2017 NATIONAL SURVEY ON DRUG USE AND HEALTH METHODOLOGICAL RESOURCE BOOK

SECTION 12: QUESTIONNAIRE DWELLING UNIT-LEVEL AND PERSON PAIR-LEVEL SAMPLING WEIGHT CALIBRATION

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

March 2019

This page intentionally left blank

2017 NATIONAL SURVEY ON DRUG USE AND HEALTH METHODOLOGICAL RESOURCE BOOK

SECTION 12: QUESTIONNAIRE DWELLING UNIT-LEVEL AND PERSON PAIR-LEVEL SAMPLING WEIGHT CALIBRATION

For questions about this report, available at https://www.samhsa.gov/data/, 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

March 2019

Recommended Citation: Center for Behavioral Health Statistics and Quality. (2019). *2017 National Survey on Drug Use and Health Methodological Resource Book, Section 12: Questionnaire Dwelling Unit-Level and Person Pair-Level Sampling Weight Calibration*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

This page intentionally left blank

Table of Contents

Chapter **Page**

List of Terms and Abbreviations ix
Overview 1
1. Introduction 3
2. Questionnaire Dwelling Unit and Pair Selection Probabilities 9
2.1 Pair Selection Probability 11
2.1.1 Case I: DUs with $S \geq 2$ 11
2.1.2 Case II: DUs with $S < 2$ 11
2.2 Questionnaire Dwelling Unit Selection Probability 13
3. Brief Description of the Generalized Exponential Model 15
4. Predictor Variables for the Questionnaire Dwelling Unit and Pair Weight Calibration via the Generalized Exponential Model 17
4.1 Questionnaire Dwelling Unit Weight Calibration 17
4.2 Pair Weight Calibration 18
5. Definition of Extreme Weights 23
5.1 Questionnaire Dwelling Unit Extreme Weight Definition 23
5.2 Person Pair Extreme Weight Definition 24
6. Weight Calibration at Questionnaire Dwelling Unit and Pair Levels 27
6.1 Phase I SDU-Level Weight Components 32
6.2 QDU Weight Components 33
6.2.1 QDU Weight Component #12: Inverse of Selection Probability of at Least One Person in the Dwelling Unit 33
6.2.2 QDU Weight Component #13: Selected QDU Poststratification to SDU-Based Control Totals 33
6.2.3 QDU Weight Component #14: Respondent QDU Nonresponse Adjustment 33
6.2.4 QDU Weight Component #15: Respondent QDU Poststratification to SDU-Based Control Totals 34
6.2.5 QDU Weight Component #16: Respondent QDU Extreme Value Adjustment 34
6.3 Pair-Level Weight Components 34
6.3.1 Pair Weight Component #12: Inverse of Selection Probability of a Person Pair in the Dwelling Unit 34
6.3.2 Pair Weight Component #13: Selected Pair Poststratification to SDU-Based Control Totals 34
6.3.3 Pair Weight Component #14: Respondent Pair Nonresponse Adjustment 35
6.3.4 Pair Weight Component #15: Respondent Pair Poststratification to SDU-Based Control Totals 35

Table of Contents (continued)

Chapter **Page**

6.3.5 Pair Weight Component #16: Respondent Pair Extreme Weight Adjustment ... 35
7. Evaluation of Calibration Weights ... 37
7.1 Response Rates ... 37
7.2 Proportions of Extreme Values and Outwinsors ... 37
7.3 Slippage Rates ... 38
7.4 Weight Adjustment Summary Statistics ... 38
7.5 Sensitivity Analysis of Drug Use Estimates ... 39
References ... 51
List of Contributors ... 53

Appendix **Page**

A Technical Details about the Generalized Exponential Model ... A-1
B Derivation of Poststratification Control Totals ... B-1
C GEM Modeling Summary for the Questionnaire Dwelling Unit Weights ... C-1
C.1 Model Group 1: Northeast ... C-9
C.2 Model Group 2: Midwest ... C-17
C.3 Model Group 3: South ... C-25
C.4 Model Group 4: West ... C-33
D Evaluation of Calibration Weights: Questionnaire Dwelling Unit-Level Response Rates ... D-1
E Evaluation of Calibration Weights: Questionnaire Dwelling Unit-Level Proportions of Extreme Values and Outwinsors ... E-1
F Evaluation of Calibration Weights: Questionnaire Dwelling Unit-Level Slippage Rates ... F-1
G Evaluation of Calibration Weights: Questionnaire Dwelling Unit-Level Weight Summary Statistics ... G-1
H GEM Modeling Summary for the Pair Weights ... H-1
H.1 Model Group 1: Northeast and South ... H-5
H.2 Model Group 2: Midwest and West ... H-15
I Evaluation of Calibration Weights: Pair-Level Response Rates ... I-1
J Evaluation of Calibration Weights: Pair-Level Proportions of Extreme Values and Outwinsors ... J-1
K Evaluation of Calibration Weights: Pair-Level Slippage Rates ... K-1
L Evaluation of Calibration Weights: Pair-Level Weight Summary Statistics ... L-1

List of Tables

Table **Page**

1.1 2013–2017 NSDUH Sample Sizes 5
1.2 Pair Domains 6
2.1 Building Blocks of the QDU and Person Pair Samples: Dwelling Units and People in the 2013–2017 NSDUHs 10
6.1 Sample Size, by Model Group at QDU and Pair Levels 32
7.1 Estimates of Totals and SEs for Domains of Interest Based on QDU Sample: 2017 40
7.2a Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Alcohol and Tobacco among Mother-Child (12 to 17) Pairs, by Mother Use: 2017 41
7.2b Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Alcohol and Tobacco among Father-Child (12 to 17) Pairs, by Father Use: 2017 42
7.3a Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Any Illicit Drug or Marijuana among Mother-Child (12 to 17) Pairs, by Mother Use: 2017 43
7.3b Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Any Illicit Drug or Marijuana among Father-Child (12 to 17) Pairs, by Father Use: 2017 44
7.4 Percentages of Youths (12 to 17) Living with a Parent Reporting Lifetime, Past Year, and Past Month Use of Alcohol and Tobacco among Parent-Child (12 to 17) Pairs, Asked Whether Their Parents Had Spoken to Them about the Dangers of Tobacco, Alcohol, or Drug Use within the Past 12 Months: 2017 45
7.5 Percentages of Youths (12 to 17) Living with a Parent Reporting Lifetime, Past Year, and Past Month Use of Any Illicit Drug and Marijuana among Parent-Child (12 to 17) Pairs, Asked Whether Their Parents Had Spoken to Them about the Dangers of Tobacco, Alcohol, or Drug Use within the Past 12 Months: 2017 46
7.6a Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Alcohol and Tobacco among Mother-Child (12 to 17) Pairs, for Mother in the Pair, Asked Whether She Had Spoken to Her Children about the Dangers of Tobacco, Alcohol, or Drug Use within the Past 12 Months: 2017 47
7.6b Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Alcohol and Tobacco among Father-Child (12 to 17) Pairs, for Father in the Pair, Asked Whether He Had Spoken to His Child about the Dangers of Tobacco, Alcohol, or Drug Use within the Past 12 Months: 2017 48
7.7a Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Any Illicit Drug and Marijuana among Mother-Child (12 to 17) Pairs, for Mother in the Pair, Asked Whether She Had Spoken to Her Child about the Dangers of Tobacco, Alcohol, or Drug Use within the Past 12 Months: 2017 49
7.7b Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Any Illicit Drug and Marijuana among Father-Child (12 to 17) Pairs, for Father in the Pair, Asked Whether He Had Spoken to His Child about the Dangers of Tobacco, Alcohol, or Drug Use within the Past 12 Months: 2017 50

