
16	Missing PPCC1217	0	11,14	1. PC1 2. PC2 3. H3	No Cases	No Cases
17	Missing PPCC1217 and PPCP1220	0	9,11,14,18	1. PC1 2. PC2 3. H3	No Cases	No Cases
18	Missing PPCC1217 and PPCP1217	0	14,17	1. PC1 2. PC2 3. H3	No Cases	No Cases
19	Missing PPCC1217, PPCP1217, and PPCP1220	0	9,14,17,18	1. PC1 2. PC2 3. H3	No Cases	No Cases
20	Missing PPCC1217 and PPCP1214	0	7,11,14,16	1. PC1 2. PC2 3. H3	No Cases	No Cases
21	Missing PPCC1217, PPCP1214, and PPCP1220	0	7,9,11,14,16, 18	1. PC1 2. PC2 3. H3	No Cases	No Cases
22	Missing PPCC1217, PPCP1214, and PPCP1217	0	7,14,16,17	1. PC1 2. PC2 3. H3	No Cases	No Cases
23	Missing PPCC1217, PPCP1214, PPCP1217, and PPCP1220	0	7,9,14,16-18	1. PC1 2. PC2 3. H3	No Cases	No Cases
24	Missing PPCC1217 and PPCC1220	9	11,12,14,15	1. PC1 2. PC2 3. H3	1-4,6,7,10,34,35: 1 2-4,6,7,10,34,35: 3 2,4,6,7,10,34,35: 1 2,6,7,37,38: 2	1-4,6,7,10,34,35: 0 2-4,6,7,10,34,35: 2
25	Missing PPCC1217, PPCC1220, and PPCP1220	0	11,14,15,18	1. PC1 2. PC2 3. H3	No Cases	No Cases

Table E.170 Constraints and Portion of the Predictive Mean Vector for Household Parent-Child Counts (continued)

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases	
					Responding Pairs	Non-Pair Members
26	Missing PPCC1217, PPCC1220, and PPCP1217	1	12,14,15,17	1. PC1 2. PC2 3. H3	1-4,6,7,10,11,13,14,31,34,35: 0 2-4,6,7,10,11,13,14,31,34,35: 0 2,4,6,7,10,11,13,14,31,34,35: 0 2,6,7,12,31,37,38: 1	No Cases
27	Missing PPCC1217, PPCC1220, PPCP1217, and PPCP1220	28	14,15,17,18	1. PC1 2. PC2 3. H3	1-4,6,7,10,11,13,14,31,32,34,35: 0 2-4,6,7,10,11,13,14,31,32,34,35: 17 2,4,6,7,10,11,13,14,31,32,34,35: 0 2,6,7,12,31,32,37,38: 3	1-4,6,7,10,11,13,14,31,32,34,35: 1 2-4,6,7,10,11,13,14,31,32,34,35: 6 2,4,6,7,10,11,13,14,31,32,34,35: 0 2,6,7,12,31,32,37,38: 1
28	Missing PPCC1217, PPCC1220, and PPCP1214	0	7,11,12,14-16	1. PC1 2. PC2 3. H3	No Cases	No Cases
29	Missing PPCC1217, PPCC1220, PPCP1214, and PPCP1220	0	7,11,14-16,18	1. PC1 2. PC2 3. H3	No Cases	No Cases
30	Missing PPCC1217, PPCC1220, PPCP1214, and PPCP1217	0	7,12,14-17	1. PC1 2. PC2 3. H3	No Cases	No Cases
31	Missing PPCC1217, PPCC1220, PPCP1214, PPCP1217, and PPCP1220	1	7,14-18	1. PC1 2. PC2 3. H3	1-4,6,7,10,11,13,14,30-32,34,35: 0 2-4,6,7,10,11,13,14,30-32,34,35: 1	No Cases
32	Missing PPCC1214	0	10,13	1. PC1 2. PC2 3. H3	No Cases	No Cases
33	Missing PPCC1214 and PPCP1220	0	9,10,13,18	1. PC1 2. PC2 3. H3	No Cases	No Cases
34	Missing PPCC1214 and PPCP1217	0	8,10,13,17	1. PC1 2. PC2 3. H3	No Cases	No Cases
35	Missing PPCC1214, PPCP1217, and PPCP1220	0	8-10,13,17,18	1. PC1 2. PC2 3. H3	No Cases	No Cases

Table E.170 Constraints and Portion of the Predictive Mean Vector for Household Parent-Child Counts (continued)

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases	
					Responding Pairs	Non-Pair Members
36	Missing PPCC1214 and PPCP1214	3	13,16	1. PC1 2. PC2 3. H3	1-5,10,11,13,14,30,33: 0 2-5,10,11,13,14,30,33: 1 2,4,5,10,11,13,14,30,33: 2	No Cases
37	Missing PPCC1214, PPCP1214, and PPCP1220	0	9,13,16,18	1. PC1 2. PC2 3. H3	No Cases	No Cases
38	Missing PPCP1214 and PPCP1217	0	8,13,16,17	1. PC1 2. PC2 3. H3	No Cases	No Cases
39	Missing PPCP1214, PPCP1217, and PPCP1220	0	8,9,13,16-18	1. PC1 2. PC2 3. H3	No Cases	No Cases
40	Missing PPCC1214 and PPCC1220	0	10,12,13,15	1. PC1 2. PC2 3. H3	No Cases	No Cases
41	Missing PPCC1214, PPCC1220, and PPCP1220	0	10,13,15,18	1. PC1 2. PC2 3. H3	No Cases	No Cases
42	Missing PPCC1214, PPCC1220, and PPCP1217	0	8,10,12,13,15,17	1. PC1 2. PC2 3. H3	No Cases	No Cases
43	Missing PPCC1214, PPCC1220, PPCP1217, and PPCP1220	0	8,10,13,15,17,18	1. PC1 2. PC2 3. H3	No Cases	No Cases
44	Missing PPCC1214, PPCC1220, and PPCP1214	0	12,13,15,16	1. PC1 2. PC2 3. H3	No Cases	No Cases
45	Missing PPCC1214, PPCC1220, PPCP1214, and PPCP1220	0	13,15,16,18	1. PC1 2. PC2 3. H3	No Cases	No Cases

Table E.170 Constraints and Portion of the Predictive Mean Vector for Household Parent-Child Counts (continued)

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases	
					Responding Pairs	Non-Pair Members
46	Missing PPCC1214, PPCC1220, PPCP1214, and PPCP1217	0	8,12,13,15-17	1. PC1 2. PC2 3. H3	No Cases	No Cases
47	Missing PPCC1214, PPCC1220, PPCP1214, PPCP1217, and PPCP1220	0	8,13,15-18	1. PC1 2. PC2 3. H3	No Cases	No Cases
48	Missing PPCC1214 and PPCC1217	0	10,11,13,14	1. PC1 2. PC2 3. H3	No Cases	No Cases
49	Missing PPCC1214, PPCC1217, and PPCP1220	0	9-11,13,14,18	1. PC1 2. PC2 3. H3	No Cases	No Cases
50	Missing PPCC1214, PPCC1217, and PPCP1217	0	10,13,14,17	1. PC1 2. PC2 3. H3	No Cases	No Cases
51	Missing PPCC1214, PPCC1217, PPCP1217, and PPCP1220	0	9,10,13,14, 17,18	1. PC1 2. PC2 3. H3	No Cases	No Cases
52	Missing PPCC1214, PPCC1217, and PPCP1214	0	11,13,14,16	1. PC1 2. PC2 3. H3	No Cases	No Cases
53	Missing PPCC1214, PPCC1217, PPCP1214, and PPCP1220	0	9,11,13,14, 16,18	1. PC1 2. PC2 3. H3	No Cases	No Cases
54	Missing PPCC1214, PPCC1217, PPCP1214, and PPCP1217	0	13,14,16,17	1. PC1 2. PC2 3. H3	No Cases	No Cases

Table E.170 Constraints and Portion of the Predictive Mean Vector for Household Parent-Child Counts (continued)

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases	
					Responding Pairs	Non-Pair Members
55	Missing PPCC1214, PPCC1217, PPCP1214, PPCP1217, and PPCP1220	1	9,13,14,16-18	1. PC1 2. PC2 3. H3	1-6,10,11,13,14,30-34: 0 2-6,10,11,13,14,30-34: 0 2,4-6,10,11,13,14,30-34: 0 2,5,6,12,30-32,36,37: 1	No Cases
56	Missing PPCC1214, PPCC1217, and PPCC1220	9	10-15	1. PC1 2. PC2 3. H3	1-7,10,33-35: 0 2-7,10,33-35: 7 2,4-7,10,33-35: 1	1-7,10,33-35: 0 2-7,10,33-35: 1
57	Missing PPCC1214, PPCC1217, PPCC1220, and PPCP1220	0	10,11,13-15, 18	1. PC1 2. PC2 3. H3	No Cases	No Cases
58	Missing PPCC1214, PPCC1217, PPCC1220, and PPCP1217	0	10,12-15,17	1. PC1 2. PC2 3. H3	No Cases	No Cases
59	Missing PPCC1214, PPCC1217, PPCC1220, PPCP1217, and PPCP1220	0	10,13-15,17, 18	1. PC1 2. PC2 3. H3	No Cases	No Cases
60	Missing PPCC1214, PPCC1217, PPCC1220, and PPCP1214	4	11-16	1. PC1 2. PC2 3. H3	1-7,10,11,13,14,30,33-35: 0 2-7,10,11,13,14,30,33-35: 2 2,4-7,10,11,13,14,30,33-35: 0 2,5-7,12,30,36-38: 2	No Cases
61	Missing PPCC1214, PPCC1217, PPCC1220, PPCP1214, and PPCP1220	0	11,13-16,18	1. PC1 2. PC2 3. H3	No Cases	No Cases
62	Missing PPCC1214, PPCC1217, PPCC1220, PPCP1214, and PPCP1217	1	12-17	1. PC1 2. PC2 3. H3	1-7,10,11,13,14,30,31,33-35: 0 2-7,10,11,13,14,30,31,33-35: 1	No Cases

Table E.170 Constraints and Portion of the Predictive Mean Vector for Household Parent-Child Counts (continued)

#	Missingness Pattern	Total Number of Cases	Logical Constraints	Predictive Mean Vector ¹	Likeness Constraints: Number of Cases	
					Responding Pairs	Non-Pair Members
63	Missing PPCC1214, PPCC1217, PPCC1220, PPCP1214, PPCP1217, and PPCP1220	29	13-18	1. PC1 2. PC2 3. H3	1-7,10,11,13,14,30-35: 1 2-7,10,11,13,14,30-35: 11 2,4-7,10,11,13,14,30-35: 0 2,5-7,12,30-32,36-38: 11 2,5-7,30-32: 1	1-7,10,11,13,14,30-35: 1 2-7,10,11,13,14,30-35: 4

¹ The predictive mean vector components are defined by the following:

1. H3 = Predicted mean count of children aged 12 to 20 with parents in the household
2. PC1 = P(there is one parent with at least one child aged 12 to 20 in the household)
3. PC2 = P(there are two parents with at least one child aged 12 to 20 in the household)

Appendix F: Quality Control Measures Used in the Imputation Procedures

This page intentionally left blank

Appendix F: Quality Control Measures Used in the Imputation Procedures

F.1 Introduction

For the 2016 National Survey on Drug Use and Health (NSDUH), several quality control (QC) measures were implemented for the imputation of demographic, substance use, substance dependence and abuse, income, health insurance, nicotine dependence, household composition (roster), and pair variables. These QC measures spanned the following three basic steps within the predictive mean neighborhood (PMN) and modified predictive mean neighborhood (modPMN) imputation methodology: (1) response propensity adjustment, (2) prediction modeling, and (3) imputation.

Checklists for the imputation of these variables were used during imputation processing for the 2016 survey and serve as formal documentation of the QC measures that were implemented. Most of these checks required both a technician check, where the individual who ran the computer program (technician) entered his or her name and the date the check was performed, and an independent reviewer check, where another person performed a check of the same items. The reviewer check ensured greater quality in the imputation procedures. Additionally, some checklist entries required the technician to document the procedures that were taken to run the programs, such as listing the variables that were dropped from the model in order to achieve model convergence.

The checks involved in each of the steps are described in detail in the following sections. Variable-specific consistency checks were often created in addition to the checks of the PMN and modPMN steps. However, they are not discussed in this appendix. The QC measures that were used in the editing process or the delivery of variables to other NSDUH teams also are not discussed.

F.2 Step 1: Response Propensity Adjustment

The first step of the PMN and modPMN methodologies is a response propensity adjustment. The item response propensity is modeled as a function of a predetermined set of covariates. The model can be thought of as a special case of the generalized exponential model (GEM)¹ developed for weighting procedures, in that imputations that are done at the item level are similar in nature to the weight adjustments made for entire units.

For PMN, the sampling weights are adjusted for item nonresponse so that the item respondent weights are representative of the entire domain of interest. These adjusted sampling weights are used in the subsequent steps, prediction modeling and hot-deck imputation. For modPMN, the adjusted sampling weights are not used in subsequent steps; however, the

¹ The GEM macro, which was written in SAS/IML® software, was developed at RTI International for weighting procedures and is described in detail in Appendix A of the person-level sampling weight calibration report in the 2016 NSDUH methodological resource book (Center for Behavioral Health Statistics and Quality, 2018a).

respondents and nonrespondents are segregated into five bins of roughly equal size based on their predicted response propensity from the model. The prediction modeling and imputation are then performed separately for each bin.

The following QC measures were conducted in Step 1:

- The output of the response propensity modeling program was checked for singularities. Any singularities that occurred were investigated, and the model was corrected by removing the correlated covariates from the model.
- The values of the convergence flags were checked to see whether the GEM converged.
- The current distance to root was checked to make sure it was close to 0, indicating that GEM converged.
- When variables were reduced from the original model, the remaining levels of variables were checked to ensure appropriateness. An example of this check was to determine whether the base variables or lower order terms were present when interactions or higher order terms existed (e.g., "age" and "age squared" must be in the model if "age cubed" is in the model).
- An indicator was calculated in the response propensity program that measured the maximum adjustment to the weights. In most cases, the adjusted weights resembled the original design weights. If the maximum adjustment was too high (usually greater than 3.00), this was likely due to an overspecified model, where the adjustment was not performing at an optimum level. Large maximum adjustments were investigated and corrected, if possible, by removing extraneous variables from the models so that any final weight adjustment applied to a respondent was within acceptable bounds.
- After the weights were adjusted, the ratio of the maximum adjusted weight to the mean adjusted weight (called the "mmratio") was computed to monitor the variation among the weights. Any mmratio value that was greater than 25 percent was noted in the response propensity program checklist.
- The unequal weighting effect (UWE) was checked before and after adjustment to ensure that there was no significant variance increase due to the nonresponse adjustment. The difference in the UWE after-adjustment value should be no more than 20 percent of the UWE before-adjustment value. The difference was fairly small in most cases, and any difference greater than 20 percent was investigated and corrected, if possible.
- When using the SAS procedure PROC MEANS to examine summary statistics, the weighted totals for the independent variables in the model were compared both before and after the adjustment. If these weighted totals were equal, the adjustment procedures worked properly.
- For modPMN, the five bins were checked to ensure that approximately 20 percent of the domain members was assigned to each bin and that the larger the bin number, the larger the number of nonrespondents (i.e., a larger bin number indicates a lower estimated response propensity from the model).

F.3 Step 2: Prediction Modeling

For PMN, the prediction modeling step, using the adjusted weights from Step 1, calculates predicted means, which are then used in the hot-deck step(s) to create neighborhoods and select donors. For modPMN, the prediction modeling step is conducted separately within each of the five response propensity bins from Step 1 using the unadjusted sampling weights (i.e., PANALWT or ANALWT). The predicted means from this step are then used in the imputation step. The dependent variable in the model is usually the variable that is undergoing imputation, or some transformation of that variable. Each model is built using only those cases within the domain with complete responses for that item. Predicted means are then calculated for all of the domain members, whether or not they were item respondents, using the values for the covariates and the estimates for the regression coefficients.

For categorical outcome variables, logistic regression models are used for the prediction models. For continuous variables, linear regression models are fit. For count variables, Poisson regression models are used. For response variables that are proportions (e.g., months on welfare), a logit transformation is applied to the proportion, and a linear regression model is utilized. For continuous nonnegative random variables (e.g., finer categories of income), a failure time model is used.

The models are discussed in detail in the imputation sections of Chapters 4 through 7 and Chapters 11 through 14 of this report. The following QC measures were conducted in Step 2:

- All models were checked for singularities and collinearities. Any singularities that occurred were investigated, and the model was corrected.
- For Poisson regression models, failure time models, and logistic models, the output was checked to ensure convergence had occurred.
- The log file was checked for "data warning" messages or other SUDAAN®-specific errors.² Under PMN, if there was a "data warning" message in the log, independent variables were dropped until the modeling procedure in SUDAAN produced no warning messages. Under modPMN, many of these warning messages are ignored under the following conditions: the algorithm converges, reasonable estimates of regression coefficients are produced, and reasonable test statistics associated with the regression coefficients are produced. Similar to the response propensity model, if the main variable was dropped, its interaction variables also were dropped.
- When variables were reduced from the original model, the remaining levels of variables were checked to ensure appropriateness.
- Output was checked to verify that everything worked properly in the regression model.
- A check on the predicted means from the model was created to ensure that each respondent in the domain had a valid nonmissing predicted mean.

² Details can be found in the *SUDAAN® Language Manual, Release 11.0.1* (RTI International, 2013).

F.4 Step 3: Imputation

The predicted means from Step 2 were used to determine the final assignments of imputed values in Step 3. Final imputed values were assigned using either a hot-deck step to select a donor or a collocated stochastic imputation process using the nonrespondent's predicted mean(s) from Step 2 (see Chapter 3 for more details).

The QC measures in this step had two objectives: (1) ensure that the imputed values were consistent with preexisting nonmissing values, and (2) ensure that the imputed values were assigned as intended. Many variables had logical constraints or restrictions placed on the final imputed value that had to be adhered to in order to prevent logical inconsistencies. Specific QC checks were implemented to ensure this; however, only the more general checks that apply to all imputations are listed in this appendix. For more detail on logical constraints and restrictions placed on imputed values, refer to Appendix E. The following QC measures were conducted in Step 3:

- Unusual imputed values were noted. If the imputed value was equivalent to one of the standard NSDUH missing value codes, this signaled a failure in the imputation process (e.g., a suitable donor could not be found), and measures were required (e.g., likeness constraints were loosened) to revise the programs so that a final imputed value could be assigned. If the imputed value was otherwise unusual, the imputation process was examined to ensure that no errors occurred.
- For variables imputed using a hot-deck step, the number of cases that were imputed within various levels of restrictiveness of the likeness constraints (as determined by the variable SMALLFLG) was noted.³
- For variables imputed using a hot-deck step, the frequency of the variable "WORKED" was checked to ensure that no values were equal to 0. Values greater than 0 signified that the imputation procedure was able to find a donor for all missing cases.
- The distribution of edited variables was compared with the distribution of imputed variables to make sure that each imputed value was within the appropriate range corresponding to the value of the edited variable.
- The imputed values were crossed with the imputation indicators to ensure that the indicators were created correctly.
- After imputation was implemented, the distribution of values for nonrespondents was checked against the distribution of values for all respondents to ensure the similarity of these two subgroups.
- It was necessary to ensure that everyone to whom the variable did not apply received a skip code for the final imputed variable. For example, all those in the 12-to-14 age group should have a "Not Applicable (N/A)" value of 99 for the imputation-revised marital status variable IRMARITSTAT.

³ Refer to Appendix E for more details about likeness constraints and the "SMALLFLG" variable.