List of Tables (continued)

Table **Page**

C.1a 2017 QDU Weight GEM Modeling Summary (Model Group 1: Northeast) C-11
C.1b 2017 Distribution of Weight Adjustment Factors and Weight Products (Model Group 1: Northeast) C-12
C.2a 2017 QDU Weight GEM Modeling Summary (Model Group 2: Midwest)..... C-19
C.2b 2017 Distribution of Weight Adjustment Factors and Weight Products (Model Group 2: Midwest)..... C-20
C.3a 2017 QDU Weight GEM Modeling Summary (Model Group 3: South) C-27
C.3b 2017 Distribution of Weight Adjustment Factors and Weight Products (Model Group 3: South)..... C-28
C.4a 2017 QDU Weight GEM Modeling Summary (Model Group 4: West) C-35
C.4b 2017 Distribution of Weight Adjustment Factors and Weight Products (Model Group 4: West)..... C-36
D.1 2017 NSDUH QDU-Level Response Rates D-1
E.1 2017 NSDUH Selected QDU-Level Proportions of Extreme Values and Outwinsors E-1
E.2 2017 NSDUH Respondent QDU-Level Proportions of Extreme Values and Outwinsors E-3
F.1 2017 NSDUH QDU-Level Slippage Rates..... F-1
G.1 2017 NSDUH Selected QDU-Level Weight Summary Statistics G-1
G.2 2017 NSDUH Respondent QDU-Level Weight Summary Statistics..... G-3
H.1a 2017 Pair Weight GEM Modeling Summary (Model Group 1: Northeast and South)..... H-7
H.1b 2017 Distribution of Weight Adjustment Factors and Weight Products (Model Group 1: Northeast and South) H-8
H.2a 2017 Pair Weight GEM Modeling Summary (Model Group 2: Midwest and West)... H-17
H.2b 2017 Distribution of Weight Adjustment Factors and Weight Products (Model Group 2: Midwest and West)..... H-18
I.1 2017 NSDUH Person Pair-Level Response Rates..... I-1
J.1 2017 NSDUH Selected Pair-Level Proportions of Extreme Values and Outwinsors J-1
J.2 2017 NSDUH Respondent Pair-Level Proportions of Extreme Values and Outwinsors J-3
J.3 2017 NSDUH Respondent Pair-Level Proportions of Extreme Values and Outwinsors J-5
K.1 2017 NSDUH Respondent Pair-Level Slippage Rates K-1
L.1 2017 NSDUH Selected Pair-Level Weight Summary Statistics..... L-1
L.2 2017 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.nr) L-4
L.3 2017 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.ps and res.pr.ev) L-6

List of Exhibits

Exhibit **Page**

1.1 QDU and Pair Sampling Weight Calibration Steps 7
4.1 Definitions of Levels for QDU-Level Calibration Modeling Variables 19
4.2 Definitions of Levels for Pair-Level Calibration Modeling Variables 20
6.1 Summary of 2017 NSDUH QDU Sample Weight Components 29
6.2 Summary of 2017 NSDUH Person Pair Sample Weight Components 30
6.3 U.S. Census Bureau Regions/Model Groups 31
C.1 Definitions of Levels for QDU-Level Calibration Modeling Variables C-2
C.2 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights C-8
C.1.1 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (sel.qdu.ps) Model Group 1: Northeast C-14
C.1.2 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 1: Northeast C-15
C.1.3 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 1: Northeast C-16
C.2.1 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (sel.qdu.ps) Model Group 2: Midwest C-22
C.2.2 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 2: Midwest C-23
C.2.3 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 2: Midwest C-24
C.3.1 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (sel.qdu.ps) Model Group 3: South C-30
C.3.2 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 3: South C-31
C.3.3 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 3: South C-32
C.4.1 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (sel.qdu.ps) Model Group 4: West C-38
C.4.2 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 4: West C-39
C.4.3 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 4: West C-40
H.1 Definitions of Levels for Pair-Level Calibration Modeling Variables H-2
H.2 Covariates for 2017 NSDUH Pair Weights H-4
H.1.1 Covariates for 2017 NSDUH Pair Weights (sel.pr.ps) Model Group 1: Northeast and South H-10
H.1.2 Covariates for 2017 NSDUH Pair Weights (res.pr.nr) Model Group 1: Northeast and South H-11
H.1.3 Covariates for 2017 NSDUH Pair Weights (res.pr.ps) Model Group 1: Northeast and South H-12
H.1.4 Covariates for 2017 NSDUH Pair Weights (res.pr.ev) Model Group 1: Northeast and South H-13

List of Exhibits (continued)

Exhibit **Page**

H.2.1 Covariates for 2017 NSDUH Pair Weights (sel.pr.ps) Model Group 2: Midwest and West..... H-20
H.2.2 Covariates for 2017 NSDUH Pair Weights (res.pr.nr) Model Group 2: Midwest and West..... H-21
H.2.3 Covariates for 2017 NSDUH Pair Weights (res.pr.ps) Model Group 2: Midwest and West..... H-22
H.2.4 Covariates for 2017 NSDUH Pair Weights (res.pr.ev) Model Group 2: Midwest and West..... H-23

List of Terms and Abbreviations

DU	Dwelling unit.
ev	Extreme value. See Sections 5.1 and 5.2 for more detail.
GEM	Generalized exponential model. See Chapter 3 for more detail.
Household-level person count	The number of pairs associated with a given domain in a given household. These counts are used as control totals in the poststratification step. See Chapter 11 in the editing and imputation report (Center for Behavioral Health Statistics and Quality, 2019a) for details on how these counts are created, and Chapter 4 for details on their use in poststratification.
IQR	Interquartile range.
Multiplicity factor	The number of pairs associated with a given respondent in a given domain. See Center for Behavioral Health Statistics and Quality (2019a) for more detail.
nr	Nonresponse.
Outwinsor	The proportion of weights trimmed after extreme value adjustment via winsorization.
Pair domain	A pair relationship where the target population is defined by one of the pair members, conditional on the attributes of the other pair member.
Pair relationship	The relationship between selected pair members.
Parent-child	A pair relationship where either both pair members identify the other as part of a parent-child relationship, or both pair members otherwise are determined to form a parent-child pair (either through other evidence or through imputation).
ps	Poststratification.
QDU	Questionnaire dwelling unit: a household where at least one member responded to the questionnaire.
res.pr.nr	Respondent pair nonresponse adjustment step. See Section 6.3.3 for more detail.
res.qdu.nr	Respondent questionnaire dwelling unit nonresponse adjustment step. See Section 6.2.3 for more detail.
res.pr.ev	Respondent pair extreme value adjustment step. See Section 6.3.5 for more detail.
res.qdu.ev	Respondent questionnaire dwelling unit extreme value adjustment step. See Section 6.2.5 for more detail.

res.pr.ps	Respondent pair poststratification adjustment step. See Section 6.3.4 for more detail.
res.qdu.ps	Respondent questionnaire dwelling unit poststratification adjustment step. See Section 6.2.4 for more detail.
SDU	Screener dwelling unit: a household where screener information is available.
sel.pr.ps	Selected person pair poststratification adjustment step. See Section 6.3.2 for more detail.
sel.qdu.ps	Selected questionnaire dwelling unit poststratification adjustment step. See Section 6.2.2 for more detail.
Sibling-sibling	A pair relationship where the pair members are siblings (either reported to be so, or otherwise determined to be so).
Spouse-spouse	A pair relationship where the pair members are either married or living together as though married (either reported to be so, or otherwise determined to be so).
SS	State sampling.
UWE	Unequal weighting effect. It refers to the contribution in the design effect due to unequal selection probability and is defined as $1 + \left[(n-1)/n\right] * CV^2$, where CV = coefficient of variation of weights and n is the sample size.
Winsorization	A method of extreme value adjustment that replaces extreme values with the critical values used for defining low and high extreme values.

Overview

This report documents the method of weight calibration used for producing the final set of questionnaire dwelling unit and pair weights for the National Survey on Drug Use and Health (NSDUH) data from 2017. The weighting team faced several challenges in this task and was able to address them by resorting to innovative modifications of certain basic statistical ideas, which are listed below.

- Under Brewer's method, high weights may occur because of small pair selection probabilities. In any calibration exercise, some treatment of extreme value (ev) in weights is needed, but there is a danger of introducing too much bias by over-treatment. In the generalized exponential model (GEM), which is described in detail in the NSDUH Methodological Resource Book person-level sampling weight calibration report (Center for Behavioral Health Statistics and Quality, 2019b), ev control is built in, but one needs to define suitable ev domains so that not too many evs are defined. If too many design variables are used to define ev domains, then each domain will be very sparse and will not be of much use in defining thresholds for ev. As in past surveys, a hierarchy of domains was defined using pair age (each pair member being in one of the three categories: 12 to 25, 26 to 49, and 50+) and number of people aged 12 to 25 in the household, state, and clusters of states (see Section 5.2 for details).
- Control of evs in weights helps reduce instability of estimates to some extent, but there is a need for methods that do not introduce much bias. Following the famous suggestion of Hajek (1971) in his comments on Basu's fabled example of circus elephants, we performed ratio adjustment (a form of poststratification) to estimated totals obtained from the household data on the number of people belonging to the pair domain of interest. This was implemented in a multivariate manner to get one set of final weights.
- In the absence of a suitable source of poststratification controls for the person pair-level weights and the household-level weights, the inherent two-phase nature of the survey design was capitalized upon to estimate these controls from the first phase of the large screener sample. The first-phase sample weight was poststratified to person-level U.S. Census Bureau counts to get more efficient estimated counts for pair and household data.
- The problem of multiplicities complicated the issue of providing one set of final weights. When dealing with person-level parameters involving drug-related behaviors among members of the same household, it is possible for an individual to manifest himself or herself in the pair sample through different pairs. To avoid overcounting, the pair weights have to be divided by multiplicity factors, which tend to be domain specific. For this reason, multiplicity factors for a key set of pair analysis domains also are produced along with a set of final calibrated pair weights.
- Missing items in the respondent questionnaire led to imputation for deriving pair relationships, multiplicity factors, and household counts for Hajek adjustments.

The calibration task described in this report has been in place, with minor modifications, since the 1999 version of NSDUH, which was then called the National Household Survey on Drug Abuse (NHSDA).[1] Results from this calibration applied to an earlier survey year were presented at the 2001 Joint Statistical Meetings. The procedures described in the proceedings papers from these presentations can serve as useful supplemental reference material on estimation in the presence of multiplicities and extreme weights (Chromy & Singh, 2001) and on GEM calibration of pair weights (Penne, Chen, & Singh, 2001). The experience of using GEM with person weights is described in an earlier proceedings paper (Chen, Penne, & Singh, 2000).

[1] The National Household Survey on Drug Abuse (NHSDA) was renamed the National Survey on Drug Use and Health (NSDUH) in the 2002 survey year.

1. Introduction

Traditionally, most household surveys have been designed either to measure characteristics of the entire household or to focus on a randomly selected respondent from among those determined to be eligible for the survey. Selecting more than one person from the same household is generally avoided because people from the same household often exhibit the same or similar characteristics and behavioral patterns. The intra-class correlation found among members of the same household leads to a clustering effect on the variance of estimates resulting in less precise estimates compared with estimates of the same sample size from a simple random sample. Selecting only one person per household avoids this clustering effect on the variance. The "one person per household" sampling approach, however, precludes the opportunity to gather information about the relationships among household members. In the National Survey on Drug Use and Health (NSDUH),[2] we allow for a richer analytic capability of a survey designed to ensure a positive pairwise probability of selection among all eligible household members in each sample household. Achieving positive probabilities for all pairs within sampled households permits unbiased estimation of the within-dwelling-unit component of variance. Besides providing efficient data collection, this sampling method also facilitates the study of the relationships of social behaviors among members of the same household. This report documents the methodology and development of calibrated weights for the second objective, the study of behavioral relationships among people residing in the same household. The report also describes the development of questionnaire dwelling unit (QDU) weights, which are of independent interest for studying household-level characteristics and also are needed for producing household count estimates of the number of people belonging to pair relationship domains for use as poststratification controls for pair weights.

NSDUH allows for estimating characteristics at the person level, pair level, and household or QDU level. This report describes the weight calibration methods used for the pair- and QDU-level respondents. As described in the person-level report, NSDUH is an annual survey of about 67,500 people selected from the civilian, noninstitutionalized population aged 12 or older from all 50 states and the District of Columbia. A coordinated sample design was developed for the 2014 through 2017 NSDUHs. The coordinated design facilitated 50 percent overlap in third-stage units (area segments) within each successive 2-year period from 2014 through 2017. This designed sample overlap slightly increased the precision of estimates of year-to-year trends because of the expected small but positive correlation resulting from the overlapping sampled area segments between successive survey years. The 50 percent overlap of segments significantly reduced segment listing costs because only one-half of the segments needed to be listed for the 2015 through 2017 surveys.

Another modification from the 2005–2013 NSDUH is a change in the sampling strategy of using 8 "large" states to obtain 3,600 respondents and 43 "small" states (including the District of Columbia) to obtain 900 respondents. The 2014–2017 survey's sample was designed to yield

[2] This report presents information from the 2017 National Survey on Drug Use and Health (NSDUH). Prior to 2002, the survey was called the National Household Survey on Drug Abuse (NHSDA).

- 4,560 completed interviews in California;
- 3,300 completed interviews each in Florida, New York, and Texas;
- 2,400 completed interviews each in Illinois, Michigan, Ohio, and Pennsylvania;
- 1,500 completed interviews each in Georgia, New Jersey, North Carolina, and Virginia;
- 967 completed interviews in Hawaii; and
- 960 completed interviews in each of the remaining 37 states and the District of Columbia.