- Any missing values were noted. This occurred when the program was unsuccessful in assigning a valid imputed value.
- Any cases where the imputed value was not consistent with preexisting nonmissing values were noted.
- It was necessary to ensure that any restrictions on the final imputed value for a given nonrespondent were honored. For example, some respondents were known to be employed, but either full-time or part-time employment status was not known. Checks were conducted to ensure that these respondents had either full-time or part-time status assigned to the employment status variable (IRWKRSTAT18), but not unemployed or other statuses.
- When multiple missingness patterns existed or different restrictions were placed on final imputed values, each pattern was treated separately, and the distribution of imputed values within each missingness pattern, or restricted imputation class, was investigated. For example, if it was known that a respondent was a past year user, both past month and past year users should have been included among the imputed values, not just past month users.

This page intentionally left blank

Appendix G: Interviewer Explanations for Overrides to Consistency Checks in Household Roster

This page intentionally left blank

Appendix G: Interviewer Explanations for Overrides to Consistency Checks in Household Roster

G.1 Introduction

In the household roster for the 2016 National Survey on Drug Use and Health (NSDUH), the interviewer was supposed to enter a roster of the respondent's entire household, which included age, gender, and the relationship to the respondent. It was not uncommon for the interviewer to enter a relationship code, age, or gender that was illogical, based on the age and gender of the respondent given in the core part of the questionnaire. Before the computer-assisted interviewing (CAI) instrument was implemented in the 1999 NSDUH, such responses would have been flagged at the data processing stage.¹ However, beginning with the 2000 survey and in every survey year since then, consistency checks have been added to the CAI instrument that allowed the interviewer, if needed, to correct the error while giving the interview. Details about these consistency checks are presented in Chapter 8.

In general, two types of consistency checks were implemented in the 2016 survey. The first type compared the entry in the roster with previously entered questionnaire information, specifically the respondent's age (CURNTAGE) and gender. The second type checked for internal consistency within the household roster. In some cases, a consistency check would be triggered even though the response was legitimate. This occurred if CURNTAGE was considered incorrect, for example, or in extremely rare family situations such as when a stepmother was younger than her stepson. With the exception of the check against the previously entered respondent's gender, the interviewer could override the consistency check and explain why the response given was correct. In some cases, the interviewer was correct in overriding the consistency check. In others, however, it was clear that the interviewer misunderstood how the roster should have been put together and that the override to the consistency check was not legitimate.

This appendix summarizes the explanations given by interviewers for consistency check overrides in the household roster. It is divided into two parts: consistency check overrides involving CURNTAGE and those involving internal consistency checks.

G.2 Override Comments from Interviewers: Comparisons with CURNTAGE

When an interviewer entered the respondent's roster entry (the "self" entry), if the age did not match the age previously entered in the questionnaire, a consistency check was triggered. The comparison was between the roster age for the "self" and CURNTAGE, which was the value of the final questionnaire-edited age (AGE) stored by Blaise.² Explanations for consistency check overrides for the variable CURNTAGE are provided in [Table G.1](#). Because CURNTAGE had the potential to change constantly throughout the questionnaire, no final variable with this

¹ Because the age and gender of the respondent given in the core part of the questionnaire were not allowed to change, the relationship code and sometimes the age of the roster member were set to bad data.

² The Blaise program is the computer program within the CAI instrument that was used to direct the respondent and interviewer through the questionnaire.

name was created. However, in most cases, the value of CURNTAGE when the roster commenced was equivalent to NEWAGE, the value of CURNTAGE after the drug modules had been completed. In theory, NEWAGE was not always equivalent to AGE, the derivation of which is described in Chapter 4.

In the 2002 survey, the explanations provided in [Table G.1](#) were not reviewed when determining AGE, nor were they reviewed when determining the final value of the age for the "self" entry in the roster. However, beginning with the 2003 survey, these explanations have been carefully reviewed. In rare cases, the final value of AGE was set to the age of the self in the questionnaire roster (the "roster age"), based on these explanations as well as other evidence, even if it disagreed with the age as it would have been calculated in prior survey years. Details about this process are provided in Chapter 4.

Even in cases where the explanation seemed clear that CURNTAGE was incorrect, the value of AGE was not always set to the roster age. In most cases, this was because the difference between CURNTAGE and the roster age was 1 year or less. A difference of 1 year was tolerated, because legitimate differences could result from a birthday occurring in the time between the drug modules and the roster.³ In other situations, the value of CURNTAGE was incorrect, but the original questionnaire-edited age was correct, so no change was necessary. In still other cases, not all the criteria that were necessary for changing the value of AGE to be equal to the roster age were met. Cases where the value of AGE was changed to roster age are denoted in the "Comment" column in bolded italics in [Table G.1](#). Otherwise, the reason for not changing the value of AGE to roster age also is shown in this column. The "Respondent's Age in Roster of Other Pair Member" column indicates whether the roster of the other pair member, if it existed, supported CURNTAGE or the override age as the respondent's age.

G.3 Override Comments from Interviewers: Internal Consistency Check Overrides

Internal consistency checks were performed on the household roster for the 2016 NSDUH. Interviewer explanations for overrides to these internal consistency checks were evaluated individually to determine their legitimacy. In the 2016 NSDUH, there were 556 interviewer explanations for overrides to these internal consistency checks. Of these, 540 were considered legitimate; no edit was applied to the age or relationship code of the roster member. The remaining 16 explanations were not considered legitimate; thus, the override was overruled, and the relationship code (and sometimes the roster member's age) was altered accordingly. Interviewer explanations for overrides to these internal consistency checks not considered legitimate are provided in [Table G.2](#). Two of these 16 explanations were the result of one roster that flagged two consistency checks with separate interviewer explanations, but only one of these consistency checks was overruled. Also provided in this table are the questionnaire-edited age of the respondent (AGE), the age and relationship to the respondent of the roster member in question, and, in the "Comment" column, a brief indication of the probable true relationship of the roster member to the respondent.

³ It was not uncommon for an interview to be conducted in more than one sitting. This could have occurred if either the respondent or the interviewer did not have enough time for the interview or otherwise could not complete the interview in a single sitting.

Table G.1 Explanations for Overrides to Consistency Checks against CURNTAGE

#	NEW AGE	Original Roster Age for Self	Screener Age	Verbatim Explanation from Field Interviewers ¹	Respondent's Age in Roster of Other Pair Member	Comment ²	AGE = Final Roster Age
1	50	50	50	R says she is 47 but the math says 46, I verified dob and its correct. R still says 47	Not in a pair	<i>AGE was changed to equal roster age</i>	50
2	67	67	67	correct age is 68. could not get computer to put in correct age based on year respondent wanted to use,	Not in a pair	<i>AGE was changed to equal roster age</i>	67
3	29	30	30	no error	Not in a pair	Diff. ≤ 1 year	29
4	33	32	32	Wen asked on last birthday said 33	Not in a pair	Diff. ≤ 1 year	33
5	43	42	43	some typo mistake, he is 42	Not in a pair	Diff. ≤ 1 year	43
6	66	67	67	66	66	Diff. ≤ 1 year	66
7	33	32	32	will be 33 in november	Not in a pair	Diff. ≤ 1 year	33
8	77	78	73	this is the right ages before she gave incorrect	Not in a pair	Diff. ≤ 1 year	77
9	61	62	62	I am just entering the info given by respondent	Not in a pair	Diff. ≤ 1 year	61
10	33	32	33	age error from ir	33	Diff. ≤ 1 year	33
11	45	44	45	error on front end 44 is correct	45	Diff. ≤ 1 year	45
12	22	21	21	21 on last bday. currently 22	Not in a pair	Diff. ≤ 1 year	22
13	30	29	30	correct age is thirty	Not in a pair	Diff. ≤ 1 year	30
14	19	18	18	female is 18 no error	19	Diff. ≤ 1 year	19
15	77	76	77	son gave the wrong age when screening	Not in a pair	Diff. ≤ 1 year	77
16	22	21	21	mistake	22	Diff. ≤ 1 year	22
17	37	36	36	respondent was wrong about ages	37	Diff. ≤ 1 year	37
18	31	32	32	r wrong answer	Not in a pair	Diff. ≤ 1 year	31
19	51	52	50	52 self	Not in a pair	Diff. ≤ 1 year	51
20	16	17	17	resp is 16 rostered wrong	16	Diff. ≤ 1 year	16
21	32	33	33	wrong age	Not in a pair	Diff. ≤ 1 year	32
22	48	47	49	R said last birthday 47 but actually 48	Not in a pair	Diff. ≤ 1 year	48
23	63	62	62	R said correct age is 63	Not in a pair	Diff. ≤ 1 year	63
24	16	15	16	per respondent	16	Diff. ≤ 1 year	16

Table G.1 Explanations for Overrides to Consistency Checks against CURNTAGE (continued)

#	NEW AGE	Original Roster Age for Self	Screener Age	Verbatim Explanation from Field Interviewers ¹	Respondent's Age in Roster of Other Pair Member	Comment ²	AGE = Final Roster Age
25	51	50	50	wrong ag	Not in a pair	Diff. ≤ 1 year	51
26	46	47	47	R unsure of age born in 1968 but gave 3 different 46, 47 , 48	46	Diff. ≤ 1 year	46
27	38	37	37	householder is 37	38	Diff. ≤ 1 year	38
28	44	45	45	R confused	Not in a pair	Diff. ≤ 1 year	44
29	30	31	30	Younger person interviewed	30	Diff. ≤ 1 year	30
30	63	62	62	wife	Not in a pair	Diff. ≤ 1 year	63
31	48	47	47	47yo male completed the interview	48	Diff. ≤ 1 year	48
32	45	44	44	R is 44	45	Diff. ≤ 1 year	45
33	46	45	45	IR is 45	Not in a pair	Diff. ≤ 1 year	46
34	20	19	20	repeated question and that is the response given by R. however, parents confirmed 19 yr old is brother or R b.	20	Diff. ≤ 1 year	20
35	22	21	22	not an error	Not in a pair	Diff. ≤ 1 year	22
36	32	31	31	Made a mistake	Not in a pair	Diff. ≤ 1 year	32
37	21	20	20	This is the R interviewing that is 20 yrs old	21	Diff. ≤ 1 year	21
38	67	68	68	just mispoke	Not in a pair	Diff. ≤ 1 year	67
39	42	43	43	husband gave wrong age	Not in a pair	Diff. ≤ 1 year	42
40	22	23	23	IR	22	Diff. ≤ 1 year	22
41	45	44	44	R gave wrong age	Not in a pair	Diff. ≤ 1 year	45
42	31	32	32	she said 32	31	Diff. ≤ 1 year	31
43	22	21	22	r's age is 22	Not in a pair	Diff. ≤ 1 year	22
44	14	13	14	r is giving age of last birthday, previous to last	14	Diff. ≤ 1 year	14
45	14	15	14	14	Not in a pair	Diff. ≤ 1 year	14
46	38	39	39	Respondant said he just turned 39	Not in a pair	Diff. ≤ 1 year	38
47	33	32	32	had a birthday	33	Diff. ≤ 1 year	33
48	15	14	14	don't remember why 15 was entered	15	Diff. ≤ 1 year	15
49	31	30	31	30	31	Diff. ≤ 1 year	31
50	43	44	43	44 next week	43	Diff. ≤ 1 year	43

Table G.1 Explanations for Overrides to Consistency Checks against CURNTAGE (continued)

#	NEW AGE	Original Roster Age for Self	Screener Age	Verbatim Explanation from Field Interviewers ¹	Respondent's Age in Roster of Other Pair Member	Comment ²	AGE = Final Roster Age
51	27	26	25	Grandmother did no know	Not in a pair	Diff. ≤ 1 year	27
52	61	62	62	IR corrected age he gave at Screen	61	Diff. ≤ 1 year	61
53	18	17	18	per respondent	Not in a pair	Diff. ≤ 1 year	18
54	22	21	22	She sd last birthdy 21	Not in a pair	Diff. ≤ 1 year	22
55	39	38	39	did not mean to	Not in a pair	Diff. ≤ 1 year	39
56	32	31	31	corect age is 32	Not in a pair	Diff. ≤ 1 year	32
57	44	43	43	R says her birthday is correct but insists she is 43 not 44 and that she has a reason for saying that but could not say	44	Diff. ≤ 1 year	44
58	25	24	24	his birthday was yesterday	25	Diff. ≤ 1 year	25
59	17	16	17	R just turned 17	17	Diff. ≤ 1 year	17
60	40	39	40	39	Not in a pair	Diff. ≤ 1 year	40
61	45	46	45	The r is actually 46	Not in a pair	Diff. ≤ 1 year	45
62	41	40	41	Interviewer is male but it is not accepting that as gender	41	Diff. ≤ 1 year	41
63	46	45	46	She says she is 45 and that she will be 46	46	Diff. ≤ 1 year	46
64	35	36	35	IR is 36. not sure why this happened	35	Diff. ≤ 1 year	35
65	23	22	22	dont know what went wrong or how to fix	Not in a pair	Diff. ≤ 1 year	23
66	16	15	16	response per IR	Not in a pair	Diff. ≤ 1 year	16
67	13	14	14	Mom said daughter was 13 but she is really 14	13	Diff. ≤ 1 year	13
68	39	38	38	38	Not in a pair	Diff. ≤ 1 year	39
69	14	13	14	respondents birthday 3/19	14	Diff. ≤ 1 year	14
70	47	48	48	she forgot she turned 48	47	Diff. ≤ 1 year	47
71	22	23	23	sr told fi he was 23 but he is 22	22	Diff. ≤ 1 year	22
72	22	23	22	birthday recently	22	Diff. ≤ 1 year	22
73	28	29	29	WRONG AGE	28	Diff. ≤ 1 year	28
74	48	47	47	turning 48 next week	48	Diff. ≤ 1 year	48

Table G.1 Explanations for Overrides to Consistency Checks against CURNTAGE (continued)

#	NEW AGE	Original Roster Age for Self	Screener Age	Verbatim Explanation from Field Interviewers ¹	Respondent's Age in Roster of Other Pair Member	Comment ²	AGE = Final Roster Age
75	60	59	60	Turning 60 soon	60	Diff. ≤ 1 year	60
76	36	35	35	wrong age	36	Diff. ≤ 1 year	36
77	23	22	21	respondent is born in 1993	Not in a pair	Diff. ≤ 1 year	23
78	33	32	33	33	33	Diff. ≤ 1 year	33
79	65	66	66	RA is 65 yrs old now!	65	Diff. ≤ 1 year	65
80	52	53	53	he is 53	Not in a pair	Diff. ≤ 1 year	52
81	31	30	30	the person is 30 not 31	Not in a pair	Diff. ≤ 1 year	31
82	17	18	18	she is 18	17	Diff. ≤ 1 year	17
83	23	24	24	23	Not in a pair	Diff. ≤ 1 year	23
84	13	12	12	actual age is 13	13	Diff. ≤ 1 year	13
85	16	15	16	16 wf	Not in a pair	Diff. ≤ 1 year	16
86	34	32	32	IR preoccupied with 5 little kids earlier and gave me wrong date of birth/age. correct is 32	32	<i>AGE was changed to equal roster age</i>	32
87	44	42	42	states she is 42	42	<i>AGE was changed to equal roster age</i>	42
88	78	76	78	he just said age	Not in a pair	Diff. > 1 year	78
89	59	57	59	don't know why it's saying 57	Not in a pair	Diff. > 1 year	59

Table G.1 Explanations for Overrides to Consistency Checks against CURNTAGE (continued)

#	NEW AGE	Original Roster Age for Self	Screener Age	Verbatim Explanation from Field Interviewers ¹	Respondent's Age in Roster of Other Pair Member	Comment ²	AGE = Final Roster Age
90	31	33	33	is male	Not in a pair	Diff. > 1 year	31
91	59	57	57	he gave the wrong birthday, confused	57	<i>AGE was changed to equal roster age</i>	57
92	22	25	25	R is 22	Not in a pair	Diff. > 1 year	22
93	31	34	31	enter wrong answer earlier at interview	Not in a pair	Diff. > 1 year	31
94	39	36	39	wife is older	Not in a pair	Diff. > 1 year	39
95	23	20	19	i made a mistake earlier. correct age is 20.	23	Diff. > 1 year, other pair member supports NEWAGE	23
96	44	40	40	The SR did not know the R B's age since she is his roommate	44	Diff. > 1 year, other pair member supports NEWAGE	44
97	48	43	43	I ckd back pages and age 43 I dont see where is 48?	43	<i>AGE was changed to equal roster age</i>	43
98	12	7	12	Nino se confundio con las edades de los hermanos	12	Diff. > 1 year, other pair member supports NEWAGE	12
99	63	58	58	she is 58 as is recorded in my tablet and confirmed by the respondent.	Not in a pair	<i>AGE was changed to equal roster age</i>	58
100	26	20	20	because 20 is right age	20	<i>AGE was changed to equal roster age</i>	20
101	24	17	17	Respondent thought question was asking about his 24YO brother	Not in a pair	<i>AGE was changed to equal roster age</i>	17

Table G.1 Explanations for Overrides to Consistency Checks against CURNTAGE (continued)

#	NEW AGE	Original Roster Age for Self	Screener Age	Verbatim Explanation from Field Interviewers ¹	Respondent's Age in Roster of Other Pair Member	Comment ²	AGE = Final Roster Age
102	48	41	48	41	Not in a pair	Diff. > 1 year	48
103	41	34	34	Respondent said his age was 41 at screening	Not in a pair	Diff. > 1 year	41
104	35	43	43	error she was born in 1973, i got a new computer and maybe dates are wrong	Not in a pair	<i>AGE was changed to equal roster age</i>	43
105	41	49	41	will correct	41	Diff. > 1 year, other pair member supports NEWAGE	41
106	36	27	35	incorrect age given	Not in a pair	Diff. > 1 year	36
107	28	18	18	male respondent was rostered in both the tablet and the computer as 18 yrs old	Not in a pair	<i>AGE was changed to equal roster age</i>	18
108	83	73	70	IR explained before that there was a mistake when she adjusted her immigration status. she chose 80 something as her age at the begining of the interview and then now she is choosing 73 as her age.	83	Diff. > 1 year, other pair member supports NEWAGE	83
109	26	16	16	I am unsure why 26 year old is on there.	16	<i>AGE was changed to equal roster age</i>	16
110	18	33	18	1	18	Diff. > 1 year, other pair member supports NEWAGE	18
111	23	3	23	age is correct that was just entered	23	Diff. > 1 year, other pair member supports NEWAGE	23
112	64	997	99	Refuse to give accurate age.	64	other pair member supports NEWAGE	64

Table G.1 Explanations for Overrides to Consistency Checks against CURNTAGE (continued)

#	NEW AGE	Original Roster Age for Self	Screener Age	Verbatim Explanation from Field Interviewers¹	Respondent's Age in Roster of Other Pair Member	Comment²	AGE = Final Roster Age
113	48	997	48	ONE PERSON LESS. DAUGHTER MOVING IN MONTH	Not in a pair	FI does not support change leave as NEWAGE	48
114	22	997	22	The program isn't letting me progress to the next question unless I suppress it	Not in a pair	FI does not support change leave as NEWAGE	22

¹ These entries came directly from the 2016 NSDUH field interviewers. Any typographical errors or misspellings were transcribed directly and were not corrected.

² "Diff." refers to the difference between CURNTAGE and the age of the "self" in the household roster, the "roster age." Bolded and italicized entries indicate that the criteria for changing the age to that given in the household roster for "self" were met.