Under a stratified design with states serving as the primary strata and state sampling (SS) regions serving as the secondary strata, census tracts, census block groups, segments within census block groups, and dwelling units (DUs) within segments were each selected using probability proportional to size sampling. Also in the 2014–2017 design, was the incorporation of census block groups at the second stage of selection to potentially reduce sampling variance and facilitate moving to an address-based sampling design in the future, if desired. NSDUH is sometimes referred to as a two-phase sample where the first phase consisted of a large number of screener dwelling units (SDUs, about 200,000) selected to ensure that various age groups (five in all: 12 to 17, 18 to 25, 26 to 34, 35 to 49, and 50+) of eligible individuals were represented adequately in the second phase. In the 2014–2017 NSDUH design, added focus (greater sample) was placed on the 26 or older age group to improve estimates of drug use and related health measures for this population. Unlike the 2005–2013 NSDUHs, which allocated state sample equally across the age categories 12 to 17, 18 to 25 and 26 or older, in the 2014–2017 design, the sample was allocated with 25 percent for 12 to 17, 25 percent for 18 to 25, 15 percent for 26 to 34, 20 percent for 35 to 49, and 15 percent for 50 or older. Information collected from SDUs also provided estimates of population controls (as in two-phase sampling) for calibration at levels (such as pair and QDU) for which suitable U.S. Census Bureau–based controls were not available. The second phase consisted of the selection of zero, one, or two people from each selected SDU using a modification of Brewer's method such that prescribed sampling rates for the five age groups in each state were achieved with high selection rates for youths (12 to 17) and young adults (18 to 25). Table 1.1 shows the eligible number of selected and responding SDUs, QDUs, pairs, and people for each of the 5 years (2013–2017). The distribution of pair data for different pairs of age groups may vary considerably (see Chapter 2 for details). It is seen that for certain age group domains, the realized sample size may not be sufficient to yield reliable estimates. Also, there may be problems of extreme weights due to small pair selection probabilities under Brewer's method that may cause instability of estimates. These and some other estimation issues related to pair data are discussed below, along with some adopted solutions.

Table 1.1 2013–2017 NSDUH Sample Sizes

Sample Unit		2013	2014	2015	2016	2017
SDU	Selected	190,067	154,533	165,328	173,149	184,266
	Completed	160,325	127,605	132,210	135,188	138,061
QDU	Selected	61,634	64,796	66,721	67,574	68,889
	Completed	48,896	49,672	50,119	50,095	50,328
Pair	Selected	27,108	26,844	27,778	28,033	28,778
	Completed	18,942	18,229	17,954	17,847	17,704
Person	Selected	88,742	91,640	94,499	95,607	97,667
	Completed	67,838	67,901	68,073	67,942	68,032

First, note that for studying drug-related behavioral relationships among members of the same household, pair data are required because the outcome variable generally is defined with respect to the specific other member selected from the household. However, the parameter of interest is generally at the person level and is not at the pair level. For example, in the parent-child pairs, one may be interested in the proportion of children who have used drugs in the past year who have parents who report talking to their child about drugs. Here the target population consists only of children, and not all possible pairs. Note that the pair-level (two people per QDU) sample forms a subsample of the larger person-level (one or two people per QDU) sample, with the QDUs themselves selected from the larger sample of SDUs. NSDUH has features of a two-phase design, which turns out to be useful for estimating calibration controls for poststratification of household-level weights and person pair-level weights. No other outside source is available for obtaining these controls. For this purpose, the screener-level household weights are poststratified to person-level census counts to obtain more efficient estimated controls for pair and household data.

In estimation for pair domains, two major problems arise: one is that of multiplicities because, for a given domain defined by the pair relationship, when the parameter of interest is at the person level, several pairs in the household could be associated with the same person. For example, analysts are interested in an outcome at the person level, the proportion of children who use drugs and whose parents report talking to them about drugs, where the focus is on the child in a parent-child pair. Several parent-child pairs in the household could be associated with the same child. If the household has two parents, the selected child has two inclusion possibilities (one with each parent) in the set of all such parent-child pairs (Center for Behavioral Health Statistics and Quality [CBHSQ], 2019a). The other problem is that of extreme weights that may arise due to small selection probabilities for certain pair age groups, which may lead to unstable estimates. Each of these issues is discussed in turn.

If several pairs in the household are associated with the same person, it is necessary to use the average measure of behavior relationships for each member, which gives rise to multiplicities. Thus, the pair weights need to be divided by the person-level multiplicity factors for each domain of interest, and, therefore, multiplicity factors need to be produced along with the final set of calibrated weights. Because it is not straightforward to create these multiplicities, analyses would have to be necessarily limited to pair relationships where the multiplicities were produced a priori. It was anticipated that analyses of interest would be limited to 14 pair domains, listed in Table 1.2. Because no multiplicity was necessary for the spouse-

spouse/partner-partner pair relationships (by definition, each pair member could have only one partner or one spouse), multiplicity factors were produced for only 12 of these domains. Note that a single pair relationship might have two domains associated with it, because the parameter of interest might be associated with only one member of the pair (the "focus" member), and the multiplicity would differ depending upon which pair member was the focus member.

Table 1.2 Pair Domains

Pair Relationship	Focus
Parent-child: parent, child aged 12–14	Parent
Parent-child: parent, child aged 12–14	Child
Parent-child: parent, child aged 12–17	Parent
Parent-child: parent, child aged 12–17	Child
Parent-child: parent, child aged 12–20	Parent
Parent-child: parent, child aged 12–20	Child
Parent-child: parent, child aged 15–17	Parent
Parent-child: parent, child aged 15–17	Child
Sibling-sibling: older sibling 15–17, younger sibling 12–14	Older sibling
Sibling-sibling: older sibling 15–17, younger sibling 12–14	Younger sibling
Sibling-sibling: older sibling 18–25, younger sibling 12–17	Older sibling
Sibling-sibling: older sibling 18–25, younger sibling 12–17	Younger sibling
Spouse-spouse or partner-partner, with or without children	No multiplicity necessary
Spouse-spouse or partner-partner, with children aged 0–17	No multiplicity necessary

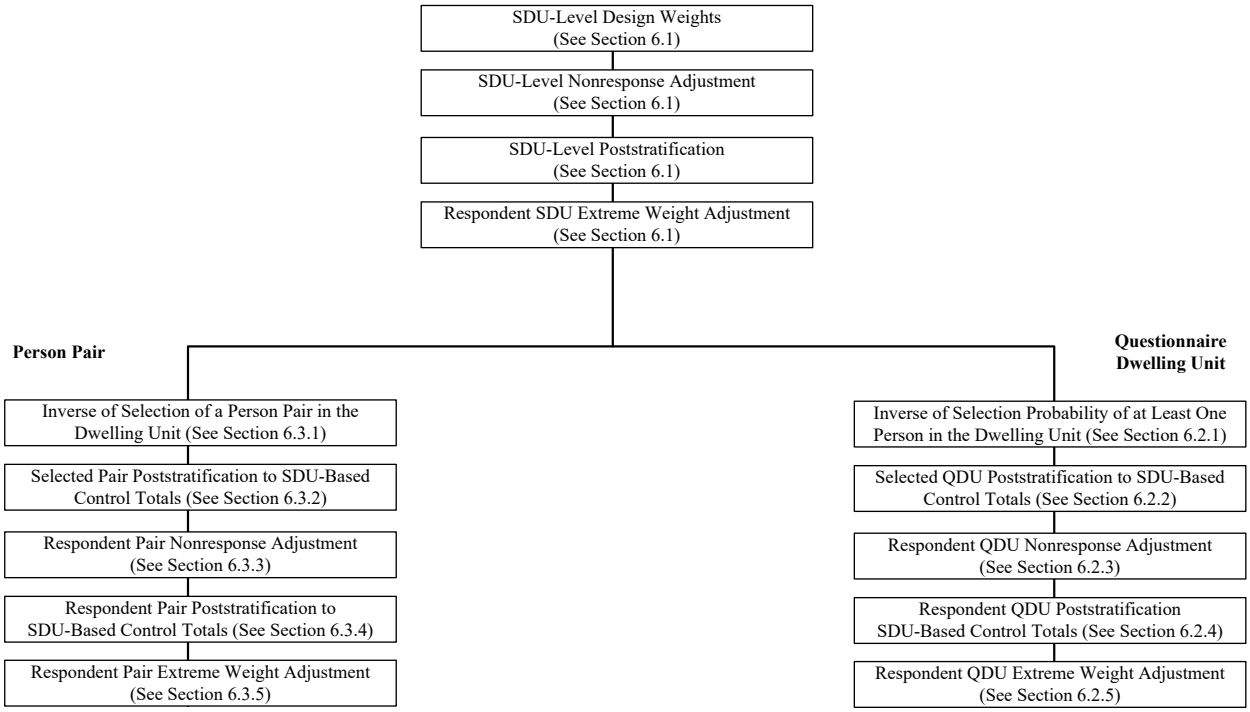
Some of the multiplicities, including counts of all possible pairs in a household for a given domain, were used for poststratification. Details are provided in Chapter 4. Additional information on the imputation of pair relationships, multiplicity factors, and household-level person counts for poststratification can be found in the NSDUH Methodological Resource Book editing and imputation report (CBHSQ, 2019a). Special consideration is required for analysis of pair-level data, and details can be found in *How To Prepare and Analyze Pair Data in the National Survey on Drug Use and Health* (CBHSQ, 2017).