Table G.2 Explanations for Overrides to Internal Consistency Checks

#	Consistency Check	AGE	Roster Member's Age and Relationship to Respondent	Verbatim Explanation from Field Interviewers ¹	Comment
1	Grandparent/Grandchild and respondent less than 30 years apart	42	22-year-old grandson	dad had kids early	Override; probable child
2	Grandparent/Grandchild and respondent less than 30 years apart	42	19-year-old grandson	dad had kids early	Override; probable child
3	Grandparent/Grandchild and respondent less than 30 years apart	26	44-year-old grandmother	see comment - it's her step grandmother	Override; probable parent
4	Multiple spouses or partners	42	1.5-year-old wife	the daughter is 18 months old	Override; probable child
5	Multiple spouses or partners	22	52-year-old husband	22	Override; probable parent-in-law
6	Respondent is 16 or younger and has a spouse or partner	14	48-year-old husband	48	Override; probable parent
7	Respondent is 16 or younger and has a spouse or partner	13	31-year-old husband	mixed up	Override; probable parent
8	Respondent is 16 or younger and has an in-law	15	51-year-old mother-in-law	thats what he said	Override; unsure of relationship (85)
9	Respondent is 16 or younger and has an in-law	16	34-year-old father-in-law	he insists hes his father in law	Override; probable parent
10	Respondent is 16 or younger and has an in-law	13	33-year-old mother-in-law	r thinks this is the right answer. I am not allowed to correct him.	Override; probable other nonrelative
11	Respondent's daughter-in-law is older than respondent	44	48-year-old daughter-in-law	r says its complicated	Override; unsure of relationship (85)
12	Respondent's female in-law is 16 or younger	34	1-year-old daughter-in-law	daughter	Override; probable child
13	Respondent's female in-law is 16 or younger	21	3-year-old daughter-in-law	that is the information he provide	Override; unsure of relationship (85)
14	Respondent's granddaughter is older than respondent	15	16-year-old granddaughter	Information given by grandmother.	Override; probable sibling
15	Respondent's son-in-law is 16 or younger	24	16-year-old son-in-law	per respondent	Override; probable child

¹ These entries came directly from the 2016 NSDUH field interviewers. Any typographical errors or misspellings were transcribed directly and were not corrected.

Appendix H: Rules for Determining Pair Relationships

This page intentionally left blank

Appendix H: Rules for Determining Pair Relationships

H.1 Rules for Determining Matching Pairs, in Priority Order

The following rules are used to determine the member of a respondent's household roster that corresponds to the other pair member. In these rules, an "age match" occurs if the questionnaire age of one pair member matches a roster age in the other pair member's roster, and a "gender match" occurs if the questionnaire gender of one of the pair members matches a roster gender in the other pair member's roster. In the table below, if the rules for Pair Member A and Pair Member B in a single row differ, then the count for that row includes the rules as listed, and the rules with Pair Member A and Pair Member B reversed. If the age and/or gender are off when finding these matches, the age and/or gender are defined by the questionnaire age and gender of the selected pair member when determining the pair domain. The rules, called match types, are listed in priority order in [Table H.1](#), along with the number of pairs to which each rule was applied. Since the 2001 survey, it was technically impossible to identify more than one roster member as the "other pair member selected," resulting in either 0 or 1 MBRSEL for each responding pair. Rules involving situations where more than one MBRSEL existed are therefore not included in this table.

Table H.1 Rules for Determining Matching Pairs, in Priority Order

Match Type	Rule		
	Pair Member A	Pair Member B	Count
0.00	Age and gender match exactly, exactly one MBRSEL in right place	Age and gender match exactly, exactly one MBRSEL in right place	13,897
2.01	Age and gender match exactly, exactly one MBRSEL in right place	Age within one, gender matches exactly, exactly one MBRSEL in right place	2,324
2.03	Age within one, gender matches exactly, exactly one MBRSEL in right place	Age within one, gender matches exactly, exactly one MBRSEL in right place	167
3.01	Age and gender match exactly, exactly one MBRSEL in right place	Age within two, gender matches exactly, exactly one MBRSEL in right place	393
3.03	Age within one, gender matches exactly, exactly one MBRSEL in right place	Age within two, gender matches exactly, exactly one MBRSEL in right place	43
3.04	Age within two, gender matches exactly, exactly one MBRSEL in right place	Age within two, gender matches exactly, exactly one MBRSEL in right place	1

Table H.1 Rules for Determining Matching Pairs, in Priority Order (continued)

Match Type	Rule		
	Pair Member A	Pair Member B	Count
4.01	Age and gender match exactly, exactly one MBRSEL in right place	Age and gender match exactly, MBRSEL missing for all roster members	372
4.03	Age within one, gender matches exactly, exactly one MBRSEL in right place	Age and gender match exactly, MBRSEL missing for all roster members	46
4.04	Age within two, gender matches exactly, exactly one MBRSEL in right place	Age and gender match exactly, MBRSEL missing for all roster members	4
4.05	Age and gender match exactly, MBRSEL missing for all roster members	Age and gender match exactly, MBRSEL missing for all roster members	25
4.06	Age and gender match exactly, exactly one MBRSEL in right place	Two matches for age and gender, MBRSEL missing for all roster members, relationship is "Other Relative"	0
4.08	Age within one, gender matches exactly, exactly one MBRSEL in right place	Two matches for age and gender, MBRSEL missing for all roster members, relationship is "Other Relative"	0
4.09	Age within two, gender matches exactly, exactly one MBRSEL in right place	Two matches for age and gender, MBRSEL missing for all roster members, relationship is "Other Relative"	0
4.10	Age and gender match exactly, MBRSEL missing for all roster members	Two matches for age and gender, MBRSEL missing for all roster members, relationship is "Other Relative"	0
4.11	Age and gender match exactly, exactly one MBRSEL in right place	Two matches for age and gender, MBRSEL missing for all roster members, roommate type relationship	0
4.13	Age within one, gender matches exactly, exactly one MBRSEL in right place	Two matches for age and gender, MBRSEL missing for all roster members, roommate type relationship	0
4.14	Age within two, gender matches exactly, exactly one MBRSEL in right place	Two matches for age and gender, MBRSEL missing for all roster members, roommate type relationship	0
4.15	Age and gender match exactly, MBRSEL missing for all roster members	Two matches for age and gender, MBRSEL missing for all roster members, roommate type relationship	0

Table H.1 Rules for Determining Matching Pairs, in Priority Order (continued)

Match Type	Rule		
	Pair Member A	Pair Member B	Count
5.01	Age and gender match exactly, exactly one MBRSEL in right place	Age matches exactly, gender off, exactly one MBRSEL in right place	20
5.03	Age within one, gender matches exactly, exactly one MBRSEL in right place	Age matches exactly, gender off, exactly one MBRSEL in right place	0
5.04	Age within two, gender matches exactly, exactly one MBRSEL in right place	Age matches exactly, gender off, exactly one MBRSEL in right place	1
5.05	Age and gender match exactly, MBRSEL missing for all roster members	Age matches exactly, gender off, exactly one MBRSEL in right place	2
5.06	Age matches exactly, gender off, exactly one MBRSEL in right place	Age matches exactly, gender off, exactly one MBRSEL in right place	0
8.01	Age and gender match exactly, exactly one MBRSEL in right place	Age within one, gender matches exactly, MBRSEL missing for all roster members	41
8.03	Age within one, gender matches exactly, exactly one MBRSEL in right place	Age within one, gender matches exactly, MBRSEL missing for all roster members	8
8.04	Age within two, gender matches exactly, exactly one MBRSEL in right place	Age within one, gender matches exactly, MBRSEL missing for all roster members	0
8.05	Age and gender match exactly, MBRSEL missing for all roster members	Age within one, gender matches exactly, MBRSEL missing for all roster members	2
8.06	Age matches exactly, gender off, exactly one MBRSEL in right place	Age within one, gender matches exactly, MBRSEL missing for all roster members	0
8.09	Age within one, gender matches exactly, MBRSEL missing for all roster members	Age within one, gender matches exactly, MBRSEL missing for all roster members	1
10.01	Age and gender match exactly, exactly one MBRSEL in right place	Age within 10, gender matches, exactly one MBRSEL in right place, excludes cases where MBRSEL could have been applied to one of closer age	263

Table H.1 Rules for Determining Matching Pairs, in Priority Order (continued)

Match Type	Rule		
	Pair Member A	Pair Member B	Count
10.03	Age within one, gender matches exactly, exactly one MBRSEL in right place	Age within 10, gender matches, exactly one MBRSEL in right place, excludes cases where MBRSEL could have been applied to one of closer age	33
10.04	Age within two, gender matches exactly, exactly one MBRSEL in right place	Age within 10, gender matches, exactly one MBRSEL in right place, excludes cases where MBRSEL could have been applied to one of closer age	7
10.05	Age and gender match exactly, MBRSEL missing for all roster members	Age within 10, gender matches, exactly one MBRSEL in right place, excludes cases where MBRSEL could have been applied to one of closer age	5
10.06	Age matches exactly, gender off, exactly one MBRSEL in right place	Age within 10, gender matches, exactly one MBRSEL in right place, excludes cases where MBRSEL could have been applied to one of closer age	0
10.09	Age within one, gender matches exactly, MBRSEL missing for all roster members	Age within 10, gender matches, exactly one MBRSEL in right place, excludes cases where MBRSEL could have been applied to one of closer age	0
10.11	Age within 10, gender matches, exactly one MBRSEL in right place, excludes cases where MBRSEL could have been applied to one of closer age	Age within 10, gender matches, exactly one MBRSEL in right place, excludes cases where MBRSEL could have been applied to one of closer age	1
11.01	Age and gender match exactly, exactly one MBRSEL in right place	Age within 10, gender matches, MBRSEL missing for all roster members, excludes cases where one of closer age could have been selected	5
11.03	Age within one, gender matches exactly, exactly one MBRSEL in right place	Age within 10, gender matches, MBRSEL missing for all roster members, excludes cases where one of closer age could have been selected	1

Table H.1 Rules for Determining Matching Pairs, in Priority Order (continued)

Match Type	Rule		
	Pair Member A	Pair Member B	Count
11.04	Age within two, gender matches exactly, exactly one MBRSEL in right place	Age within 10, gender matches, MBRSEL missing for all roster members, excludes cases where one of closer age could have been selected	1
11.05	Age and gender match exactly, MBRSEL missing for all roster members	Age within 10, gender matches, MBRSEL missing for all roster members, excludes cases where one of closer age could have been selected	0
11.06	Age matches exactly, gender off, exactly one MBRSEL in right place	Age within 10, gender matches, MBRSEL missing for all roster members, excludes cases where one of closer age could have been selected	0
11.09	Age within one, gender matches exactly, MBRSEL missing for all roster members	Age within 10, gender matches, MBRSEL missing for all roster members, excludes cases where one of closer age could have been selected	0
12.01	Age and gender match exactly, exactly one MBRSEL in right place	Everything missing	10
12.03	Age within one, gender matches exactly, exactly one MBRSEL in right place	Everything missing	2
12.04	Age within two, gender matches exactly, exactly one MBRSEL in right place	Everything missing	0
12.05	Age and gender match exactly, MBRSEL missing for all roster members	Everything missing	0
12.06	Age matches exactly, gender off, exactly one MBRSEL in right place	Everything missing	0
12.09	Age within one, gender matches exactly, MBRSEL missing for all roster members	Everything missing	0

Table H.1 Rules for Determining Matching Pairs, in Priority Order (continued)

Match Type	Rule		
	Pair Member A	Pair Member B	Count
13.01	Age and gender match exactly, exactly one MBRSEL in right place	Gender and reported household sizes match exactly, age missing, MBRSEL missing for all roster members	62
13.03	Age within one, gender matches exactly, exactly one MBRSEL in right place	Gender and reported household sizes match exactly, age missing, MBRSEL missing for all roster members	3
13.04	Age within two, gender matches exactly, exactly one MBRSEL in right place	Gender and reported household sizes match exactly, age missing, MBRSEL missing for all roster members	1
13.05	Age and gender match exactly, MBRSEL missing for all roster members	Gender and reported household sizes match exactly, age missing, MBRSEL missing for all roster members	0
13.06	Age matches exactly, gender off, exactly one MBRSEL in right place	Gender and reported household sizes match exactly, age missing, MBRSEL missing for all roster members	0
13.09	Age within one, gender matches exactly, MBRSEL missing for all roster members	Gender and reported household sizes match exactly, age missing, MBRSEL missing for all roster members	0
15.01	Age and gender match exactly, exactly one MBRSEL in right place	Multiple matches on age, gender, and relationship code; MBRSEL missing for all roster members; does not matter which match is picked	2
15.03	Age within one, gender matches exactly, exactly one MBRSEL in right place		0
15.04	Age within two, gender matches exactly, exactly one MBRSEL in right place	Multiple matches on age, gender, and relationship code; MBRSEL missing for all roster members; does not matter which match is picked	0
15.05	Age and gender match exactly, MBRSEL missing for all roster members	Multiple matches on age, gender, and relationship code; MBRSEL missing for all roster members; does not matter which match is picked	0

Table H.1 Rules for Determining Matching Pairs, in Priority Order (continued)

Match Type	Rule		
	Pair Member A	Pair Member B	Count
15.06	Age matches exactly, gender off, exactly one MBRSEL in right place	Multiple matches on age, gender, and relationship code; MBRSEL missing for all roster members; does not matter which match is picked	0
15.09	Age within one, gender matches exactly, MBRSEL missing for all roster members	Multiple matches on age, gender, and relationship code; MBRSEL missing for all roster members; does not matter which match is picked	0
16.01	Age and gender match exactly, exactly one MBRSEL in right place	Multiple matches on gender and age, MBRSEL missing for all roster members, pick one where relationship code matches	0
18.01	Age and gender match exactly, exactly one MBRSEL in right place	Multiple matches on gender, age within one, MBRSEL missing for all roster members, pick one where relationship code matches	0
18.03	Age within one, gender matches exactly, exactly one MBRSEL in right place	Multiple matches on gender, age within one, MBRSEL missing for all roster members, pick one where relationship code matches	0
18.04	Age within two, gender matches exactly, exactly one MBRSEL in right place	Multiple matches on gender, age within one, MBRSEL missing for all roster members, pick one where relationship code matches	0
18.05	Age and gender match exactly, MBRSEL missing for all roster members	Multiple matches on gender, age within one, MBRSEL missing for all roster members, pick one where relationship code matches	0
18.06	Age matches exactly, gender off, exactly one MBRSEL in right place	Multiple matches on gender, age within one, MBRSEL missing for all roster members, pick one where relationship code matches	0
18.09	Age within one, gender matches exactly, MBRSEL missing for all roster members	Multiple matches on gender, age within one, MBRSEL missing for all roster members, pick one where relationship code matches	0
19.01	Age and gender match exactly, exactly one MBRSEL in right place	Age within one, gender off, one MBRSEL, only two in household	0
19.03	Age within one, gender matches exactly, exactly one MBRSEL in right place	Age within one, gender off, one MBRSEL, only two in household	0

Table H.1 Rules for Determining Matching Pairs, in Priority Order (continued)

Match Type	Rule		
	Pair Member A	Pair Member B	Count
19.04	Age within two, gender matches exactly, exactly one MBRSEL in right place	Age within one, gender off, one MBRSEL, only two in household	0
19.05	Age and gender match exactly, MBRSEL missing for all roster members	Age within one, gender off, one MBRSEL, only two in household	0
19.06	Age matches exactly, gender off, exactly one MBRSEL in right place	Age within one, gender off, one MBRSEL, only two in household	0
19.09	Age within one, gender matches exactly, MBRSEL missing for all roster members	Age within one, gender off, one MBRSEL, only two in household	0
19.11	Age within one, gender off, one MBRSEL, only two in household	Age within one, gender off, one MBRSEL, only two in household	0
20.00	No match, but no relationship codes are missing, and none involve domains of interest	No match, but no relationship codes are missing, and none involve domains of interest	17
21.01	Age and gender match exactly, exactly one MBRSEL in right place	Age matches exactly, gender off, MBRSEL missing for all roster members	0
21.03	Age within one, gender matches exactly, exactly one MBRSEL in right place	Age matches exactly, gender off, MBRSEL missing for all roster members	0
21.04	Age within two, gender matches exactly, exactly one MBRSEL in right place	Age matches exactly, gender off, MBRSEL missing for all roster members	0
21.05	Age and gender match exactly, MBRSEL missing for all roster members	Age matches exactly, gender off, MBRSEL missing for all roster members	0
21.06	Age matches exactly, gender off, exactly one MBRSEL in right place	Age matches exactly, gender off, MBRSEL missing for all roster members	0
21.09	Age within one, gender matches exactly, MBRSEL missing for all roster members	Age matches exactly, gender off, MBRSEL missing for all roster members	0
22.01	Age and gender match exactly, exactly one MBRSEL in right place	No match at all (often paired respondent is missing from roster)	49

Table H.1 Rules for Determining Matching Pairs, in Priority Order (continued)

Match Type	Rule		
	Pair Member A	Pair Member B	Count
22.03	Age within one, gender matches exactly, exactly one MBRSEL in right place	No match at all (often paired respondent is missing from roster)	12
22.04	Age within two, gender matches exactly, exactly one MBRSEL in right place	No match at all (often paired respondent is missing from roster)	4
22.05	Age and gender match exactly, MBRSEL missing for all roster members	No match at all (often paired respondent is missing from roster)	4
22.06	Age matches exactly, gender off, exactly one MBRSEL in right place	No match at all (often paired respondent is missing from roster)	0
22.09	Age within one, gender matches exactly, MBRSEL missing for all roster members	No match at all (often paired respondent is missing from roster)	0
25.00	No match at all	No match at all	18

H.2 Rules for Identifying Pair Relationships among Pairs

[Table H.2](#) summarizes the rules used to identify the pair relationships, using the relationship codes and questionnaire ages of the two pair members. Because the child (12 to 17)-parent and child (12 to 20)-parent relationships can be derived from relationships created using 12- to 14-year-olds, 15- to 17-year-olds, and 18- to 20-year-olds, these latter relationships are the ones referenced in the rules. The variable PAIRREL, which is the next to last column of the table, identifies the pair relationship as defined by [Table 14.3](#) in the main body of this report. As with the rules for identifying which members of the roster belong to the pair, these rules—also called priority conditions because of their hierarchical nature—are shown in priority order. In the headers, the moniker "A" refers to pair member A, and "B" refers to pair member B. The relationship between A and B is described in the columns "A-B Relationship," from the perspective of pair member A ("B to A, According to A") and the perspective of pair member B ("A to B, According to B"). Any constraints on the pair members (other than FIPE3) are provided in the columns "Constraint on A" and "Constraint on B." These constraints include age constraints, where a range of ages (e.g., 12 to 17) indicates that the value of the questionnaire edited age (AGE) is between the numbers shown. Also in this column, "child" and "children" are defined as (a) roster member(s) with nonmissing ages less than 18. The question FIPE3 asks if the respondent is the parent of a selected 12- to 17-year-old. The responses provided in the table are either "yes" or "no." The column for RELMATCH indicates the quality of the match between pair members, as defined in [Table 14.4](#) in the main body of this report. In the table, blank cells mean that no restrictions were placed on that variable to determine the pair relationship.

Table H.2 Rules for Identifying Pair Relationships among Pairs

Priority Condition	A-B Relationship		Constraint on A	Constraint on B	FIPE3 (A)	FIPE3 (B)	PAIRREL	REL-MATCH
	B to A, According to A	A to B, According to B						
1	Parent	Child	12-14				1	1
	Child	Parent		12-14				
2	Parent	Child	15-17				2	1
	Child	Parent		15-17				
3	Parent	Child	18-20				3	1
	Child	Parent		18-20				
4	Parent	Child	21+				4	1
	Child	Parent		21+				
5	Sibling	Sibling	12-14	15-17			5	1
	Sibling	Sibling	15-17	12-14				
6	Sibling	Sibling	12-17	18-25			6	1
	Sibling	Sibling	18-25	12-17				
7	Sibling	Sibling	No constraints, after considering #5 & #6					1
8	Spouse/partner	Spouse/partner	≥ 1 child	≥ 1 child			8	1
9	Spouse/partner	Spouse/partner	0 children, no bad data	0 children, no bad data			9	1
10	Spouse/partner	Spouse/partner	≥ 1 child	0 children, some bad data			8	1.5
	Spouse/partner	Spouse/partner	0 children, some bad data	≥ 1 child				
11	Spouse/partner	Roommate/nonrelative	≥ 1 child both sides, equal number each side					3
	Roommate/nonrelative	Spouse/partner	≥ 1 child both sides, equal number each side					

Table H.2 Rules for Identifying Pair Relationships among Pairs (continued)

[illegible]

Table H.2 Rules for Identifying Pair Relationships among Pairs (continued)

Priority Condition	A-B Relationship		Constraint on A	Constraint on B	FIPE3 (A)	FIPE3 (B)	PAIRREL	REL-MATCH
	B to A, According to A	A to B, According to B						
23	Child	Missing	A less than 10 yrs. older th. B	15-17	No		14	0
	Missing	Parent						
24	Child	Missing		15-17			2	2
	Missing	Parent						
25	Parent	Missing	18-20				3	2
	Missing	Parent		18-20				
26	Child	Missing		18-20			3	2
	Missing	Child	18-20					
27	Parent	Missing	21+				4	2
	Missing	Parent		21+				
28	Child	Missing		21+			4	2
	Missing	Child	21+					
29	Sibling	Missing	12-14	15-17			5	2
			15-17	12-14				
	Missing	Sibling	12-14	15-17				
			15-17	12-14				
30	Sibling	Missing	12-17	18-25			6	2
			18-25	12-17				
	Missing	Sibling	12-17	18-25				
			18-25	12-17				
31	Sibling	Missing	No constraints, after considering #24, #25					2
	Missing	Sibling	No constraints, after considering #24, #25					
32	Spouse/partner	Missing	≥ 1 child	No spouse in roster			8	2
	Missing	Spouse/partner	No spouse in roster	≥ 1 child				

Table H.2 Rules for Identifying Pair Relationships among Pairs (continued)

Priority Condition	A-B Relationship		Constraint on A	Constraint on B	FIPE3 (A)	FIPE3 (B)	PAIRREL	REL-MATCH
	B to A, According to A	A to B, According to B						
33	Spouse/partner	Missing	0 children, no bad data	No spouse in roster			9	2
	Missing	Spouse/partner	No spouse in roster	0 children, no bad data				
34	Spouse/partner	Missing	After #27, #28, no constraints	No spouse in roster			10	2
	Missing	Spouse/partner	No spouse in roster	After #27, #28, no constraints				
35	Grandchild	Missing	A at least 20 years older than B					2
	Missing	Grandparent						
	Grandparent	Missing	B at least 20 years older than A					
	Missing	Grandchild						
36	Roommate/boarder/other relative/nonrelative/in-laws	Missing			No		12	2
	Missing	Roommate/boarder/other relative/nonrelative/in-laws			No			
37	Roommate/boarder/other relative/nonrelative/in-laws	Missing					13	2
	Missing	Roommate/boarder/other relative/nonrelative/in-laws						
38	Nonmissing	Child	12-14			Yes	1	3
39	Nonmissing	Parent		12-14	Yes		1	3
40	Child	Nonmissing		12-14	Yes		1	3
41	Parent	Nonmissing	12-14			Yes	1	3
42	Nonmissing	Child	15-17			Yes	2	3

Table H.2 Rules for Identifying Pair Relationships among Pairs (continued)

Priority Condition	A-B Relationship		Constraint on A	Constraint on B	FIPE3 (A)	FIPE3 (B)	PAIRREL	REL-MATCH
	B to A, According to A	A to B, According to B						
43	Nonmissing	Parent		15-17	Yes		2	3
44	Child	Nonmissing		15-17	Yes		2	3
45	Parent	Nonmissing	15-17			Yes	2	3
46	Parent	Roommate/boarder/ other relative/ nonrelative	12-14			No	13	3
						Missing	15	4
	Roommate/boarder/ other relative/ nonrelative	Parent		12-14	No		13	3
					Missing		15	4
47	Parent	Roommate/boarder/ other relative/ nonrelative	15-17			No	13	3
						Missing	16	4
	Roommate/boarder/ other relative/ nonrelative	Parent		15-17	No		13	3
					Missing		16	4
48	Parent	Roommate/boarder/ other relative/ nonrelative	18-20				17	4
	Roommate/boarder/ other relative/ nonrelative	Parent		18-20			17	4
49	Parent	Roommate/boarder/ other relative/ nonrelative	21+				18	4
	Roommate/boarder/ other relative/ nonrelative	Parent		21+			18	4

Table H.2 Rules for Identifying Pair Relationships among Pairs (continued)

Priority Condition	A-B Relationship		Constraint on A	Constraint on B	FIPE3 (A)	FIPE3 (B)	PAIRREL	REL-MATCH
	B to A, According to A	A to B, According to B						
50	Nonmissing, not sibling	Child	12-14	21-75		No	13	3
			12-14, exactly one parent	21-75, exactly one spouse		Missing	1	3
			12-14, 0 or 2 parents, or B has 0 or 2 spouse	21-75, 0 or 2 spouses, or A has 0 or 2 parents		Missing	15	4
	Child	Nonmissing, not sibling	21-75	12-14	No		13	3
			21-75, exactly one spouse	12-14, exactly one parent	Missing		1	3
			21-75, 0 or 2 spouses, or A has 0 or 2 parents	12-14, 0 or 2 parents, or B has 0 or 2 spouse	Missing		15	4
51	Nonmissing, not sibling	Child	15-17	24-75		No	13	3
			15-17, exactly one parent	24-75, exactly one spouse		Missing	2	3
			15-17, 0 or 2 parents, or B has 0 or 2 spouse	24-75, 0 or 2 spouses, or A has 0 or 2 parents		Missing	16	4
	Child	Nonmissing, not sibling	24-75	15-17	No		13	3
			24-75, exactly one spouse	15-17, exactly one parent	Missing		2	3
			24-75, 0 or 2 spouses, or A has 0 or 2 parents	15-17, 0 or 2 parents, or B has 0 or 2 spouse	Missing		16	4

Table H.2 Rules for Identifying Pair Relationships among Pairs (continued)

[illegible]

Table H.2 Rules for Identifying Pair Relationships among Pairs (continued)

Priority Condition	A-B Relationship		Constraint on A	Constraint on B	FIPE3 (A)	FIPE3 (B)	PAIRREL	REL-MATCH
	B to A, According to A	A to B, According to B						
56	Spouse	Sibling	Ages neither 12-14/15-17 nor 12-17/18-25, both sides have parents or spouses				7	3
	Sibling	Spouse						
57	Other relative	Sibling	Both sides have 2 parents, ages of oldest parents on either side differ by > 5 years, age of youngest parents on either side differ by > 5 years				13	3
	Sibling	Other relative						
58	Nonmissing, not child	Sibling	15-17				19	4
	Sibling	Nonmissing, not child	12-14					
59	Nonmissing, not parent	Sibling	12-14				19	4
	Sibling	Nonmissing, not parent	15-17					
60	Nonmissing, not child	Sibling	18-25				20	4
	Sibling	Nonmissing, not child	12-17					
61	Nonmissing, not parent	Sibling	12-17				20	4
	Sibling	Nonmissing, not parent	18-25					
62	Nonmissing, not child	Sibling	Ages neither 12-14/15-17 nor 12-17/18-25, A older than B				21 21	4
	Sibling	Nonmissing, not child	Ages neither 12-14/15-17 nor 12-17/18-25, B older than A					4
63	Nonmissing, not parent	Sibling	Ages neither 12-14/15-17 nor 12-17/18-25, B older than A				21 21	4
	Sibling	Nonmissing, not parent	Ages neither 12-14/15-17 nor 12-17/18-25, A older than B					4

Table H.2 Rules for Identifying Pair Relationships among Pairs (continued)

Priority Condition	A-B Relationship		Constraint on A	Constraint on B	FIPE3 (A)	FIPE3 (B)	PAIRREL	REL-MATCH
	B to A, According to A	A to B, According to B						
64	Sibling	Roommate, in-law, grandparent, grandchild, boarder, other relative, nonrelative	At least one is between 18 and 20				13	3
	Roommate, in-law, grandparent, grandchild, boarder, other relative, nonrelative	Sibling	At least one is between 18 and 20					
65	Sibling	Unusual in-law code	12-20	26 or older			13	3
	Unusual in-law code	Sibling	26 or older	12-20				
66	Spouse/partner	Not a child, parent, or sibling	≥ 1 child aged < 18	No spouse			22	4
	Not a child, parent, or sibling	Spouse/partner	no spouse	≥ 1 child aged < 18				
67	Spouse/partner	Not a child, parent, or sibling	15 or older, 0 children, no bad data	15 or older, no spouse			23	4
	Not a child, parent, or sibling	Spouse/partner	15 or older, no spouse	15 or older, 0 children, no bad data				
68	Grandparent, grandchild	Not grandparent, not grandchild					25	4
	Not grandparent, not grandchild	Grandparent, grandchild						
69	Any codes	Any codes	No constraints	No constraints			14	0

Appendix I: Conditions Used to Reconcile Differing Multiplicity Counts between Pair Members

This page intentionally left blank

Appendix I: Conditions Used to Reconcile Differing Multiplicity Counts between Pair Members

I.1 Introduction

To determine multiplicity counts, counts were obtained from each pair member. The count from the pair member who was the focus member of the domain was considered the direct count, and the count from the other pair member was considered the indirect count. Typically, these counts were in agreement, and the determination of the final multiplicity count was straightforward, provided neither roster had bad data codes. The strategy also was usually clear if one pair member had bad data in the household roster, or had a 0 count when the pair relationship precluded a value of 0. The count from the pair member with good, nonzero data was usually preferred in those cases. If the bad data was limited to bad relationship codes, then the member with good data was selected only if substituting the appropriate relationship codes for the bad data codes would have given a total that was equal to the count from the pair member with good data. There were instances where bad data codes existed in the roster, and this condition did not apply. Finally, there were instances where neither pair member had bad data in their rosters, yet their counts still disagreed. The rules that were used to reconcile these disagreeing counts are outlined in this appendix.

Note that the reconciliation of differing counts was necessary for parent-child and sibling-sibling pairs but was not necessary for spouse-spouse pairs, since the multiplicity count for spouse-spouse pairs was always 1. As noted in Section 14.3.1, it was technically possible for a respondent to have multiple spouses, but these situations were not accounted for.

I.2 Parent-Child Counts

For parent-child counts, the screener and the FIPE3 variable were used to help reconcile disagreeing counts. The rules follow below, separated by the member of focus:

Parent-child pairs, child focus. The multiplicity counts in this domain reflected the selected child's parents and in most cases had values of 1 or 2. If neither side had bad relationship codes, and if the direct count exceeded the indirect count, the following rules applied:

1. The direct count might have exceeded the indirect count because one parent had left or entered the household between interviews. In this case, the ages in the rosters were matched to the screener roster to determine which count to believe. This was done in two ways. First, the total number of roster members between ages 30 and 39, 35 and 44, and 40 and 49 were compared between pair members and the screener. The pair member with age range counts closest to the screener was the one whose parent-child count was chosen for the final count. If neither side had age range counts equal to the screener, then the pair member with a parent-child count equal to the total number of screener roster members between ages 26 and 64 was chosen as the final count.

2. The direct count might have exceeded the indirect count because the selected parent did not consider the other "parent" a spouse or live-in partner. If the pair relationship was not imputed, the indirect count was selected. However, if the pair relationship was imputed and the older pair member called the younger pair member a child, then the older pair member considered the child's "true" parent as not a spouse or live-in partner, even though he or she claimed the "true" parent's children. In this case, the direct count was used (the child's adjusted count).

If the direct count was exceeded by the indirect count, then the child listed only one parent, and the parent listed a spouse (a "stepparent") or live-in partner in the household roster. The following rules applied:

1. The indirect count might have exceeded the direct count because the selected child did not accept a stepparent or live-in partner as his or her parent. If this stepparent or live-in partner was the other respondent selected, this was considered a parent-child pair based on the response of the "parent" to the FIPE3 question. If the FIPE3 question was answered "yes," and the RELMATCH variable had a value of 3, then the indirect count was selected as the multiplicity count. If the FIPE3 question was answered "no," the pair was not considered a parent-child pair and was not considered for these counts. Finally, if the FIPE3 question was not answered, the respondent was considered a "parent" if he or she was a stepparent. If the respondent was a live-in partner, the determination of the pair relationship was left to imputation. The multiplicity count was set to the indirect count to account for the possibility that the pair relationship would be imputed as parent-child.
2. Suppose the selected child did not accept a stepparent or live-in partner as his or her parent (as above), but the other respondent selected was the "true" or "original" parent. In this case, the stepparent or live-in partner was identified only in the "original" parent's roster, so there was no way to determine how the stepparent or live-in partner would have answered the FIPE3 question. A stepparent was considered a "parent" even if the child did not view him or her this way so that the indirect count was used. The case of live-in partners was less clear, and these cases were left to imputation.
3. If age range counts between the two pair members and the screener matched across a variety of age ranges (30 to 39, 40 to 49, and 50 to 59), but the child's roster had a bad relationship code among roster members of potential parent age (15 or older), or the child's roster had a value of MBRSEL that did not match what was finally determined to be the child's parent, then the multiplicity count for the parent—the indirect count—was selected as the final count.

Parent-child pairs, parent focus. The multiplicity counts in this domain reflected the selected parent's children and were limited to have values of at least 1. If neither side had bad relationship codes, the following rules applied:

1. If the count of children in the household within the relevant age ranges differed between the pair members, but one side had a count of children equal to the same

count from the screener roster, then the multiplicity count that corresponded to the pair member with the same count of children as the screener was used.

2. If the count of children in the household within the relevant age ranges differed between the pair members, and both sides had a multiplicity count that exceeded the count of all children from the screener roster, then the number of children in the screener roster was used as the multiplicity count. If the screener roster had missing exact ages, then the minimum multiplicity count from the two pair members' rosters was used as the final count.
3. The direct count and indirect count might differ because either the child listed a sibling that the parent considered "another relative" or the parent listed a child that the child considered "another relative." In either case, the parent was the one to answer the FIPE3 question. Because of this, the multiplicity count from the parent's perspective was selected as the final count, provided that the counts of children in the household within the relevant age ranges for each pair member were equal.
4. After considering the above situations, the multiplicity counts might have still differed without a resolution of which count should have been chosen as the final count. This occurred because the counts of children in the household differed between pair members, each of which differed from the screener count. Moreover, multiplicity counts did not exceed the screener age range count. In this instance, if one of the multiplicity counts equalled the screener age range count, then this multiplicity was selected as the final count. However, if this was not the case, then upper and lower bounds were created and the final multiplicity was left to imputation.

Because of the hierarchical nature of these counts, parent-child counts for 12- to 17-year-old and 12- to 20-year-old children could sometimes be derived if the 12- to 14-year-old parent-child count was already determined for both child focus and parent focus counts. In particular, if one pair member's count for 12- to 17-year-old children or 12- to 20-year-old children equalled or exceeded the final parent-child count for 12- to 14-year-old children and the other did not, then the pair member's count that equalled or exceeded the 12- to 14-year-old count was chosen as the final count.

I.3 Sibling-Sibling Counts

Although there were two types of sibling-sibling pairs under consideration, each associated with two domains, the same rules could be applied to all four domains. When the older sibling was the focus, the multiplicity count was a count of the number of siblings within the younger age group (12 to 14 or 12 to 17). Conversely, the multiplicity count was the number of siblings in the older age group (15 to 17 or 18 to 25) when the younger sibling was the focus. Deciding how to assign a final multiplicity count often involved looking at a count of household members within the age range of the siblings being counted. For example, if the older sibling was the focus and the age ranges were 12 to 14 and 15 to 17, the number of household members aged 12 to 14 were counted. The following general rules applied if the multiplicity counts for each pair member disagreed:

1. The counts disagreed if a household member left or entered the household between interviews. As before, the roster that was closest to the screener was used to

- determine the count. In particular, depending upon the domain, the count of household members within the age range of the siblings being counted was compared between each pair member and the screener. The multiplicity count from the pair member with the count closest to the screener was used, provided that the member had no bad relationship codes within the relevant age range.
2. If the counts of household members within the age range of the siblings being counted differed between pair members and those counts were both exceeded by the screener count, then the multiplicity associated with the pair member with the age range count closest to the screener was chosen, provided that the member had no bad relationship codes within the relevant age range.
 3. In some cases, the counts of household members within the age range of the siblings being counted were the same for the two pair members, but the multiplicity counts disagreed.
 - a. If one pair member had bad relationship codes and the other did not, the disagreement could have been due to the bad relationship codes. If the sums of the multiplicity count and the number of bad relationship codes were equal across pair members, then the final count was set to equal the multiplicity of the pair member who did not have bad relationship codes.
 - b. If one pair member identified the other as "sibling" but the other pair member did not reciprocate, then imputation was required to establish whether the relationship was sibling-sibling. The count associated with the pair member who indicated that the other pair member was a sibling should have been chosen as the final count. In effect, this was done by taking the maximum of the two pair members' counts.
 4. If the counts of household members within the age range of the siblings being counted disagreed and both exceeded the screener count of household members within the relevant age range, then the multiplicity count was set to the screener count. If the screener roster had missing exact ages, then the minimum multiplicity count from the two pair members' rosters was used as the final count.
 5. If the differing multiplicity counts could not be reconciled with the above rules, upper and lower bounds for the true multiplicity were determined using the two multiplicity counts, as well as the counts of children within relevant age ranges in both pair member's rosters and the screener roster. In rare cases, the values for these bounds were equal. These cases were investigated, and if the reasons were legitimate, then the final multiplicity count was set to this value. Otherwise, the final multiplicity was left to imputation.

Appendix J: Conditions Used to Reconcile Differing Household-Level Person Counts between Pair Members

This page intentionally left blank

Appendix J: Conditions Used to Reconcile Differing Household-Level Person Counts between Pair Members

J.1 Introduction

Household-level person counts for a particular domain were obtainable using the multiplicity counts if the pair belonged to a pair relationship that fit into that domain, provided only one family unit was in the household. No reconciliation between pair members was necessary in that case, since the reconciliation had already been done with the multiplicity counts. Other counts were obtained from single respondents for whom no reconciliation was necessary. This appendix discusses the conditions used to reconcile differing household-level person counts when the pair belonged to a pair relationship that corresponded to different pair domains than the one being counted. Typically, the counts between the two pair members were in agreement, and the determination of the final household-level count did not involve a reconciliation of counts, though assigning a final count meant ensuring that pair relationships were not hidden due to the relationships of the two pair members to other household members.¹ A similar situation occurred if one pair member had bad data in the household roster. The count from the pair member with good data was usually preferred in those cases, provided pair relationships of interest were not hidden. If bad data existed in either household roster, but the bad data was limited to bad relationship codes, then the member with good data was selected only if substituting the appropriate relationship codes for the bad data codes would have given a total that was equal to the count from the pair member with good data. There were instances where bad data codes existed in the roster, and this condition did not apply. There were other exceptions as well. Finally, there were instances where neither pair member had bad data in their rosters, yet their counts still disagreed. In this appendix, the rules that were used to assign a final count, as well as to reconcile disagreeing counts, are outlined. For each pair domain, a set of general rules are given, each with specific conditions required for the general rule to be implemented. Within each general condition, if at least one of the specific conditions was not satisfied, upper and lower bounds were determined and the final count was left to imputation.

J.2 Parent-Child Counts

For parent-child counts where the pairs were not parent-child pairs of interest (e.g., sibling-sibling pairs, parent-child pairs where the child was 21 or older, etc.), the screener was used to help reconcile disagreeing counts. The rules follow below, separated by the member of focus:

Parent-child pairs, child focus. For the child-focus counts, the count is of the number of children of a parent in the household. The following general rules applied:

1. In most cases, the counts of children in the relevant age range with parent(s) in the household (abbreviated below as children with parent(s) in the household) for the two

¹ If a roster pointed to a household size of one, this was considered "bad data" since both pair members in the household were survey respondents.