A resolution to the extreme weight problem is to use a Hajek-type modification (Hajek, 1971). This modification essentially entails calibration (like poststratification) to controls for the number of people in households belonging to each domain of interest. These controls can be obtained from the larger sample of singles and pairs (i.e., one or two people selected from DUs). Note, however, that the multiplicity factor, being domain specific, renders the calibration adjustment factor domain specific. This raises the question of finding one set of calibration weights for use with all domains or outcome variables. To get around this problem, a multivariate calibration with respect to a key set of pair domains was performed. This type of poststratification then was followed by a repeat poststratification to further control the extreme weights by imposing separate bound restrictions on the initially identified extreme weights.

The generalized exponential model (GEM) method (Folsom & Singh, 2000) was used for calibration of both QDU- and pair-level design weights through several steps of adjustment as shown in Exhibit 1.1. In GEM, treatment of extreme value (ev) weights is built in via the definition of lower and upper bounds for the extreme weights. For pair data, there was a problem

defining suitable domains for defining extreme weights, as explained in the following paragraphs.

Exhibit 1.1 QDU and Pair Sampling Weight Calibration Steps



In dealing with extreme weights, it is assumed that they arise due to design (due to an imperfect frame, assignment of very small selection probabilities to some units, or a big weight adjustment factor after calibration) so that they make the sample representative of the population and, hence, do not introduce bias. The only problem is that they may lead to highly unstable estimates similar to the problem of Basu's circus elephants[3] (Hajek, 1971). So, we need to perform some treatment (such as winsorization[4]) within suitably defined extreme weight domains such that these domains contain units possibly from different strata but with similar sample selection probabilities to avoid the occurrence of extreme weights due to a mix of

[3] A circus owner had 50 elephants, and wanted to estimate the total weight to help him make arrangements for shipping. To save time, he only wanted to weigh Sambo (an average sized elephant), and use 50 times its weight as an estimate. However, the circus statistician, being highly conscious of the optimality and unbiasedness of the Horvitz-Thompson (HT) estimator, objected about the potential bias of his estimate because of the purposive selection. Instead, he suggested random selection of an elephant with a very high probability of 99/100 for Sambo, and the rest including Jumbo (the biggest in the herd) with probability 1/4900 each. The circus owner was very unhappy with the statistician's response of 100/99 times the Sambo's weight as the estimate if Sambo got selected in this random draw, and was outraged with the response of 4900 times the Jumbo's weight if Jumbo happened to get selected. It was obvious to the owner that this new estimator was extremely poor, although he didn't know anything about its unbiasedness. The story had an unhappy ending with the circus statistician losing his job. To alleviate the instability of the HT-estimator, Hajek suggested to multiply it by 50 divided by inverse of the selection probability, which reduces simply to 50 times the weight of the selected elephant.

[4] Winsorization is a method of extreme value adjustment that replaces extreme values with the critical values used for defining low and high extreme values.

different designs. The domains must be large enough (e.g., at least size 30) to be able to define evs according to the domain-specific weight distribution. Any ev treatment to increase precision of estimates would introduce some bias. However, this bias can be reduced considerably if the ev treatment is performed under calibration controls. This is what the built-in ev control in GEM tries to accomplish.

It follows that the definition of extreme weight domains should depend on factors that affect the selection probabilities of units in the sample, such as state- and age-specific sampling rates, segment selection probabilities, pair age-specific selection probabilities, and household composition. If one tries to define extreme weight domains by taking account of all these factors via cross-classification, it will lead to too many domains with insufficient observations. That is why it is difficult to define suitable extreme weight domains for pair data. In the case of person-level weights, it was less difficult, because state by age group suitably captured the extreme weight domain requirements. The definition of extreme weight domains for pair-level weighting used in the 2017 survey was the same as the one used in the 1999–2016 surveys. The domains were defined as the cross-classification of state, pair age,[5] and number of people aged 12 to 25 in a household. In particular, the pair age was defined by the age groups of each pair member according to the age categories of 12 to 25, 26 to 49, and 50 or older (resulting in six pair age categories), and the number of people aged 12 to 25 were categorically defined as zero, one, and two or more. For more details, see Chapter 5.

[5] Pair age in this case should not be confused with the modeling term, which has a finer level breakdown.

2. Questionnaire Dwelling Unit and Pair Selection Probabilities

Similar to the 1999–2001 National Household Surveys on Drug Abuse (NHSDAs) and the 2002–2016 National Surveys on Drug Use and Health (NSDUHs),[6] the 2017 NSDUH had a two-phase design and used a computer-assisted interviewing method. There were five stages of selection: census tracts, census block groups, segments within census block groups, dwelling units (DUs) within segments, and people within dwelling units. Any two survey-eligible people had some nonzero chance of being selected and, when both were selected, they formed a within-household pair. This design feature is of interest to NSDUH researchers because, for example, it allows analysts to examine how the drug use propensity of an individual (in a family) relates to the drug use propensity of other members residing in the same dwelling unit (Center for Behavioral Health Statistics and Quality, 2018).

For the 1999–2001 surveys, the method used for selecting pairs was as follows. For a given DU, if the sum of the age-specific selection probabilities was larger than 2, then the individual person-selection probabilities were ratio adjusted downward to make the sum equal to 2. If the sum was less than 2, the difference between 2 and the sum of the probabilities was evenly distributed over three dummy people so that the sum of the person probabilities was made to equal 2. Brewer's method was then applied to select a person pair. If the selected pair consisted of two real people, then both people were selected. If the selected pair consisted of one real person and one dummy person, then the real person was selected. If the selected pair consisted of two dummy people, no one was selected from that DU.

Starting with the 2002 NSDUH and continuing through 2017, the pair-sampling algorithm was modified to increase the number of pairs selected in the sample. DUs with the sum of person selection probabilities greater than or equal to 2 were treated the same as in previous survey years. However, DUs where the sum of person-level selection probabilities was less than 2 received a slightly different treatment that increased the chance for selecting a pair of real people. Section 2.1 describes the selection process for both types of DUs.

Table 2.1 provides a summary of these NSDUH sampling units: eligible and completed screening dwelling units (SDUs), selected and completed questionnaire dwelling units (QDUs), selected and completed person interviews, and selected and completed person pairs, as well as their response rates. Using Brewer's method, zero, one, or two individuals were selected per household. Those SDUs where at least one person was selected were counted as the selected QDUs. A QDU where two people were selected and both had completed interviews was considered to be a completed person pair. The table provides a breakdown by age group at the person level and age group by selection group (none, single, or pair) at the person pair level.