- sides agreed. However, both sides had to meet the following conditions for the final count to be set to one of the sides:
- Either no bad ages with the relevant relationship codes and no bad relationship codes within the relevant age ranges, or the counts of children with parent(s) in the household were equal to the screener age counts, or a side with good data indicated siblings within the relevant age range living together in a household without parents;
 - No situations where parents were not identified in the household, but some in the household had bad relationship codes and were old enough to be parents;
 - No counts of one child in the relevant child-age range when both members of the pair were in that range and the children were siblings;
 - No pairs where the ages of the identified parents did not match, the pair members were not siblings, and both sides had relationship codes signifying "other relative" or a nonrelative, indicating more than one family unit in the household;² and
 - The household size was greater than 1 and was nonmissing on both sides.
2. The counts of children with parent(s) in the household might have agreed even though the above conditions were not met. The final count of children with parent(s) in the household could still have been set to one of the sides, if any one of the following was true:
- If the number of children within the relevant age ranges matched across both rosters and the screener and (at least) one side had all good age and relationship codes, provided the equal counts did not refer to different children;³
 - If both sides had a count of zero children with parent(s) in the household, both had a roster, and (at least) one side had all good age and relationship codes;
 - If both sides had a count of zero children with parent(s) in the household, both had a roster, and the number of respondents who were old enough to be parents in the household was zero according to the screener; or
 - If the counts of children with parent(s) in the household that agreed with each other equalled or exceeded the count of the number of children from the screener within the relevant age ranges.
3. The counts of children with parent(s) in the household might have agreed with a value of 1. If both pair members were children within the relevant age range, and both indicated they had parents even though the children were siblings, then they were not included in each other's rosters, but they were obviously in the screener roster, so the final count of children with parent(s) in the household was set to 2.

² Codes that indicate "other relative" or a nonrelative are 7 (roommate), 8 (child-in-law), 10 (parent-in-law), 12 (boarder), 13 (other relative), and 14 (other nonrelative).

³ This was determined by excluding situations where the ages of the identified parents did not match, the pair members were not siblings, and both sides had relationship codes signifying "other relative" or a nonrelative, indicating more than one family unit in the household.

4. If one pair member did not have a valid roster but the other member did, the final count of children with parent(s) in the household was set to the other pair member's count under the following conditions:
 - No counts of one child with parent(s) in the household when both members of the pair were children in the relevant age range and the children were siblings, and
 - Either:
 - There were no bad relationship codes within the relevant child-age ranges and the respondent identified parents or children in the household,
 - There were no children within the relevant age range, or
 - No parents were identified in the household and nobody in the roster older than the respondent had a bad relationship code.
5. If one pair member did not have a valid roster but the other member did, and the above conditions were not met, it was still possible to use the other pair member's count of children with parent(s) in the household, if that count was 0, under any of the following conditions. Either:
 - The other roster was valid, did not have any bad ages, and had no ages in the relevant age range;
 - The other roster also was bad but the screener roster was valid and did not have any ages in the relevant age range; or
 - The respondent identified both grandchildren and grandparents in the roster where the "grandchild" relationship code(s) were incorrectly entered into the respondent's household roster. The "grandchildren" that these relationship codes were referring to were not the respondent's grandchildren, but, rather, they were the respondent's grandparent's grandchildren.⁴
6. When two different family units were in the household, the determination of the final count of children with parent(s) in the household had to be treated separately. This could have included the multigenerational families referred to earlier and the two siblings both with children in the relevant age range living in the household. The latter was more easily identified if it was not a parent-child pair (e.g., a cousin-cousin pair). The sum of the two counts of children with parent(s) in the household (one count might be 0) was used as the final count, provided the following conditions were satisfied on both sides:
 - There were no bad ages or relationship codes within the relevant age ranges;
 - Both had counts of children with parent(s) in the household pointing to two or fewer parents, meaning that the two family units were not identifiable on one side;

⁴ This condition has not manifested itself since the 2001 survey. With the addition of a new consistency check added since the 2001 survey to address grandparent/grandchild code inconsistencies, this condition could be observed only if a respondent overrode this consistency check, which has not happened.

- The number of identified parents was not equal to the total number of household members older than 25 in the household on either side, meaning that parents could correspond to roster members identified by other relationship codes;
- The number of identified children was not equal to the total number within the relevant age range in the household on either side, meaning that children with parents could correspond to roster members identified by other relationship codes; and
- There were not three generations in the household with first and second generation parents both having children in the appropriate age range. This was already accounted for by the counts for one or both sides.

If the pair was a parent-child pair, the final count was determined using imputation.

7. Two family units might be in the household but the conditions given in item #6 were not met. If there were no bad ages or relationship codes within the relevant age ranges (for both children and parents), the two families in the household might have been already accounted for when the counts of children with parent(s) in the household were determined for each side. The maximum of the two counts was used as the final count if the household members in the roster older than 25 (of parental age) were either both equal to the number of household members older than 25 in the screener roster or both different than the number of members older than 25 in the screener roster. However, if the number of household members older than 25 in the screener roster was equal to the number of members older than 25 in one of the pair member's rosters but not the other, then the count of children with parent(s) in the household corresponding to the pair member with a roster matching the screener roster (among household members of potential parental age) was used as the final count of children with parent(s) in the household.
8. If one pair member did not have a valid roster and the pair member with a valid roster was within the valid age range and was a sibling to the other pair member, but the count of children with parent(s) in the household from his roster was only 1, then the final count was set to 2.
9. If the pair relationship was not parent-child nor was it sibling-sibling, but one side had nonzero counts of children with parent(s) in the household and the other did not, it was necessary to decide who to believe. This occurred often because one of the respondents was a relative outside the nuclear family unit—like a cousin or aunt/uncle—whose own parents did not live in the household, or the respondent was a boarder.⁵ Selecting either the zero count or nonzero count in this instance required that the following conditions were met:
 - The respondent with a zero count of children with parent(s) in the household did not identify parents in the roster or he or she identified parents but was older than 20 and had no bad relationship codes within the relevant age ranges, and

⁵ Even if there was disagreement between the respondents about whether a boarder or other family member was in fact a sibling, parent, or child, this would have been resolved at the pair relationship stage where it would have been determined whether this was in a domain of interest.

- Either the respondent with a nonzero count of children with parent(s) in the household had siblings or children within the relevant age range, or the respondent himself or herself was within that age range (with a count of 1).

When one count of children with parent(s) in the household was zero and the other was nonzero, the nonzero count was used under the following conditions:

- The respondent pair member with a nonzero count also did not have bad relationship codes within the relevant age ranges, and
- Either:
 - The count of children within the relevant age range in the household for the nonzero count pair member matched that of the zero count pair member, and the count of children with parent(s) in the household did not exceed the screener count of children within the relevant age range;
 - The count of children in the household within the relevant age range for the nonzero count pair member matched that of the screener;
 - The count of children in the household within the relevant age range for the zero count pair member matched that of the screener because a child was (or children were) listed as 11 years old in the nonzero count pair member's roster, when he or she (they) should have been 12 (according to the zero count pair member's and the screener roster) so that the final count was the nonzero count with this child (these children) added;
 - The respondent with a zero count had no household members with a family-type relationship code and the reported household sizes of the two pair members were equal (indicating that it was unlikely that anyone had entered or left the household between interviews);
 - The respondent with a nonzero count showed a parent-child relationship existed in the household, but the respondent with a zero count did not because he was not related to the other household members. However, the count of children within the relevant age range in the household for the zero count was closer to the screener age count. Nevertheless, the nonzero count was equal to or less than the screener age count; or
 - The other conditions had not already established a nonzero count, but a count for a subset age group had already been established as nonzero. For example, if the count for 12- to 14-year-olds was nonzero, then the 12- to 17-year-old count had to be nonzero.

The zero count of children with parent(s) in the household was used if the zero-count respondent had no bad relationship codes at all, and either:

- The household age composition among the relevant age ranges for the zero count pair member more closely matched the screener, or
- The pair was a grandparent-grandchild pair with an adult child of the grandparent living in the household. The nonzero count resulted from an assumption that a

respondent's adult child and grandchild within the relevant age range were a parent-child pair. If the grandchild identified the grandparent's child as "other relative" and did not identify any parents, this indicated that the grandparent's adult child was an uncle/aunt of the grandchild, not a parent.

10. If the pair relationship was not parent-child nor was it sibling-sibling, but one side had nonzero counts of children with parent(s) in the household and the other did not, taking the side that was closest to the screener sometimes meant that the count of children with parent(s) from neither pair member was chosen. As with the previous item, a zero count and a nonzero count often occurred because one of the respondents was a relative outside the nuclear family unit—like a cousin or aunt/uncle—whose own parents did not live in the household, or the respondent was a boarder. If neither the zero count nor the nonzero count was chosen, the final count could still have been determined using either the screener count, the count of children within the relevant age range for the respondent with a zero count, or one less than the nonzero count. One of these was chosen, provided that the following conditions were met:
 - The respondent with a zero count of children with parent(s) in the household did not identify parents in the roster or he or she identified parents but was older than 20 and had no bad relationship codes within the relevant age ranges, and
 - Either the respondent with a nonzero count of children with parent(s) in the household had siblings or children within the relevant age range, or the respondent himself or herself was within that age range (with a count of 1).

The screener count was chosen if either:

- The respondent pair member with a nonzero count also did not have bad relationship codes within the relevant age ranges. The count of children within the relevant age range in the household for the nonzero count pair member matched that of the zero count pair member, and the nonzero count exceeded the screener count of children within the relevant age range.
- The respondent with a nonzero count showed a parent-child relationship existed in the household, but the respondent with a zero count did not because he was not related to the other household members. However, the count of children within the relevant age range in the household for the respondent with the zero count was closer to the screener age count, and the nonzero count exceeded the screener count of children.

In situations where a respondent with a zero count had a roster more closely resembling that of the screener, but the screener included a household member within the relevant age range who was not part of the immediate family, neither the nonzero count of children with parent(s) in the household nor the screener count of children within the relevant age range could be used—a different count had to be used. Two strategies were employed:

- For the respondent with a nonzero count of children with parent(s) in the household, the nonzero count was the same as the count of children within the relevant age range in the household, but it exceeded the number-of-children count

for the zero-count respondent. However, the count of children within the relevant age range for the zero-count respondent, which was not zero, was closer to the screener age count than the nonzero-count respondent.

- If the count of children within the relevant age range for the zero-count respondent was the same as the nonzero count of children with parent(s) in the household, the number-of-children count for the zero-count respondent could not be used, since the nonzero count included a household member that was not in the appropriate age range at the time of screening. One less than the nonzero count of children with parent(s) in the household was therefore chosen as the final count.
11. Other situations with a zero and nonzero count did not necessarily mean that the relationship was something other than parent-child or sibling-sibling. This was usually due to one pair member having missing relationship codes for the roster member that would have been identified as a parent (i.e., relationship codes for roster members in a parental age range). If the count for the pair member with the entirely good roster was equal to the number within the relevant child age range for the pair member with bad relationship codes in the roster, the nonzero count was selected.
 12. The two counts of children with parent(s) in the household might have disagreed where both were nonzero and both exceeded the screener count of children within the relevant age range. For the screener count to be chosen as the final household count of children with parent(s) in the household, the following conditions had to be met:
 - The pair member's household rosters had to have different numbers of children within the relevant age range,
 - The pair relationship could be neither parent-child nor sibling-sibling with a zero screener count of children within the relevant age range(s),
 - The total number within the screener roster (where the minimum age was 12 years) had to be at least two, and
 - The number of children in the screener roster within the relevant age range was valid and at least as large as the final count of children with parents in the household for the next smallest age range.
 13. The two counts might have disagreed because one side had bad relationship codes within the relevant age range and the other did not. If the sum of the number of bad relationship codes with the smaller count equalled the larger count, the larger count was chosen.
 14. The two counts might have disagreed because they disagreed on the ages of one or more household members, even though each respondent's count included all the children in their respective roster. If the roster for one respondent more closely matched the screener in terms of the distribution of ages within the roster, then that respondent's count was chosen.
 15. The two counts might have disagreed because they disagreed on the ages of one or more household members and each respondent's count included all the children in their respective roster, but neither was closer to the screener count. If the screener

- count differed from each respondent's count by the same amount, was greater than 1 but less than the other, then the screener count was used as the final count.
16. If the pair relationship was parent-child and the parent-child counts were associated with the same age range, then the household-level person counts were obtained using the parent-focus multiplicity counts corresponding to the appropriate age range. However, this did not occur if the age range for the pair relationship differed from the age range for the parent-child counts. If the pair relationship was imputed to be parent-child or it was deemed parent-child even though the child did not consider the parent a "parent," but the parent answered the FIPE3 question, then the nonzero count was used as the final count.
 17. If, after all the above tests were done to find the final count, the minimum possible and maximum possible counts—considering both questionnaire rosters and the screener roster—were the same, then the final count was set to that value.
 18. Remaining disagreeing counts were left to imputation, with appropriate bounds set on the imputed value.

Parent-child pairs, parent focus. For the parent-focus counts, the count is of the number of parents of at least one child in the household. The child-focus parent-child counts are processed first, so if the child-focus parent-child counts are 0, it necessarily means that the parent-focus counts will also be 0. Nonzero child-focus counts also point to nonzero parent-focus counts. After setting counts to 0 where necessary, the following general rules applied:

1. In most cases, the counts of parents with children in the household for the two sides agreed. However, both sides had to meet the following conditions for the final count to be set to one of the sides:
 - No situations where both pair members were children in the relevant age range but were in a spouse-spouse pair relationship and both identified the same roster member as a parent,
 - The household size was greater than 1 and nonmissing on both sides, and
 - Either:
 - No bad relationship codes for household members of an age to be parents,
 - The total count was 2 for two parents, or
 - The total count plus the number of grandparents equalled the total number of household members aged 26 or older, according to the screener roster.

Note that it was not necessary to check for bad relationship codes in the child age ranges, since it was already known that the count had to be at least 1, and the number of children was not important for the parent counts.

2. The counts of parents with children in the household might have agreed even though the above conditions were not met. The final count could still have been set to one of the sides if it was a sibling-sibling pair, and the bad codes in the parental age range were on one side only. This would indicate that the side with bad codes were not missing parental codes.

3. If one pair member did not have a valid roster but the other member did, the final count of parents with children in the household was set to the other pair member's count if there were no bad relationship codes and no roster members with bad age and bad gender values. Other circumstances called for setting the final count to 0, which would necessarily be the case if the child-focus counts were 0.
4. When two different family units were in the household, the determination of the final count of parents with children in the household had to be treated separately. This could have included multigenerational families or two siblings both with children in the relevant age range living in the household. The latter was more easily identified if it was not a parent-child pair (e.g., a cousin-cousin pair). The sum of the two counts (one count might be 0) was used under the following conditions:
 - There were no bad ages or relationship codes within the relevant age ranges,
 - Both pair members had counts pointing to 2 or fewer parents, meaning that the two family units were not identifiable on a side,
 - The number of identified parents was not equal to the total number of household members older than 25 on either side, meaning that parents could correspond to roster members identified by other relationship codes, and
 - There were not three generations in the household, with first and second generation parents both having children in the appropriate age range. This was already accounted for by the counts for one or both sides.
5. Two family units might be in the household but the conditions given in item #4 were not met. If there were no bad ages or relationship codes within the relevant age ranges (for both children and parents), the two families in the household might have been already accounted for when the counts of parents with children in the household were determined for each side. The maximum of the two counts was used as the final count if the household members older than 25 (of parental age) in the roster were either both equal to the number of members older than 25 in the screener roster or both different than the number of members older than 25 in the screener roster. However, if the number of household members older than 25 in the screener roster was equal to the number of members older than 25 in one of the pair member's rosters but not the other, then the count of parents with children in the household corresponding to the pair member with a roster matching the screener roster (among household members of potential parental age) was used as the final count of children with parent(s) in the household.
6. If the pair relationship was a spouse-spouse pair and one of the pair members was within the relevant age range and had a positive count, then the count for that pair member was taken as the final count, provided there were no bad relationship codes in that roster for roster members aged 18 or older.⁶
7. The two counts might have disagreed with one nonzero count and the other equal to zero. Due to the fact that the counts of parent(s) in the household with children were

⁶ For this condition, either the count for the other pair member was 0 or the count for the pair members was equal.

determined first and that the zero counts were handled separately, the final count of parents with children in the household determined at this stage of processing had to be nonzero. Counts arising from two or more families in the household also were handled in previous code. Hence, the final count had to be one or two parents.⁷ The nonzero count was chosen as the final count if one of the following conditions were met:

- The count was 1 and there were no bad ages with the relevant relationship codes and no bad relationship codes within the relevant age ranges, or
 - The count was 2.
8. The two counts might have disagreed where the number of roster members aged 26 or older disagreed between the two pair members. In these situations, one count was 1, and the other count was 2. The final count corresponded to the pair member with the number of roster members aged 26 or older closest to the screener number of roster members aged 26 or older, under the following conditions:
 - The difference between the screener count of the number of household members aged 26 or older and the pair members' counts of this number of household members was not the same between the two pair members,
 - Neither pair member had bad ages in their rosters, and
 - Each pair member either had no bad relationship codes in his or her roster or had a nonzero count with no bad relationship codes among respondents aged 26 or older.
 9. The two counts might have disagreed if the bad relationship codes referred to missing parental codes. If one side had no bad relationship codes, and the sum of the number of bad relationship codes and the count on the side with the bad codes was equal to the count on the side with no bad relationship codes, then the count from the side with no bad relationship codes was used as the final count.
 10. The two counts might have disagreed where one count was 2 and the other was 3. Since households with two family units had already been considered, the maximum number of parents possible was two, so the final count was set to 2.
 11. If the pair relationship was parent-child and the parent-child counts were associated with the same age range, then the household-level person counts were obtained using the child-focus multiplicity counts corresponding to the appropriate age range.
 12. If, after all the above tests were done to find the final count, the minimum possible and maximum possible counts—considering both questionnaire rosters and the screener roster—were the same, then the final count was set to that value.
 13. Remaining disagreeing counts were left to imputation, with appropriate bounds set on the imputed value.

⁷ This precluded the extremely unlikely possibility that the pair member with a zero count masked a situation where three parents in a single family unit lived in the household (two biological parents and a stepparent).

J.3 Sibling-Sibling Counts

The logic for the sibling-sibling counts did not depend upon whether the lower age range was 12 to 14 or 12 to 17 or whether the upper age range was 15 to 17 or 18 to 25. It also did not depend upon which pair member was the focus, though for the household-level person counts, the older member focus counts were the only ones considered. Hence, the counts of interest are of roster members in the upper age range. As with the parent-child pairs, the multiplicity counts could be used if the pair relationship was a sibling-sibling pair of interest. However, the counts had to be determined for all other pairs. The rules follow below, separated by the member of focus:

1. In most cases, the counts for the two sides agreed. However, both sides had to meet the following conditions for the final count to be set to one of the sides:
 - The pair could not be a sibling-sibling pair, where both respondents were in the upper age range, and could not have a younger sibling in the lower age range, and the count was 1. (This refers to a sibling-sibling pair that would not constitute a domain of interest.)
 - No bad relationship codes in the lower range if the count was 0.
 - Either:
 - No bad relationship codes in the upper range, or
 - The count matched the screener age count.
 - The household size was greater than 1 and nonmissing on both sides.
2. The counts might have agreed even though the above conditions were not met. The count could still have been set to one of the sides if any one of the following conditions was true:
 - If the number of children matched across both rosters and the screener for both the upper and lower age ranges, or
 - If the count was 0 and one of the following two conditions was true:
 - Neither side had bad relationship codes or ages, or
 - The number of household members aged 26 or older in the screener roster was zero.
3. If one pair member did not have a valid roster but the other member did, the final count was set to the other pair member's count under the following conditions:
 - No bad relationship codes within the lower age range when the count was 0.
 - Either:
 - There were no bad relationship codes within the upper age range,
 - The count was equal to the screener age count within the upper age range, or
 - The count was 0, and the count of household members in the lower age range was 0.