[6] This report presents information from the 2017 National Survey on Drug Use and Health (NSDUH). Prior to 2002, the survey was called the National Household Survey on Drug Abuse (NHSDA).

Table 2.1 Building Blocks of the QDU and Person Pair Samples: Dwelling Units and People in the 2013–2017 NSDUHs

	2013			2014			2015			2016			2017		
Domain	**Sel.[1]**	**Resp.[2]**	**% Rate[3]**	**Sel.[1]**	**Resp.[2]**	**% Rate[3]**	**Sel.[1]**	**Resp.[2]**	**% Rate[3]**	**Sel.[1]**	**Resp.[2]**	**% Rate[3]**	**Sel.[1]**	**Resp.[2]**	**% Rate[3]**
DUs															
Total DUs Screened	190,067	160,325	84.35	154,533	127,605	82.57	165,328	132,210	79.97	173,149	135,188	78.08	184,266	138,061	74.92
QDUs															
Total QDUs	61,634	48,896	79.33	64,796	49,672	76.66	66,721	50,119	75.12	67,574	50,095	74.13	68,889	50,328	73.06
People															
Total People	88,742	67,838	76.44	91,640	67,901	74.10	94,499	68,073	72.04	95,607	67,942	71.06	97,667	68,032	69.66
12–17	27,630	22,494	81.41	21,392	17,007	79.50	21,859	16,911	77.36	22,323	17,081	76.52	22,750	17,026	74.84
18–25	28,921	22,214	76.81	21,726	16,449	75.71	23,211	17,097	73.66	22,836	16,435	71.97	23,707	16,469	69.47
26–34	8,210	6,310	76.86	14,004	10,252	73.21	14,720	10,446	70.96	15,022	10,528	70.08	15,140	10,416	68.80
35–49	12,566	9,058	72.08	19,065	13,590	71.28	19,341	13,304	68.79	19,988	13,572	67.90	20,280	13,639	67.25
50+	11,415	7,762	68.00	15,453	10,603	68.61	15,368	10,315	67.12	15,438	10,326	66.89	15,790	10,482	66.38
Non-Pairs[4]															
Total Non-Pairs	133,217	29,954	N/A	90,443	31,443	N/A	104,432	32,165	N/A	107,155	32,248	N/A	109,283	32,624	N/A
0,0	98,691	N/A	N/A	62,809	N/A	N/A	65,489	N/A	N/A	67,614	N/A	N/A	69,172	N/A	N/A
Total Singletons	34,526	29,954	86.76	37,952	31,443	82.85	38,943	32,165	82.60	39,541	32,248	81.56	40,111	32,624	81.33
0, 12–17	9,420	8,574	91.02	4,850	4,704	96.99	5,244	5,014	95.61	5,144	4,997	97.14	5,155	4,997	96.94
0, 18–25	10,535	9,475	89.94	7,250	6,647	91.68	7,583	7,102	93.66	7,647	6,895	90.17	7,858	7,079	90.09
0, 26–34	3,914	3,367	86.02	7,460	6,034	80.88	7,726	6,166	79.81	8,045	6,270	77.94	7,987	6,247	78.21
0, 35–49	4,506	3,736	82.91	8,074	6,450	79.89	8,093	6,320	78.09	8,442	6,596	78.13	8,601	6,679	77.65
0, 50+	6,151	4,802	78.07	10,318	7,608	73.74	10,297	7,563	73.45	10,263	7,490	72.98	10,510	7,622	72.52
Pairs															
Total Pairs[5]	27,108	18,942	69.88	26,844	18,229	67.91	27,778	17,954	64.63	28,033	17,847	63.66	28,778	17,704	61.52
12–17, 12–17	4,535	3,609	79.58	3,070	2,407	78.40	2,962	2,253	76.06	3,199	2,386	74.59	3,261	2,368	72.62
12–17, 18–25	3,662	2,754	75.20	2,443	1,832	74.99	2,571	1,795	69.82	2,548	1,774	69.62	2,679	1,758	65.62
12–17, 26–34	811	621	76.57	1,297	941	72.55	1,299	939	72.29	1,281	883	68.93	1,338	894	66.82
12–17, 35–49	3,834	2,756	71.88	5,530	3,940	71.25	5,654	3,888	68.77	5,829	3,930	67.42	5,845	3,870	66.21
12–17, 50+	833	571	68.55	1,132	776	68.55	1,167	769	65.90	1,123	725	64.56	1,211	771	63.67
18–25, 18–25	5,478	3,795	69.28	3,743	2,585	69.06	4,043	2,654	65.64	3,958	2,512	63.47	4,167	2,467	59.20
18–25, 26–34	1,034	690	66.73	1,378	870	63.13	1,577	975	61.83	1,429	886	62.00	1,443	839	58.14
18–25, 35–49	1,561	990	63.42	1,906	1,180	61.91	2,092	1,186	56.69	2,013	1,134	56.33	2,084	1,176	56.43
18–25, 50+	1,173	715	60.95	1,263	750	59.38	1,302	731	56.14	1,283	722	56.27	1,309	683	52.18
26–34, 26–34	822	581	70.68	1,356	865	63.79	1,492	870	58.31	1,518	905	59.62	1,551	871	56.16
26–34, 35–49	489	307	62.78	737	442	59.97	716	408	56.98	788	445	56.47	810	445	54.94
26–34, 50+	318	163	51.26	420	235	55.95	418	218	52.15	443	234	52.82	460	249	54.13
35–49, 35–49	857	512	59.74	1,160	658	56.72	1,158	635	54.84	1,213	627	51.69	1,233	628	50.93
35–49, 50+	462	245	53.03	498	262	52.61	470	232	49.36	490	213	43.47	474	213	44.94
50+, 50+	1,239	633	51.09	911	486	53.35	857	401	46.79	918	471	51.31	913	472	51.70

DU = dwelling unit; N/A = not applicable; QDU = questionnaire dwelling unit.

[1] Selected pairs are based on the screener age.

[2] Respondent pairs are based on the questionnaire age and comprise only respondent people.

[3] These rates are unweighted and based only on the total selected and total responding counts of pairs.

[4] Non-pairs are completed screening dwelling units where either zero or one person was selected.

[5] Total pairs are housing units where two people were selected.

2.1 Pair Selection Probability

2.1.1 Case I: DUs with $S \geq 2$

For a given DU, if the sum of the age-specific person selection probabilities (S) was larger than 2, then the selection probability was ratio adjusted by a multiplicative adjustment factor so that all probabilities were scaled down to sum to exactly 2. Now, Brewer's method sets the pairwise selection probabilities at

$$P_{h(ij)} = \left[\frac{P_{h(i)}P_{h(j)}}{K}\right]\left[\frac{1}{1-P_{h(i)}} + \frac{1}{1-P_{h(j)}}\right] \tag{2.1}$$

by setting K at

$$K = 2 + \sum \frac{P_{h(i)}}{1-P_{h(i)}}, \tag{2.2}$$

where $i = i^{th}$ person in household h (whose selection probability depends on his or her age category: 1, 2, 3, 4, or 5) and

$j = j^{th}$ person in household h (whose selection probability depends on his or her age category: 1, 2, 3, 4, or 5),

where age category 1 corresponds to people aged 12 to 17, 2 to people aged 18 to 25, 3 to people aged 26 to 34, 4 to people aged 35 to 49, and 5 to people aged 50 or older.

The sum of the pairwise selection probabilities taken over all unique pairs will be guaranteed to be exactly 1.

$$\sum_i \sum_{j>i} P_{h(ij)} = 1 \tag{2.3}$$

It also guarantees that the sum of the pairwise selection probabilities for an individual is equal to the individual's selection probability

$$\sum_{j \neq i} P_{h(ij)} = P_{h(i)} \tag{2.4}$$

for all values of i.

Note that the above scheme always selects a pair of two eligible people.

2.1.2 Case II: DUs with $S < 2$

If the sum S of person-level selection probabilities was less than 2, the method used in survey years 1991–2001 consisted of dividing $2 - S$ equally among the three dummy people added

to the household, and then used Brewer's method to select a pair, as in Case I. However, if the household had two or more people, we preferred a pair of real people to have a greater chance of being selected. To achieve this goal, the individual selection probabilities, $P_{h(i)}$, were scaled upward by the factor F_s such that their sum came close to but did not exceed 2 and such that each person selection probability did not exceed the maximum allowed probability of 0.99. Thus, denoting the revised person selection probabilities by $P'_{h(i)}$, the factor F_s is given by

$$F_s = Min\left\{\frac{T(\lambda)}{S}, \frac{0.99}{\max\left\{P_{h(i)}\right\}}\right\}, \tag{2.5}$$

where $T(\lambda) = S + \lambda(2 - S)$ and λ is set to 0.5. Note that if λ is chosen as 0, then $F_s = 1$ and the selection scheme would follow that of Case I. The individual person probabilities are scaled upward by the factor F_s so they either sum to 2 or sum as close to 2 as possible. Denote S' as the sum of the selection probability after scale adjustment by F_s. If S' is exactly 2, then dummy people are not needed. If S' is less than 2, then three dummy people are added to the DU.

Now, for Brewer's method, set the pairwise selection probabilities similar to (2.1), as

$$P'_{h(ij)} = \left[\frac{P'_{h(i)}P'_{h(j)}}{K'}\right]\left[\frac{1}{1-P'_{h(i)}} + \frac{1}{1-P'_{h(j)}}\right] \tag{2.6}$$

by setting K' at

$$K' = 2 + \sum_i \frac{P'_{h(i)}}{1-P'_{h(i)}}, \tag{2.7}$$

where $P'_{h(i)}$ and $P'_{h(j)}$ are the selection probabilities adjusted by the scaling factor F_s,

where $i = i^{th}$ person in the household (whose selection probability depends on his or her age category: 0, 1, 2, 3, 4, or 5),

$j = j^{th}$ person in the household (whose selection probability depends on his or her age category: 0, 1, 2, 3, 4, or 5), and

where age category 0 corresponds to dummy people, and categories 1 to 5 are defined as in Case I.

Note that we now have $\sum_{j \neq i} P'_{h(ij)} = P'_{h(i)}$. To maintain the original person selection probabilities despite the scale adjustment by F_s, we modified Brewer's method as follows. First, draw a random number, R, from a uniform (0,1) distribution. If $R \leq 1/F_s$, then select a pair using Brewer's method based on formula (2.6). However, if $R > 1/F_s$, then no people are selected from the

household. In this way, the probability for selecting a pair (*i,j*) in household *h* becomes $P^*_{h(ij)} = P'_{h(ij)}/F_s$, which, in turn, gives the original person selection probabilities, $P_{h(i)}$. Unlike Case I, where a pair of eligible people was always selected, this adjusted selection scheme allows for zero, one, or two people to be selected from a DU.

2.2 Questionnaire Dwelling Unit Selection Probability

A dwelling unit was considered a selected QDU if it had completed the screening interview and had at least one person selected for the questionnaire interview. QDUs with at least one respondent were considered respondent QDUs.

The QDU selection probability was defined as

$$P^*_h = \left(1 - P^*_{h(00)}\right), \tag{2.8}$$

where $P^*_{h(00)}$ is the probability of not selecting any person. For the DUs with an unadjusted sum of age-specific selection probabilities larger than or equal to 2 (Case I), $P^*_{h(00)}$ is 0. It follows from Section 2.1, under Case II, $P^*_{h(00)}$ can be calculated as

$$P^*_{h(00)} = \left(1 - \frac{1}{F_s}\right) + \frac{3}{F_s}\left[\frac{P'_{h(0)}P'_{h(0)}}{K'}\right]\left[\frac{1}{1-P'_{h(0)}} + \frac{1}{1-P'_{h(0)}}\right], \tag{2.9}$$

where $P'_{h(0)}$ is the selection probability of a dummy person when person selection probabilities are adjusted by F_s.

This page intentionally left blank

3. Brief Description of the Generalized Exponential Model

In survey practice, design-based weights are typically adjusted in three steps: (1) for extreme values (ev) via winsorization, (2) for nonresponse (nr) via weighting classes, and (3) for poststratification (ps) via raking ratio adjustments. If weights are not treated for extreme values, the resulting estimates, although unbiased, will tend to have low precision. The bias introduced by winsorization is alleviated to some extent through ps. The nr adjustment is a correction for bias introduced in estimates based only on responding units, and ps is an adjustment for coverage (typically undercoverage) bias and variance reduction due to correlation between the study and control (usually demographic) variables.

There are limitations in the existing methods of weight adjustment for ev, nr, and ps. It would be desirable to adjust for bias introduced in the ev step (when extreme weights are treated via winsorization) in that the sample distribution for various demographic characteristics is preserved. For the nr step, there are general raking type methods, such as the scaled constrained exponential model developed by Folsom and Witt (1994), where the lower and upper bounds can be suitably chosen by use of a separate scaling factor. The factor is set as the inverse of the overall response propensity. It would be desirable to have a model for the nr adjustment factor so that the desired lower and upper bounds on the factor are part of the model. Note that the lower bound on the nr adjustment factor should be one, as it is interpreted as the inverse of the probability of response for a particular unit. For the ps step, on the other hand, the general calibration methods of Deville and Särndal (1992), such as the logit method, allow for built-in lower (L) and upper (U) bounds (for ps, typically $L < 1 < U$). However, it would be desirable to have nonuniform bounds (L_k, U_k) depending on the unit k such that the final adjusted weight, w_k, could be controlled within certain limits. An important application of this feature would be weight adjustments in the presence of ev to allow some control on the final adjustment of the initially identified extreme values.

A modification of the earlier method of the scaled constrained exponential model of Folsom and Witt (1994), termed as the method of the generalized exponential model (GEM) and proposed by Folsom and Singh (2000), provides a unified approach to the three weight adjustments for ev, nr, and ps, and it has the desired features mentioned above. The functional form of the GEM adjustment factor is provided in Appendix A. It generalizes the logit model of Deville and Särndal (1992), typically used for ps, such that the bounds (L, U) may depend on k. Thus, it provides a built-in control on ev during both ps and nr adjustments. In addition, the bounds are internal to the model and can be set to chosen values (e.g., $L_k = 1$ in the nr step). If there is a low frequency of ev in the final ps, then a separate ev step may not be necessary.

In fitting GEM to a particular problem, the choice of a large number of predictor variables along with tight bounds will have an impact on the resulting unequal weighting effect (UWE) and the proportion of extreme values. In practice, this leads to somewhat subjective considerations of trade-off between the target set of bounds for a given set of factor effects and the target UWE and the target proportion of extreme values. It also may be beneficial to look at

the proportion of "outwinsors" (a term coined to signify the extent of residual weights after winsorization), which is probably more realistic in determining the robustness of estimates in the presence of extreme values.