4. If one pair member did not have a valid roster but the other member did, and the above conditions were not met, it was still possible to use the other pair member's count under the following conditions:
 - The count was 0,
 - The number of children in either the lower or upper age ranges was 0 with no bad ages in the roster.
5. If neither pair member had a valid roster, it was occasionally still possible to assign a final count. If the number of children in the screener roster in either the lower or upper age ranges was zero and the screener roster was valid, then it was not possible for a sibling-sibling pair in the relevant age ranges to be selected and the final count to be set to 0.
6. When two different sets of siblings were in the household, the determination of the final count had to be treated separately. The two sets of siblings refer to siblings where both parents from one set differ from the parents of the other set. The sum of the two counts (one count might be 0) was used, provided the following conditions were satisfied for both pair members:
 - The sum of counts of the number of sibling-sibling pairs equalled or exceeded at least one of the counts of household members in the upper age range for the screener roster or either of the pair member's rosters.
 - There were no bad relationship codes within the upper age ranges.
 - There were no bad relationship codes within the lower age range, or the count was nonzero.
7. If the counts from the two pair members did not agree, the following rules were used to assign the appropriate count, provided no bad relationship codes were evident in either age range on either side. These conditions are hierarchical, in that subsequent conditions require that the previous condition was not met.
 - If the number within the upper age range was the same on both sides, but the number in the lower age range was not, then the side with the number in the lower age range equal to the number in the screener roster within the lower age range was chosen. (In all cases, one side had a zero count and the other did not. This captured situations where it was necessary to discern whether the zero count was due to no children in the lower age range on one side and whether the screener also had no children in that range.)
 - For one pair member, the number of children in either the lower age range or the upper age range did not agree with the number in the screener roster in that range. However, for the other pair member, the number within both age ranges agreed with the screener count. The count was set to the side that agreed with the screener.
 - For both pair members, the numbers within the lower age range were either both zero or both positive. The number within the upper age range did not agree between pair members, but one pair member agreed with the screener. The final count was set to the count for that pair member.

- In the rosters for both pair members and the screener, the numbers within the upper age range for at least one of the three were nonzero but not necessarily equal. The numbers within the lower age range were not equal across any of the three rosters. The pair member with the number of children in the upper age range closest to the screener was selected.
8. If the counts from the two pair members did not agree, but one side had bad relationship codes within the upper age range and the other did not have bad relationship codes, and the sum of the count and the number of bad relationship codes on one side was equal to the count for the pair member with the good roster, then the count for the pair member with the good roster was selected.
 9. If the counts from the two pair members did not agree, and the above conditions were not met, in many cases this was due to one of the pair members not being part of the immediate family unit, in which case his or her count was automatically 0. To identify these cases and assign the count to the other pair member, the following conditions had to be satisfied:
 - The pair relationship did not indicate an identifiable family-type relationship (e.g., sibling-sibling, parent-child, spouse-spouse, or grandparent-grandchild relationship).
 - Either:
 - One pair member did not have any relationship codes indicating parent, child, sibling, spouse, grandchild, or grandparent;
 - The other pair member had at least one relationship code indicating a relationship other than parent, child, sibling, spouse, grandchild, or grandparent;
 - For the pair member with family codes, either no bad relationship codes were within both the upper and lower age ranges or no bad relationship codes were within the upper age range, and the count was positive; or
 - There were no bad relationship codes within both the upper and lower age ranges for either pair member.
 10. If one pair member had no bad relationship codes within both the upper and lower age ranges, but the other member had some bad codes, then the count associated with the pair member with no bad codes was selected if the count of immediate family members (parent, child, sibling, spouse, grandchild, or grandparent) was the same as the count of household members within both the lower and upper age ranges.
 11. If one pair member had a zero count due to having no household members within the upper age range, but the number of household members within that age range was nonzero for both the screener and the other pair member (though not necessarily equal), and the count for the other pair member was equal to the number of household members within the upper age range for that pair member, then a nonzero count was selected. If the number of household members within that age range in the screener roster was nonzero, then that number was chosen as the final count. Otherwise, the

number of household members within the upper age range for the pair member with nonzero count was selected as the final count.

12. If the pair was a spouse-spouse pair, one count might have been zero while the other was nonzero because the spouse-spouse pair still lived with the parents of one pair member, and the pair member's younger siblings also lived in the household. In this case, the nonzero count was selected if the number of immediate family members (parent, child, sibling, spouse, grandchild, or grandparent) in the roster for the pair member with the zero count was less than his or her total household size.
13. In some cases, one pair member called the other pair member a parent or child, but the other pair member did not reciprocate. In the case of a child who did not reciprocate the parent's identification of him or her as a child, the child's count was always less than the parent's count. By the same token, in the case of a parent who did not reciprocate the child's identification of him or her as a parent, the parent's count was always less than the child's count. If the pair relationship was imputed to be "parent-child," then the pair member who did not acknowledge a parent-child relationship was overruled, and the maximum count of the two pair members was selected as final.

J.4 Spouse-Spouse Counts (with or without Children)

The multiplicity counts were not useful in the logic for the spouse-spouse household counts, since the spouse-spouse multiplicity counts were always 1.⁸ If the household size was one, or the number of respondents aged 15 or older in the household was one or zero, then the final household person count was set to 0 since no spouse-spouse pairs could reside under those limits. If two family units had been previously identified in the household, the following rules were used to determine the final household person count:

1. When two different family units were already identified in the household, then two different parent sets were being referenced (one of the parent sets was often a single parent). The sum of the two counts (one count might be 0) was used, provided neither pair member had grandparents or grandchildren identified. This was to prevent spouse-spouse pairs from being counted twice, which would happen if grandparents were also parents of children younger than 18 years of age. If two family units were multigenerational families, then the final count was obtained by taking the maximum of the two pair members' counts.
2. It was possible for two different spouse-spouse pairs to be in the household, even though two different family units had not been identified. The final count was set to 2, even though two family units had not been previously identified, under the following conditions:
 - The pair relationship was not a spouse-spouse pair, and the total household size was at least four; and

⁸ In rare cases, an individual might identify two spouses in the household. As noted in Section 14.2.2, the true multiplicity count in these cases was not determined; rather, the multiplicity count was set to 1, due to the complexity of determining the appropriate multiplicity count and the rarity of the occurrence of multiple spouses.

- Either:
 - Both sides identified a spouse,
 - Both sides identified a partner, or
 - One side identified a parent and the other side identified a parent-in-law.
- 3. If the conditions for the previous item were not met, it was still possible for two different spouse-spouse pairs to be in the household, even though two different family units were not previously identified. The final count was set to 2 under the following conditions:
 - One pair member had two parents with valid ages and both ages differed from the age of the spouse of the other pair member, and
 - The pair relationship was either sibling-sibling or a pair that was not a pair of interest.

Otherwise, reconciling the counts to a nonmissing value always required the following condition: There was no potential for two or more couples in the household that were not already obviously identified, whereby one of the pair members had at least four roster members of at least 15 years of age. This respondent had grandchildren younger than 18 years of age, did not have children-in-law, and had household members aged 12 or older who were not children, grandchildren, siblings, children, parents, spouses, or partners. For all remaining cases where a final household count needed to be assigned—in addition to the above condition—the final count was assigned using the following rules:

- 4. Among the majority of pairs, the counts for the two sides agreed. However, both sides had to meet the following conditions for the final count to be set to one of the sides:
 - The pair could not be a spouse-spouse pair where both respondents had a spouse or both respondents had a partner,
 - No bad relationship codes for roster members aged 15 or older for either pair member,
 - The number of spouse-spouse pairs was either one or zero for both pair members,
 - The household size was greater than 1 and nonmissing on both sides,
 - One pair member had at least two household members aged 15 or older, and
 - There were not two spouse-spouse pairs in the household according to the conditions given in item #3.
- 5. The counts might have agreed even though the above conditions were not met. The count could still have been set to one of the sides if any one of the following was true:
 - One pair member was younger than 18 and had no bad relationship codes for roster members aged 18 or older, but he or she did have bad relationship codes for roster members between the ages of 15 and 17 years old. The other pair member had no bad relationship codes for roster members aged 15 or older.
 - One pair member had a single bad relationship code, and no other relationship codes could match it to make it a couple (i.e., the pair member did not have a

single identified parent, grandparent, parent-in-law, or child-in-law). The other pair member had no bad relationship codes.

- One pair member had bad relationship codes among roster members aged 15 or older or had bad ages, and the other had no bad ages or relationship codes, where the pair member with no bad roster entries had the same number of household members aged 15 or older as the screener. The pair member with the bad roster entries would not have had the same age composition as the screener if the number of roster members aged 15 or older was added to the number of roster members with bad ages.
 - One pair member had bad relationship codes among roster members aged 15 or older or had bad ages, and the other had no bad ages or relationship codes, where all the relationship codes for the pair member with no bad roster entries were immediate family codes (child, parent, sibling, spouse, partner, grandparent, or grandchild). For the pair member with bad roster entries, all the existing relationship codes were immediate family codes.
6. For those cases where the pair was imputed to be a spouse-spouse pair and both sides agreed that only one spouse-spouse pair was in the household, the count was set to 1 if any one of the following conditions was true:
- Both sides had fewer than four people older than 15 in the household, or
 - One side had fewer than four people older than 15 in the household, and the other side had no bad relationship codes among roster members aged 15 or older
7. If one pair member did not have a valid roster but the other member did, the final count was set to the other pair member's count under any one of the following conditions:
- There were no bad relationship codes among roster members aged 15 or older, or
 - There were no bad relationship codes among roster members aged 18 or older and the pair member had parents.
8. If the count of the number of spouse-spouse pairs did not agree between the two pair members, it could have been because a couple entered the household or otherwise materialized after screening. The smaller count was chosen as the final count in this instance, which was identified if the following conditions were satisfied:
- The screener count of roster members aged 12 or older was no larger than the count of roster members aged 12 or older in the roster of the pair member with the smaller spouse-spouse count.
 - The screener count of roster members aged 12 or older was smaller than the count of roster members aged 12 or older in the roster of the pair member with the larger spouse-spouse count.
 - The difference between the screener count of roster members aged 12 or older and the count of roster members aged 12 or older in the questionnaire rosters of the pair members was smallest with the pair member with the smaller spouse-spouse count.

9. If the count of the number of spouse-spouse pairs did not agree between the two pair members, it could have been because a couple left the household or otherwise dissolved after screening. The larger count was chosen as the final count in this instance, which was identified if the following conditions were satisfied:
 - The screener count of roster members aged 12 or older was no larger than the count of roster members aged 12 or older in the roster of the pair member with the larger spouse-spouse count.
 - The screener count of roster members aged 12 or older was larger than the count of roster members aged 12 or older in the roster of the pair member with the smaller spouse-spouse count.
10. In many cases where the count of the number of spouse-spouse pairs did not agree between the two pair members, one side had a zero count and the other did not. The nonzero count was selected if the pair member associated with the zero count was not a close relative or somehow did not identify a spouse, partner, two parents, or two grandparents. The following conditions were required to select the nonzero count:
 - The pair member with a nonzero count either identified a spouse, a partner, two parents, or two grandparents.
 - The number of roster members aged 15 or older associated with the nonzero count pair member was no larger than the corresponding number associated with the zero count pair member.
 - If the side associated with the nonzero count identified a spouse, partner, or two parents, the following additional conditions were required:
 - The number of roster members between the ages of 26 and 44 was the same between the two pair members.
 - The number of roster members between the ages of 30 and 49 was the same between the two pair members.
 - The number of roster members between the ages of 35 and 54 was the same between the two pair members.
 - The number of roster members between the ages of 40 and 59 was the same between the two pair members.
 - If the side associated with the nonzero count identified two grandparents, the following additional condition was required:
 - The number of roster members aged 50 or older was the same between the two pair members.
11. The counts might not agree because a pair member's partner did not consider the other pair member's family as his or her own family. If at least one side identified a partner and the maximum count was 1, then the maximum was selected if both pair members had the same number of household members aged 15 or older. Otherwise, if the pair members had a different number of household members aged 15 or older, the count belonging to the pair member with a count of household members aged 15 or older closer to that of the screener was used as the final count.

12. The counts might not agree because a pair member had two grandparents and an uncle/aunt husband-wife pair in the household. The maximum was selected if the pair member associated with the smaller count had a grandparent and had at least two roster members who were neither parents, siblings, children, spouses, partners, or grandparents, and the pair member with the larger count had children-in-law.
13. The count of the number of spouse-spouse pairs might not agree because one of the pairs was a sibling and sibling-in-law, and there are no codes for sibling-in-law. The maximum count was selected if the pair member with the smaller count did not have a spouse or partner but did have siblings aged 15 or older, and there were household members in his or her roster that were not parents, children, siblings, spouses, partners, grandchildren, or grandparents.
14. The count of the number of spouse-spouse pairs might not agree because one side had no nuclear family or grandparent-grandchild relationship codes, and one of the selected respondents was not in a child-parent, child-grandparent, or spouse-spouse relationship. The maximum count was selected if the following conditions were met:
 - The pair member's roster associated with the minimum count (usually 0) had no children, parents, siblings, spouses, partners, grandchildren, or grandparents among respondents aged 12 or older; and
 - The pair member's roster associated with the maximum count had some roster members who were not children, parents, siblings, spouses, partners, grandchildren, or grandparents.

Note that this condition also nabbed cases where the relationship codes were not correctly identified on one pair member's roster. This occurred rarely, but when it did, the minimum count was 1 and the maximum count was 2.

15. The count of the number of spouse-spouse pairs might not agree because the pair members were siblings, and one sibling did not consider a stepparent or parent's partner as a "parent." The maximum count was selected if the following conditions were met:
 - The pair members were siblings,
 - The pair member associated with the maximum count had two parents,
 - The pair member associated with the minimum count had one parent, and
 - The roster associated with the pair member with the maximum count had more immediate family members (children, parents, siblings, spouses, partners, grandchildren, or grandparents) than the roster associated with the other pair member.
16. The count of the number of spouse-spouse pairs might not agree because the household changed after screening, which was not accounted for by previous conditions. In general, the count with a household composition closest to the screener was selected. The age composition was defined by looking at age classes. The count for a given pair member was selected if any of the following properties held:

- The number of roster members between the ages of 26 and 44 for that pair member matched the screener count within the same age range, which differed from the corresponding count for the other pair member.
 - The number of roster members between the ages of 30 and 49 for that pair member matched the screener count within the same age range, which differed from the corresponding count for the other pair member.
 - The number of roster members between the ages of 35 and 54 for that pair member matched the screener count within the same age range, which differed from the corresponding count for the other pair member.
 - The number of roster members between the ages of 40 and 59 for that pair member matched the screener count within the same age range, which differed from the corresponding count for the other pair member.
17. In some cases, neither pair member's household composition matched that of the screener. In that case, the household roster closest to that of the screener was selected. The maximum was selected if the number of screener roster members aged 12 or older exceeded the corresponding count from the questionnaire rosters of both pair members, which also differed from each other.
18. The counts might not agree because, on the rare occasion, one pair member in a spouse-spouse pair identified two grandparents of a different gender. Since there is no code for grandparents-in-law, they could not be identified, so the maximum count was selected. The following conditions were required:
- The pair was a spouse-spouse pair.
 - The pair member with the maximum count had two grandparents of a different gender, and the pair member with the minimum count did not have any.
- The assumption here, of course, is that the grandparents of a different gender were in fact a spouse-spouse pair. There was no way to check whether a grandfather was the father's father and the grandmother was the mother's mother, for example.
19. Even though the household composition may match in terms of ages across the screener roster and the two pair members' rosters, the counts may disagree where two spouse-spouse pairs were clearly identified by one pair member but not the other. This may be because one of the in-laws was incorrectly identified on one side, or because a partner was not considered an in-law by a responding pair member, or because a partner did not consider other family members as "in-laws." The following conditions were required for the maximum count to be selected:
- The number of screener roster members aged 12 or older matched the corresponding count from the questionnaire rosters of both pair members.
 - The pair member with the maximum number of spouse-spouse pairs had a spouse or partner and also had two parents.
 - There were no bad relationship codes among roster members aged 15 or older on either pair member's roster.

20. If the counts for each pair member were not equal but the number of roster members aged 12 or older was the same between the two pair members, and the count for one pair member was the maximum possible in the household, then that number was selected as the final count. This condition was applied only after all other conditions, including conditions where the final count was ambiguous, had already been applied.
21. After accounting for all other rules, if the number of spouse-spouse pairs was still missing, but the lower and upper bounds for imputation were equal to each other, then the final household-level person count was set to one of those bounds.

J.5 Spouse-Spouse Counts (with Children)

The household counts for spouse-spouse counts with children obviously depended upon the counts obtained for spouse-spouse counts with or without children. The first two rules described in this section were determined directly from the spouse-spouse counts or from the household size, and no reconciliation of counts was necessary:

1. For a sizable proportion of cases, clearly no couples with children could be in the household, either because the spouse-spouse count was 0 or the household size was two or less. In these cases, the final spouse-spouse-with-children count was set to 0.
2. An additional small number of cases also could be readily determined by looking at the spouse-spouse count. If one pair member had a spouse-spouse-with-children count that exceeded the final spouse-spouse count, but the other pair member had a spouse-spouse-with-children count that was equal to or smaller than the final spouse-spouse count, then the final spouse-spouse-with-children count was set to the pair member's count that was consistent with the final spouse-spouse count.

The remainder of cases involved households with at least one spouse-spouse couple. After assigning values for the conditions described above, the assignment of values for these cases was done using the rules described in the rest of this section. If two family units had been previously identified in the household, the following rule was used to determine the final household person count:

3. When two different family units were already identified in the household, then two different parent sets were being referenced (one of the parent sets was often a single parent). The sum of the two counts (one count might be 0) was used, provided the spouse-spouse count was greater than 1. In that event, the maximum count was used.

Otherwise, reconciling the counts to a nonmissing value always required the following condition: There was no potential for two or more couples in the household that were not already obviously identified, whereby one of the pair members had at least four roster members of at least 15 years of age. This respondent had grandchildren younger than 18 years of age, did not have children-in-law, and had household members aged 12 or older who were not children, grandchildren, siblings, children, parents, spouses, or partners. For all remaining cases where a final household count needed to be assigned—in addition to the above condition (unless specifically noted below)—the final count was assigned using the following rules:

4. For cases that were not already determined by looking at the previous two conditions, the counts for the two pair members (if there were two pair members) were equal in

the vast majority of cases. The final count could be set to each pair member's count under the following conditions:

- Both pair members had valid rosters.
 - Either:
 - The counts were nonzero and equal to the final spouse-spouse count, or
 - There were no bad relationship codes for roster members younger than 18, and one of the following conditions held for at least one pair member:
 - The pair member's roster had no bad relationship codes for roster members aged 15 or older,
 - The pair member was older than 18 and had neither children nor siblings younger than 18 (covers zero counts since no bad codes were for members younger than 18), or
 - The pair member was younger than 18 and did not have parents, but there was one bad relationship code among roster members older than 18 in that pair member's roster (covers zero counts since only one bad relationship code could potentially be a single parent but not a pair of parents making a couple).
5. The pair members might both have had zero counts, but the above conditions did not apply. The final count could still have been 0 if the age counts for both pair members and the screener indicated nobody lived in the household who was younger than 18 and there were no bad roster ages. (In this case, it was not necessary to check for the potential of two or more family units in the household.)
6. The counts for both pair members might still have agreed with nonzero counts, even though none of the previous conditions applied. The final count could still have been set to one of the pair member's counts if the pair relationship was imputed to be a spouse-spouse pair with children.
7. If one pair member did not have a valid roster but the other member did, the final count was set to the other pair member's count under one of the following conditions:
- The count for the pair member with the valid roster was nonzero and equal to the final spouse-spouse count, or
 - There were no bad relationship codes for roster members younger than 18, and one of the following conditions held for the pair member with the valid roster. Either:
 - The pair member's roster had no bad relationship codes for roster members aged 15 or older,
 - The pair member was older than 18 and had neither children nor siblings younger than 18 (covers zero counts since no bad codes were for members younger than 18), or
 - The pair member was younger than 18 and did not have parents, but there was one bad relationship code among roster members older than 18 in that pair

member's roster (covers zero counts since only one bad relationship code could potentially be a single parent but not a pair of parents making a couple).

8. The pair member with the valid roster might have had a zero count, but the above conditions did not apply. The final count could still have been 0 if the age counts for both the pair member with the valid roster and the screener indicated nobody lived in the household who was younger than 18 and there were no bad roster ages. (In this case, it was not necessary to check for the potential of two or more family units in the household.)
9. If the spouse-spouse-with-children counts disagreed in the same manner as the spouse-spouse counts disagreed, then the choice was obvious: Use the count that corresponded to the correct spouse-spouse count. (In this case, it was not necessary to check for the potential of two or more family units in the household.) Details follow:
 - If the spouse-spouse-with-children counts were equal to the spouse-spouse counts for both pair members, even though they were unequal to each other, then the final spouse-spouse-with-children count was set to the final spouse-spouse count.
 - If the spouse-spouse counts exceeded the spouse-spouse-with-children counts by one for each pair member, even though they were unequal to each other, then the final spouse-spouse-with-children count was set to one less than the final spouse-spouse count.
10. Based on earlier conditions, households without couples were already excluded. Households with a possibility of two or more couples were also excluded. If the pair relationship was parent-child and at least one count was nonzero, then the identified couple corresponded to the parent-child relationship. The maximum of the counts was selected under the following conditions:
 - The sum of counts from the two pair members was 1.
 - Either:
 - The relationship was parent-child where the child was between the ages of 12 and 17, or
 - The relationship was parent-child where the child was between the ages of 18 and 20 and the child had siblings younger than 18.
11. In some cases, two couples were identified in the household where the household was multigenerational (one member of the younger couple was in a parent-child relationship with the older couple). If a sibling to the pair member in the younger couple was selected, or if a member of the younger couple was selected who "married into" the family, then he or she was not able to identify the nephews, nieces, brothers-in-law, or sisters-in-law—which could point to an appropriate accounting of all the couples with children—because of the relationship codes that were available. The maximum of the two counts was selected under the following conditions:
 - There were two couples in the household, as identified by the final spouse-spouse count.
 - The difference between the pair members' counts was 1.

- Either:
 - The pair member with the smaller count had a spouse or partner and the pair member with the larger count had parents in the household, or
 - The pair member with the smaller count had parents-in-law or children-in-law in the household.
12. If a couple was in a marriage/partnership that occurred after an earlier marriage, the partner might not have considered the partner's children as his or her children, but the child (who also was selected) considered the spouse/partner a parent. Even though the pair relationship was not parent-child, these cases were still counted as spouse-spouse with children since they consisted of the children of one spouse/partner. The maximum count was selected under the following conditions:
- The pair relationship was not one of interest.
 - One count was 0 and the other count was 1.
 - The pair member with the zero count had a spouse or partner.
 - The pair member with the nonzero count had parents.
 - The spouse-spouse final count was nonmissing.
13. The counts might have been unequal because children younger than 18 left, entered, or otherwise materialized or disappeared in the household after screening and between the time of the interviews. In general, the count was selected that corresponded to the pair member with a household composition closest to the screener household composition. If one pair member did not have children in the household and the other pair member did, the following conditions were required for the count corresponding to the pair member with a household composition closest to the screener:
- One pair member had a nonzero count of children younger than 18 and the other pair member had a zero count of children younger than 18.
 - Either:
 - The screener composition indicated that children younger than 18 were in the household, whereupon the nonzero count was selected, or
 - The screener composition indicated that no children younger than 18 were in the household, whereupon the zero count was selected.
14. The counts might have been unequal with a count of 0 and a count of 1 because a pair member with a count of 0 was not part of the immediate family unit. The nonzero count was used under the following conditions:
- The pair relationship was not a parent-child, sibling-sibling, spouse-spouse, or grandparent-grandchild relationship.
 - Both pair members had relationship codes that were not parent, child, sibling, spouse, partner, grandparent, or grandchild codes among roster members who were aged 12 or older.

15. The counts might have been unequal because of bad relationship codes among roster members younger than 18. The following rules were used to determine if the count associated with the pair member did not have bad relationship codes:
- The number of roster members younger than 18 was the same between both pair members.
 - The side with the smaller count had at least one bad relationship code for roster members younger than 18.
16. If, after considering all of the general conditions given above, the count was left to imputation, it was still possible that the lower and upper bounds were equal. In this instance, the final count was set to one of the bounds.

Appendix K: Priority Conditions Used to Create Household-Consistent Covariates

This page intentionally left blank

Appendix K: Priority Conditions Used to Create Household-Consistent Covariates

K.1 Household Size

In [Table K.1](#), blank entries indicate that no conditions were required for that set of variables. The reported household size variable is QD54, and the edited household size variable is TOTPEOP, which cannot differ from the raw variable by more than 1. Any variable suffixed by "A" indicates that the variable corresponds to the value for pair member "A." A similar comment can be made with regard to the suffix "B." For example, "QD54A" reflects the reported household size for pair member A. The quality-of-roster counts are considered in the column "any roster missing?" The variables GOODAGEA and GOODAGEB are the total number of cases in the roster with valid ages. The variables that appear in the table are TGOODAGA and TGOODAGB, the total number of cases in the roster with valid ages, incorporating the minimum possible counts within the age categories 12 to 17, 18 to 25, 26 to 34, 35 to 49, and 50 or older. Finally, the variable used to describe the screener household size is SHHSIZE. The conditions used to create the variable HHSIZE resulted in no missing values for this variable, and thus no imputation was required. The first column in [Table K.1](#) shows the hierarchical priority condition, with the frequency of occurrence for each priority condition in parentheses.

Table K.1 Priority Conditions Used to Create Household-Consistent Household Size

Priority Condition (Frequency)	Relationship of QD54A & QD54B	Relationship of TOTPEOPA & TOTPEOPB	Relationships Involving Age Range Variables	Any Roster Missing?	Screener Roster Characteristics	HHSIZE Equals:
1 (16,546)	Equal, both > 1, both nonmissing	Equal, both > 1, both nonmissing				TOTPEOPA
2 (0)	Equal, both > 1, both nonmissing	TOTPEOPB one more than TOTPEOPA	$TGOODAGA \leq QD54A$	A: no		QD54A
3 (0)	Equal, both > 1, both nonmissing	TOTPEOPA one more than TOTPEOPB	$TGOODAGB \leq QD54B$	B: no		QD54B
4 (0)	Equal, both > 1, both nonmissing	TOTPEOPA one more than TOTPEOPB	$TGOODAGA = TGOODAGB$ $TGOODAGA \leq TOTPEOPA$		SHHSIZE not equal to QD54A	TOTPEOPA
			$TGOODAGA = TOTPEOPA$		No condition	
5 (0)	Equal, both > 1, both nonmissing	TOTPEOPB one more than TOTPEOPA	$TGOODAGA = TGOODAGB$ $\leq TOTPEOPB$		SHHSIZE not equal to QD54B	TOTPEOPB
			$TGOODAGB = TOTPEOPB$		No condition	
6 (0)	Equal, both > 1, both nonmissing	Within one of each other			SHHSIZE at least as large or larger than screener roster, equal to QD54A	SHHSIZE
7 (0)	A: missing or 1 B: not missing > 1	A: missing or 1 B: not missing > 1, not equal to QD54B	$QD54B \geq TGOODAGB$		$SHHSIZE \geq 2$, closer to QD54B than TOTPEOPB	QD54B
8 (7)	A: missing or 1 B: not missing > 1	A: missing or 1 B: not missing > 1	$TGOODAGB \leq TOTPEOPB$ (no bad roster ages if equal)		$SHHSIZE \geq 2$, TOTPEOPB is as close as QD54B	TOTPEOPB
9 (0)	A: missing or 1 B: not missing > 1	A: missing or 1 B: not missing > 1	$TGOODAGB \leq SHHSIZE$		$TGOODAGB \leq SHHSIZE$	SHHSIZE
10 (1)	A: missing or 1 B: not missing > 1	A: missing or 1 B: not missing > 1				TGOODAGB

Table K.1 Priority Conditions Used to Create Household-Consistent Household Size (continued)

Priority Condition (Frequency)	Relationship of QD54A & QD54B	Relationship of TOTPEOPA & TOTPEOPB	Relationships Involving Age Range Variables	Any Roster Missing?	Screener Roster Characteristics	HHSIZE Equals:
11 (0)	A: not missing, > 1 B: missing or 1	A: not missing, > 1, not equal to QD54A B: missing or 1	$QD54A \geq TGOODAGA$		$SHHSIZE \geq 2$, closer to QD54A than TOTPEOPA	QD54A
12 (5)	A: not missing, > 1 B: missing or 1	A: not missing, > 1 B: missing or 1	$TGOODAGA \leq TOTPEOPA$ (no bad roster ages if equal)		$SHHSIZE \geq 2$, TOTPEOPA is as close as QD54A	TOTPEOPA
13 (0)	A: not missing, > 1 B: missing or 1	A: not missing, > 1 B: missing or 1	$TGOODAGA \leq SHHSIZE$		$TGOODAGA \leq$ $SHHSIZE$	SHHSIZE
14 (0)	A: not missing, > 1 B: missing or 1	A: not missing, > 1 B: missing or 1				TGOODAGA
15 (0)	Both missing or 1	Both missing or 1			$SHHSIZE \geq 2$, SHHSIZE at least as large or larger than screener roster	SHHSIZE
16 (32)	Not equal, both > 1	TOTPEOPB = QD54B	A: At least one age range variable less than min. ¹ B: Age range variables all same or larger than min.			QD54B
		TOTPEOPA = QD54A	B: At least one age range variable less than min. A: Age range variables all same or larger than min.			QD54A
17 (1)	Not equal, both > 1		A: At least one age range variable less than min. B: At least one age range variable less than min.		Age range variables all same or larger than min.	SHHSIZE

Table K.1 Priority Conditions Used to Create Household-Consistent Household Size (continued)

Priority Condition (Frequency)	Relationship of QD54A & QD54B	Relationship of TOTPEOPA & TOTPEOPB	Relationships Involving Age Range Variables	Any Roster Missing?	Screener Roster Characteristics	HHSIZE Equals:
18 (1,068)	Not equal, both > 1	QD54A is equal to at least one of TOTPEOPA or TOTPEOPB	A: Age range variables all same or larger than min., no bad roster ages		SHHSIZE at least as large or larger than screener roster, equal to QD54A	QD54A
		QD54B is equal to at least one of TOTPEOPA or TOTPEOPB	B: Age range variables all same or larger than min., no bad roster ages		SHHSIZE at least as large or larger than screener roster, equal to QD54B	QD54B
19 (1)	Not equal, both > 1	QD54A is equal to at least one of TOTPEOPA or TOTPEOPB	A: At least one age range variable less than min., or some bad roster ages		SHHSIZE at least as large or larger than screener roster, equal to QD54A	A: Maxima for each age range between given count and min.
		QD54B is equal to at least one of TOTPEOPA or TOTPEOPB	B: At least one age range variable less than min., or some bad roster ages		SHHSIZE at least as large or larger than screener roster, equal to QD54B	B: Maxima for each age range between given count and min.
20 (0)	Not equal, both > 1	Not equal, both >1	TGOODAGA = TGOODAGB, TGOODAGA = QD54A	A: no B: no		QD54A
			TGOODAGA = TGOODAGB, TGOODAGA = QD54B	A: no B: no		QD54B
21 (0)	Not equal, both > 1	Not equal, both >1	TGOODAGA = QD54A TGOODAGB > QD54B	A: no B: no	SHHSIZE > QD54B	QD54A
			TGOODAGB = QD54B TGOODAGA > QD54A	A: no B: no	SHHSIZE > QD54A	QD54B

Table K.1 Priority Conditions Used to Create Household-Consistent Household Size (continued)

Priority Condition (Frequency)	Relationship of QD54A & QD54B	Relationship of TOTPEOPA & TOTPEOPB	Relationships Involving Age Range Variables	Any Roster Missing?	 Screener Roster Characteristics	HHSIZE Equals:
22 (1)	Not equal, both > 1	Not equal, both > 1	TGOODAGA > GOODAGEA, TGOODAGB > GOODAGEB, TGOODAGA = SHHSIZE TGOODAGB = SHHSIZE	A: no B: no	TGOODAGA = SHHSIZE, TGOODAGB = SHHSIZE	SHHSIZE
23 (0)	Not equal, both > 1	Not equal, both > 1	TGOODAGA > GOODAGEA, TGOODAGB > GOODAGEB, TGOODAGA = TGOODAGB	A: no B: no		TGOODAGA
24 (4)	Not equal, both > 1	Not equal, both > 1		A: no B: no	SHHSIZE = sum of maxima for each age group across pair members	SHHSIZE
25 (175)	Not equal, both > 1	Not equal, both > 1		A: no B: no	SHHSIZE ≥ 2 , at least as large or larger than screener roster, closer to one of the QD54's	QD54A if SHHSIZE closer to A, QD54B if closer to B
					SHHSIZE ≥ 2 , at least as large or larger than screener roster, equidistant between the QD54's	QD54 of oldest pair member

Table K.1 Priority Conditions Used to Create Household-Consistent Household Size (continued)

Priority Condition (Frequency)	Relationship of QD54A & QD54B	Relationship of TOTPEOPA & TOTPEOPB	Relationships Involving Age Range Variables	Any Roster Missing?	Screener HH SIZE Characteristics	HH SIZE Equals:
26 (3)	Not equal, both > 1	Not equal, both > 1		A fewer than B	SHHSIZE $\geq$ 2, at least as large or larger than screener roster, closer to QD54A than QD54B	QD54A
				B fewer than A	SHHSIZE $\geq$ 2, at least as large or larger than screener roster, closer to QD54B than QD54A	QD54B
				No condition	SHHSIZE $\geq$ 2, at least as large or larger than screener roster, equidistant between the QD54's	QD54 of oldest pair member
				No condition	SHHSIZE $\geq$ 2, at least as large or larger than screener roster	SHHSIZE
27 (3)			At least 3 of the age range variables are missing		SHHSIZE $\geq$ 2, at least as large or larger than screener roster	SHHSIZE

¹ "Min." refers to the minimum possible within each age range based upon the ages of the two pair members.

K.2 Age Variables

Table K.2 illustrates the hierarchical priority conditions ("priorities") used to create a new household-consistent 12 to 17 age group count. Similar priority conditions are used for the 0 to 11, 12 to 14, 12 to 20, 18 to 25, 26 to 34, 35 to 49, 50 or older, and 15 or older age groups. In this table, blank entries indicate that no priority conditions were required for that set of variables. As with the previous table, a variable followed by "A" indicates that the variable corresponds to the value for pair member "A." A similar comment can be made with regard to "B."

As stated earlier, the variables GOODAGEA and GOODAGEB are the total number of cases in the roster with valid ages, and the variables TGOODAGA and TGOODAGB are also the total number of cases in the roster with valid ages, but if the original adjusted count is less than the minimum required, the original count is replaced by the minimum within the age categories 12 to 17, 18 to 25, 26 to 34, 35 to 49, and 50 or older. As noted in Section 14.3.1, these counts are adjusted so that the roster ages match what was entered in each pair member's questionnaire. Hence, AGE1217A is the adjusted count of 12- to 17-year-olds for pair member A, and AGE1217B is the adjusted count of 12- to 17-year-olds for pair member B. If AGE1217A or AGE1217B is less than the minimum possible, the count is replaced by the minimum, which is given by TAG1217A and TAG1217B, respectively. Otherwise, AGE1217A and TAG1217A are equivalent, as are AGE1217B and TAG1217B. The sum of AGE011A, AGE1217A, AGE1825A, AGE2634A, AGE3549A, and AGE50PA is GOODAGEA. Similarly, the sum of TAG011A, TAG1217A, TAG1825A, TAG2634A, TAG3549A, and TAG50PA is TGOODAGA. The same can be said for GOODAGEB and TGOODAGB.

The final 12 to 17 age count is denoted by AGE1217. The screener age count, denoted by SAGE1217, is used only if the age counts in each pair member's roster cannot conform to the minimum necessary or otherwise are not possible to incorporate. If after all edits the count for AGE1217 is missing, but the counts for other age groups are not missing, and the counts for the 0 to 11 age group are the same for both pair members, then the sum of the counts for the other age groups, plus the minimum possible for AGE1217, are given by EXC1217. If other means fail to determine the appropriate value for the age count, match measures are used. These are measures that summarize the quality of the match between the two pair members. A match label of "0" indicates a perfect match, where the pair member's roster has a household member who is identified as the other pair member with a perfect match on age and gender and is indicated as the other pair member by the MBRSEL variable. There are several levels of match measures where a lower number signifies a better quality match. These measures are explained in detail in Section 14.2.2.1. As a final check, if the age group counts do not equal HHSIZE, and the counts for the pair members are unequal, then the count is set to missing. As with Table K.1, the first column in Table K.2 shows the hierarchical "priority," with the frequency of occurrence for each priority in parentheses, for the AGE1217 count. In most cases, the frequencies corresponding to the other age ranges were the same as the frequency for AGE1217. In those cases where the frequency differed, footnotes provide details of the differences.

Table K.2 Priority Conditions Used to Create Household-Consistent Age Variables (Using AGE1217)

Priority Condition (Frequency)	Relationships Involving TOTPEOP, GOODAGE, and HHSIZE	Relationships Involving AGE1217A, AGE1217B	Relationships Involving Other Age Groups	Relationships Involving Screener Counts	Quality of Roster Measures	AGE1217 Equals:
1 (1) ¹	GOODAGEA = GOODAGEB, GOODAGEA = TOTPEOPA, GOODAGEB = TOTPEOPB GOODAGEB = HHSIZE, all nonmissing, all > 1	AGE1217A < min. (minimum), AGE1217B ≥ min.				AGE1217B
2 (0) ²		AGE1217B < min. AGE1217A ≥ min.				AGE1217A
3 (0)		AGE1217A < min. AGE1217B < min.		SHHSIZE = HHSIZE, SAGE1217 ≥ min.		SAGE1217
4 (7) ³		AGE1217A = AGE1217B, both ≥ min.	Another count except 12-17 < min.			AGE1217A
5 (1) ⁴		AGE1217A not equal to AGE1217B, both ≥ min.	AGE1825A < min., AGE1825B ≥ min.			AGE1217B
6 (0) ⁵			AGE1825B < min., AGE1825A ≥ min.			AGE1217A
7 (0)			Another count except 12-17 < min.		Fewer roster entries missing in A than B	AGE1217A
8 (0)					Fewer roster entries missing in B than A	AGE1217B
9 (0) ⁶					A & B: none missing A has better match measure than B	AGE1217A

Table K.2 Priority Conditions Used to Create Household-Consistent Age Variables (Using AGE1217) (continued)

Priority Condition (Frequency)	Relationships Involving TOTPEOP, GOODAGE, and HHSIZE	Relationships Involving AGE1217A, AGE1217B	Relationships Involving Other Age Groups	Relationships Involving Screener Counts	Quality of Roster Measures	AGE1217 Equals:
10 (0) ⁷	GOODAGEA = GOODAGEB, GOODAGEA = TOTPEOPA, GOODAGEB = TOTPEOPB, GOODAGEB = HHSIZE, all nonmissing, all > 1	AGE1217A not equal to AGE1217B, both $\geq$ min.	Another count except 12-17 < min.		A & B: none missing B has better match measure than A	AGE1217B
11 (0) ⁸					A & B: none missing Age A $\geq$ Age B	AGE1217A
					A & B: none missing Age B > Age A	AGE1217B
12 (0) ⁹						Missing
13 (15,566)		AGE1217A = AGE1217B	All other counts equal across pair members			AGE1217A
14 (618)		At least one age group has an unequal count between pair members		A: all age counts are equal to their screener counterparts	No missing roster entries on either side	AGE1217A
				B: all age counts are equal to their screener counterparts	No missing roster entries on either side	AGE1217B
15-22 (76)					A & B: none missing A has better match measure than B	AGE1217A
					A & B: none missing B has better match measure than A	AGE1217B
23 (88)					A & B: none missing Age A $\geq$ Age B	AGE1217A
					A & B: none missing Age B > Age A	AGE1217B

Table K.2 Priority Conditions Used to Create Household-Consistent Age Variables (Using AGE1217) (continued)

Priority Condition (Frequency)	Relationships Involving TOTPEOP, GOODAGE, and HHSIZE	Relationships Involving AGE1217A, AGE1217B	Relationships Involving Other Age Groups	Relationships Involving Screener Counts	Quality of Roster Measures	AGE1217 Equals:
24 (4)	GOODAGEA = GOODAGEB, GOODAGEA = TOTPEOPA, GOODAGEB = TOTPEOPB, GOODAGEB = HHSIZE, all nonmissing, all > 1	At least one age group has an unequal count between pair members			Fewer roster entries missing in A than B A has good match measure (labels 0-7)	AGE1217A
				Fewer roster entries missing in B than A B has good match measure (labels 0-7)	AGE1217B	
25 (0)					Fewer roster entries missing in A than B A is older than B	AGE1217A
				Fewer roster entries missing in B than A B is older than A	AGE1217B	
26 (0)					Fewer roster entries missing in A than B B is older than A	AGE1217B
				Fewer roster entries missing in B than A A is older than B	AGE1217A	
27 (0)					A & B: same number of roster entries missing (> 0) A is older than B	AGE1217A
28 (0)					A & B: same number of roster entries missing (> 0) B is older than A	AGE1217B

Table K.2 Priority Conditions Used to Create Household-Consistent Age Variables (Using AGE1217) (continued)

Priority Condition (Frequency)	Relationships Involving TOTPEOP, GOODAGE, and HHSIZE	Relationships Involving AGE1217A, AGE1217B	Relationships Involving Other Age Groups	Relationships Involving Screener Counts	Quality of Roster Measures	AGE1217 Equals:
29 (0)	GOODAGEA = TOTPEOPA, GOODAGEB = TOTPEOPB, GOODAGEA = HHSIZE, GOODAGEB not equal to HHSIZE	AGE1217A < min. AGE1217B = min.				AGE1217B
30 (0)		AGE1217B < min. AGE1217A = min.				AGE1217A
31 (0)		AGE1217A < min. AGE1217B < min.		SAGE1217 $\geq$ min.		SAGE1217
32 (0)		AGE1217A = AGE1217B, both $\geq$ min.				AGE1217A
33 (0)		AGE1217A not equal to AGE1217B	AGE1825A < min. AGE1825B $\geq$ min.			AGE1217B
34 (0)			AGE1825B < min. AGE1825A $\geq$ min.			AGE1217A
35 (0)					Fewer roster entries missing in A than B	AGE1217A
36 (0)					Fewer roster entries missing in B than A	AGE1217B
37 (0)					A & B: same number of roster entries missing (> 0) A has good match measure (labels 0-7)	AGE1217A
38 (0)					A & B: same number of roster entries missing (> 0) B has good match measure (labels 0-7)	AGE1217B

Table K.2 Priority Conditions Used to Create Household-Consistent Age Variables (Using AGE1217) (continued)

Priority Condition (Frequency)	Relationships Involving TOTPEOP, GOODAGE, and HHSIZE	Relationships Involving AGE1217A, AGE1217B	Relationships Involving Other Age Groups	Relationships Involving Screener Counts	Quality of Roster Measures	AGE1217 Equals:
39 (0)	GOODAGEA = TOTPEOPA, GOODAGEB = TOTPEOPB, GOODAGEA = HHSIZE, GOODAGEB not equal to HHSIZE	AGE1217A not equal to AGE1217B			A & B: same number of roster entries missing (> 0) A is older than B	AGE1217A
					A & B: same number of roster entries missing (> 0) B is older than A	AGE1217B
40 (0)		Priority conditions 29-39 not met				Missing
41 (0)		AGE1217 missing after priority conditions 29-40 invoked, other age range counts not missing				HHSIZE - sum of other age counts
42 (727)		Priority conditions 29-41 not met				AGE1217A
43 (0)	GOODAGEA = TOTPEOPA, GOODAGEB = TOTPEOPB, GOODAGEB = HHSIZE, GOODAGEA not equal to HHSIZE	AGE1217A < min. AGE1217B = min.				AGE1217B
44 (0)		AGE1217B < min. AGE1217A = min.				AGE1217A
45 (0)		AGE1217A < min. AGE1217B < min.		SAGE1217 ≥ min.		SAGE1217
46 (0)		AGE1217A = AGE1217B, both ≥ min.				AGE1217A
47 (0)		AGE1217A not equal to AGE1217B	AGE1825A < min. AGE1825B ≥ min.			AGE1217B
48 (0)			AGE1825B < min. AGE1825A ≥ min.			AGE1217A

Table K.2 Priority Conditions Used to Create Household-Consistent Age Variables (Using AGE1217) (continued)

Priority Condition (Frequency)	Relationships Involving TOTPEOP, GOODAGE, and HHSIZE	Relationships Involving AGE1217A, AGE1217B	Relationships Involving Other Age Groups	Relationships Involving Screener Counts	Quality of Roster Measures	AGE1217 Equals:
49 (0)	GOODAGEA = TOTPEOPA, GOODAGEB = TOTPEOPB, GOODAGEB = HHSIZE, GOODAGEA not equal to HHSIZE	AGE1217A not equal to AGE1217B			Fewer roster entries missing in A than B	AGE1217A
50 (0)					Fewer roster entries missing in B than A	AGE1217B
51 (0)					A & B: same number of roster entries missing (> 0) B has good match measure (labels 0-7)	AGE1217B
52 (0)					A & B: same number of roster entries missing (> 0) A has good match measure (labels 0-7)	AGE1217A
53 (0)					A & B: same number of roster entries missing (> 0) A is older than B	AGE1217A
					A & B: same number of roster entries missing (> 0) B is older than A	AGE1217B
54 (0)		Priority conditions 43-53 not met				Missing
55 (0)		AGE1217 missing after priority conditions 43-54 invoked, other age range counts not missing				HHSIZE - sum of other age counts
56 (486)		Priority conditions 43-55 not met				AGE1217B

Table K.2 Priority Conditions Used to Create Household-Consistent Age Variables (Using AGE1217) (continued)

Priority Condition (Frequency)	Relationships Involving TOTPEOP, GOODAGE, and HHSIZE	Relationships Involving AGE1217A, AGE1217B	Relationships Involving Other Age Groups	Relationships Involving Screener Counts	Quality of Roster Measures	AGE1217 Equals:
57 (213)	TGOODAGA = HHSIZE					TAG1217A
	TGOODAGB = HHSIZE					TAG1217B
58 (48) ¹⁰	SHHSIZE = HHSIZE	AGE1217A, AGE1217B ≤ SAGE1217		AGE1217A & B ≤ SAGE1217		SAGE1217
59 (1) ¹¹	SHHSIZE = HHSIZE, HHSIZE = EXC1217	AGE1217 missing	Other counts not missing, AGE011A equals AGE011B			MIN1217
60 (7) ¹²	Previous priority conditions for HHSIZE, TOTPEOP, GOODAGE, not met, either the two TOTPEOP's > 0, or SHHSIZE = HHSIZE			AGE1217A equals SAGE1217		AGE1217A
				AGE1217B equals SAGE1217		AGE1217B

Table K.2 Priority Conditions Used to Create Household-Consistent Age Variables (Using AGE1217) (continued)

Priority Condition (Frequency)	Relationships Involving TOTPEOP, GOODAGE, and HHSIZE	Relationships Involving AGE1217A, AGE1217B	Relationships Involving Other Age Groups	Relationships Involving Screener Counts	Quality of Roster Measures	AGE1217 Equals:
61 (0)	Previous priority conditions for HHSIZE, TOTPEOP, GOODAGE, not met, SHHSIZE = HHSIZE	AGE1217 missing	At least 3 of the other counts missing			SAGE1217
99 (0)	All prior conditions were not met.					Missing

¹ The following frequencies were observed for priority condition #1: AGE3549, AGE50p, AGE1214, and AGE15p = 0; AGE1825 = 2; AGE011 = 7.

² The following frequencies were observed for priority condition #2: AGE1825, AGE2634, AGE3549, and AGE1220 = 1; AGE50p = 2.

³ The following frequencies were observed for priority condition #4: AGE011 = 0; AGE1825 and AGE2634 = 4; AGE3549 and AGE1220 = 5; AGE1214 = 9.

⁴ The frequency of priority condition #5 for AGE011, AGE3549, AGE50p, AGE1214, and AGE15p = 0.

⁵ The frequency of priority condition #6 for AGE3549 and AGE15p = 1.

⁶ The following frequencies were observed for priority condition #9: AGE2634 and AGE3549 = 1.

⁷ The frequency of priority condition #10 for AGE15p = 1.

⁸ The frequency of priority condition #11 for AGE1825, AGE2634, AGE3549, and AGE1220 = 1.

⁹ The following frequencies were observed for priority condition #12: AGE011 = 2.

¹⁰ The following frequencies were observed for priority condition #58: AGE011 = 35; AGE2634 = 43; AGE1825, AGE1214, AGE1220, and AGE15p = 46; AGE50p = 49.

¹¹ The following frequencies were observed for priority condition #59: AGE1825, AGE3549, AGE1214, AGE1220, and AGE15p = 0; AGE011=3.

¹² The following frequencies were observed for priority condition #60: AGE011 and AGE15p = 1; AGE2635 and AGE3549 = 4; AGE1825 and AGE50p = 6.

This page intentionally left blank

Appendix L: Creation of Household-Level and Person-Level Files

This page intentionally left blank

Appendix L: Creation of Household-Level and Person-Level Files

L.1 Introduction

For the 2016 administration of the National Survey on Drug Use and Health (NSDUH), a person was randomly selected for an interview through a five-stage sample selection process. States were first stratified into a total of 750 state sampling (SS) regions. Within each of these SS regions, a sample of census tracts was selected (i.e., the first stage of selection) with probabilities proportional to a composite size measure and with minimum replacement. For the second stage of selection, adjacent census block groups within selected census tracts were aggregated to meet the minimum dwelling unit (DU) requirements. One census block group was selected per sampled census tract with probability proportionate to a composite size measure and with minimum replacement. Within each selected census block group, adjacent census blocks were combined to form the third-stage sampling units or area segments. One area segment was selected within each sampled census tract with probability proportional to population size.¹ Once the sample segments were selected, specially trained field staff visited areas and created lists of all eligible DUs within the sample segment boundaries. These lists served as the frames for the fourth stage of sample selection. After the DUs were selected within each segment, an interviewer visited each selected DU to obtain a roster of all people aged 12 or older. This roster information was then used to select zero, one, or two people from the household at the fifth stage of sample selection.

At the end of the survey year, a household-level file and a person-level file were created to record the information obtained from the sampling processes described above. The household-level and person-level files were used in the final creation of the person-level and pair-level analysis weights. In addition, the person-level file was later subset into a smaller data file that contained only respondents who were considered "completed" cases; this file was used for analysis. Refer to Section L.3 for the definition of a completed case.

L.2 Dwelling Unit-Level Eligibility and Completeness Criteria

Before proceeding with the fifth stage of sample selection (i.e., within selected households), a set of rules was used to determine whether a DU was eligible to be selected. Eligibility of the DU was recorded in the binary variable DUELIG, where a value of 1 indicated eligibility. Two examples of DUs deemed to be ineligible included those defined as "vacant" and those determined to be "not a primary residence."

Occasionally, DUs were eligible but failed to complete the screening process. Reasons for not completing the screening process were recorded, including situations such as "language barrier," "refusal," and "denied access." Completeness of the screening process for the DU was recorded in the binary variable DUCOMP, where a value of 1 indicated completeness. For the

¹ Segments consist of clusters of the geographic aggregated adjacent census blocks. SS regions were formed through geographically partitioning each state into roughly equal-sized regions based on a composite size measure. The sample design report in the 2016 NSDUH methodological resource book (Center for Behavioral Health Statistics and Quality, 2017c) contains more information regarding the sample design.

segments where all the DUs were from denied-access areas, such as gated communities, an adjustment was made in the final household-level file. Although the field interviewers could not obtain an accurate count of DUs from denied-access areas, these DUs were considered eligible. Therefore, DU information from the U.S. Census Bureau for these areas was used in the household-level file.

During the third stage of sampling, it was possible to select a sample segment more than once because samples were selected with replacement. These duplicated segments had different segment IDs (SEGIDs) for each duplicate. However, one SEGID contained all the DU information and the other had none. The number of eligible DUs was split as evenly as possible between the two SEGIDs, and this information was updated in the household-level and person-level files.

L.3 Person-Level Eligibility and Completeness Criteria

During screening, respondents were asked to identify all eligible household members so that only eligible individuals were listed and, therefore, potentially selected. Eligibility was determined according to the criteria provided in Section L.1. Eligible respondents at the time of screening were recorded in the binary variable PRELIG, which had a value of 1 if the household member was eligible. Respondents who were selected were recorded in the binary variable PRSEL, where 1 indicated a selected individual. It was possible for an individual to be selected, but at the time of the interview, to be determined ineligible. Examples of changes from eligibility to ineligibility included "the selected person turned out not to be a permanent resident in the DU" and "roster error." If this occurred, the value of PRELIG was changed from 1 to 0.

A summary of the number of selected, eligible, and completed dwelling units is shown in [Table L.1](#). The number of eligible, selected, and interviewed individuals also is summarized in the table.

Table L.1 NSDUH Household, Person Eligibility, and Completed Interview Counts: 2016

	Selected Dwelling Units	Eligible Dwelling Units	Completed Screenings (Dwelling Units)	Eligible People	Selected People	Interviewed People	Completed Cases (Interviews)
CAI	205,589	173,149	135,188	286,377	95,607	68,018	67,942

CAI = computer-assisted interviewing.

To be considered a completed case for purposes of analysis, a respondent had to provide "yes" or "no" answers to the cigarette usage gate question and to at least 9 of the following additional drug usage gate questions: (1) chewing tobacco, (2) snuff, (3) cigars, (4) alcohol, (5) marijuana, (6) cocaine (in any form), (7) heroin, (8) hallucinogens, (9) inhalants, (10) pain relievers, (11) tranquilizers, (12) stimulants, and (13) sedatives.² Unlike the paper-and-pencil interviewing questionnaire in 1999 and surveys prior to 1999, no logical inference could be made from information within a section if the gate question was not answered. This was because the computer-assisted interviewing instrument routed respondents out of a section if the gate question was not answered. Completeness of the survey for eligible individuals was recorded in

² For more details on the "usable case rule," see Section 2.2.3.

the binary variable PRCOMP, which had a value of 1 if the respondent was a completed case, and 0 if not. For a summary of the number of completed cases in the 2016 survey, see [Table L.1](#).

L.4 Variables in the Household-Level and Person-Level Files

This section documents some of the important person-level variables that were created for the household-level and person-level files.

Screeners-level demographic variables were created from the screener roster information in the household-level and person-level files. XAGE was the screener age, which either could be "continuous" (single-year ages) or categorical. A respondent could choose to give an age category instead of the actual age. The age categories with their accompanying codes were 199 = 12 to 17 years old; 299 = 18 to 25 years old; 399 = 26 to 34 years old; 499 = 35 to 49 years old; and 599 = 50 years old or older. Screener race (XRACE1-XRACE6), screener Hispanicity (XHISP), and screener gender (XSEX) also were produced from the screener roster information. XRACE1 through XRACE6 were indicator variables representing white, black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, and other, respectively. The household-level variable PAIRSEL represented the number of people within each age group selected from a DU. It was a 20-level variable indicating whether zero, one, or two individuals were selected from the five age groups (12 to 17, 18 to 25, 26 to 34, 35 to 49, and 50 or older) in a given household. (If two people were selected from the household, this variable indicated the age groups of both pair members.) Similar to PAIRSEL, the household-level variable PAIRRESP had 20 levels, which indicated whether zero, one, or two people completed the interviews from the five age groups within a household.

As described in the sample design report in the 2016 NSDUH methodological resource book (Center for Behavioral Health Statistics and Quality, 2017c), states were partitioned into SS regions, which were further partitioned into clusters of adjacent blocks called "segments." The variable SEGID (segment ID number) was a two-letter state abbreviation followed by a two-digit SS region and a two-letter segment identifier, with the fifth character indicating the survey year and the sixth character being the initial quarter assignment of the segment. Census region (REGION) was a four-level geographic variable recoded from the respondent's state of residence. The four levels were Northeast, Midwest, South, and West. The population density variables PDEN10 and PDEN210³ classified respondents according to their living situation, whether it was in a rural or urban area, and, if urban, the size of the urban area. They were used to categorize segments where the respondents lived based on 2010 census data and the December 2009 Core-Based Statistical Area (CBSA) classifications.⁴ PDEN210 had five levels: segment in CBSA with 1 million or more people; segment in CBSA with 250,000 to 999,999 people; segment in CBSA with fewer than 250,000 people; segment in urban area but not in CBSA; and segment in rural area (not in CBSA and not in urban area). PDEN10 had three levels: segment in CBSA with 1 million or more people; segment in CBSA with fewer than 1 million people; and segment not in CBSA.

³ PDEN10 and PDEN210 were similar to PDEN/PDEN2 and PDEN00/PDEN200 in previous sample designs. The suffix "10" indicates that the variables were derived from the 2010 census.

⁴ CBSAs include metropolitan and micropolitan statistical areas as defined by the Office of Management and Budget (2009).

The variable STATE represented the Federal Information Processing Standards (FIPS) state codes for the 50 states and the District of Columbia. The variable STATE was created at the sampling stage and did not contain any missing values. The variables VESTR and VEREP were created to capture the sampling design structure. Each SS region appeared in a different variance estimation stratum (VESTR) every quarter. Two replicates (VEREP) were defined within each variance stratum. Each replicate consisted of four segments, one for each quarter of data collection. The segment-level variable RURURB10⁵ is derived from the 2010 census block-level designation of rural or urban.⁶ In the NSDUH sample, if 100 percent of the blocks are rural, the segment is defined as rural (RURURB10 = 1). If one or more of the blocks within a segment is urban, the segment is defined as urban (RURURB10 = 2). Other sampling variables such as DIVISION, SSREGION, GQTYPE, ID, STNAME, STUSAB, and QUARTER⁷ also were included in the household-level and person-level files.

⁵ RURURB10 was similar to RURORURB and RURURB00 in previous sample designs. The suffix "10" indicates that this variable was derived from the 2010 census.

⁶ The census classifies as urban all blocks located within urbanized areas (UA) and urban clusters (UC). UAs and UCs generally consist of core census block groups or blocks that have a population density of at least 1,000 people per square mile and surrounding census blocks that have an overall density of at least 500 people per square mile. In addition, under certain conditions, less densely settled territory may be part of each UA or UC.

⁷ For more details on these sampling variables, refer to the sample design report in the 2016 NSDUH methodological resource book (Center for Behavioral Health Statistics and Quality, 2017c).