A large increase in the number of predictor variables in GEM typically would result in a higher UWE, thus indicating a possible loss in precision. This was checked by comparing SUDAAN-based standard errors of a key set of estimates computed from two sets of calibration models, one baseline using only the main effects and the other using the final model. The results are presented in Chapter 7.

To implement GEM, several steps need to be followed: (1) define and create all the covariates; (2) define the extreme weights; (3) fit the GEM model. The details of practical aspects of GEM implementation can be found in Chapters 4 and 5 of this report and Chapter 4 of the National Survey on Drug Use and Health Methodological Resource Book person-level sampling weight calibration report (Center for Behavioral Health Statistics and Quality, 2019b).

4. Predictor Variables for the Questionnaire Dwelling Unit and Pair Weight Calibration via the Generalized Exponential Model

We note that unlike the person-level weight calibration, the control totals for the questionnaire dwelling unit (QDU)-level and person pair-level poststratification are not available from the U.S. Census Bureau. A way around this problem is to take advantage of the two-phase nature of the design, in which the screener data provide a large sample containing demographic information that can be used to derive control totals for the QDU-level and person pair-level sampling weight calibrations, as well as for the selected person poststratification adjustment. The stability of control totals from the screener dwelling unit (SDU)-level data can be improved by poststratification of the SDU sample using person-level counts from the census. This was indeed done and is documented in the National Survey on Drug Use and Health Methodological Resource Book person-level sampling weight calibration report (Center for Behavioral Health Statistics and Quality, 2019b).

4.1 Questionnaire Dwelling Unit Weight Calibration

After the nonresponse and poststratification adjustments at the SDU level, which are common to the person-level weight calibration, the QDU sample weights were adjusted in three steps: poststratification of selected QDUs, nonresponse adjustment of respondent QDUs, and poststratification of respondent QDUs. The set of initially proposed predictor variables for these adjustments using the generalized exponential model (GEM) were set to be common and to correspond to those used for the SDU nonresponse and poststratification adjustments. The variables are of two types: Those used for SDU nonresponse adjustment are 0/1 indicators, while those used for SDU poststratification adjustment are counting variables. The variables of the first type (0/1 indicators) are population density,[7] group quarters, race/ethnicity of householder, percentage of people in segment who are black or African American, percentage of people in segment who are Hispanic or Latino, percentage of owner-occupied dwelling units (DUs) in segment, segment-combined median rent and housing value, and household type. Variables of the second type (counting variables) represent the number of eligible people within each DU who fall into the various demographic categories of race, age group, Hispanicity, and gender. Note that the state and quarter variables are represented as both binary and counting variables. Thus, not only are DU counts within a specific state or quarter in the QDU sample controlled to the corresponding totals obtained from the SDU sample, but also counts of people living in the DUs in the QDU sample are controlled to totals from the SDU sample. These person-level totals match the census estimates because of the SDU-level poststratification to census counts. It may be noted that in the poststratification of selected QDUs and the nonresponse adjustment of the respondent QDUs steps, demographic information from screener data was used in defining

[7] Population density, percentage of people in segment who are black or African American, percentage of people in segment who are Hispanic or Latino, percentage of owner-occupied dwelling units in segment, and segment-combined median rent and housing value were defined using 2010 U.S. Census Bureau data.

covariates, whereas in the poststratification of the selected QDUs step, questionnaire demographic information was used.

Exhibit 4.1 lists all predictor variables proposed for QDU-level calibration and identifies them as counting, binary, or both. Various main effects and higher-level factor effects based on the predictor variables were included in the GEM modeling. As stated previously, all adjustment steps at the QDU level used a common set of proposed predictor variables.

4.2 Pair Weight Calibration

Like QDU, the initial set of weight components in pair weight calibration are the same as the set obtained from the SDU-level weight calibration. The SDU-calibrated weight is multiplied by the pair-level design weight, which in turn was adjusted in four steps: poststratification of selected pairs, nonresponse adjustment of respondent pairs, poststratification of respondent pairs, and the extreme weight adjustment of respondent pairs. All the adjustment steps for pair weights utilized the same set of initially proposed predictor variables, which included a subset of those used for the person-level nonresponse adjustment. This included segment characteristic variables, such as population density, percentage of people in segment who are black or African American, percentage of people in segment who are Hispanic or Latino, percentage of owner-occupied DUs in segment, and segment-combined median rent and housing value. Also included were pair-specific covariates, such as the demographic characteristics of pair age, pair race/ethnicity, and pair gender, as well as dwelling unit characteristics, such as race/ethnicity of householder, household type, household size, and group quarters indicators. State and quarter indicators were included as well. However, for two-factor effects, instead of individual state, state/region was used because of insufficient sample size. This resulted in a 12-level variable where the eight largest sample states were kept separate, and the remainder of states were grouped according to the four census regions. All variables were defined as 0/1 indicators. These proposed predictor variables and their levels are shown in Exhibit 4.2.

In the poststratification of selected pairs and the nonresponse adjustment of respondent pairs, screener data were used in the definition of the pair-specific variables such as pair age, pair race/ethnicity, and pair gender, whereas in the poststratification and extreme weight adjustment of respondent pairs, these variables were obtained from the questionnaire. For the latter case, in addition to the variables described above, indicator covariates corresponding to selected pair domains were included to perform Hajek-type ratio adjustments via weight calibration, as mentioned in Chapter 1. The selected pair domains were limited to 10 of the 14 pair domains listed in Chapter 1. (Parent-child pairs where the child was in the 15- to 17-year-old age range and sibling-sibling pairs with focus on the younger child were not included in the poststratification.) The inclusion of these pair domain covariates led to the use of two sets of control totals in the modeling. Details of the construction of these control totals can be found in Appendix B.

Exhibit 4.1 Definitions of Levels for QDU-Level Calibration Modeling Variables

Age[a]
1: 12–17, 2: 18–25, 3: 26–34, 4: 35–49, 5: 50+[1]

Gender[a]
1: Male, 2: Female[1]

Group Quarter Indicator[b]
1: College Dorm, 2: Other Group Quarter, 3: Non-Group Quarter[1]

Hispanicity[a]
1: Hispanic or Latino, 2: Non-Hispanic or Latino[1]

Household Size[a]
Continuous Variable Count of Individuals Rostered with DU

Household Type (Ages of People Rostered within DU)[b]
1: 12–17, 18–25, 26+; 2: 12–17, 18–25; 3: 12–17, 26+; 4: 18–25, 26+; 5: 12–17, 6: 18–25; 7: 26+[1]

Percentage of Owner-Occupied Dwelling Units in Segment (% Owner-Occupied)[b]
1: 50–100%,[1] 2: 10–<50%, 3: 0–<10%

Percentage of Segments That Are Black or African American[b]
1: 50–100%, 2: 10–<50%, 3: 0–<10%[1]

Percentage of Segments That Are Hispanic or Latino[b]
1: 50–100%, 2: 10–<50%, 3: 0–<10%[1]

Population Density[b]
1: MSA 1,000,000 or More, 2: MSA Less than 1,000,000, 3: Non-MSA Urban, 4: Non-MSA Rural[1]

Quarter[a,b]
1: Quarter 1, 2: Quarter 2, 3: Quarter 3, 4: Quarter 4[1]

Race (3 Levels)[a]
1: White,[1] 2: Black or African American, 3: Other

Race (5 Levels)[a]
1: White,[1] 2: Black or African American, 3: American Indian or Alaska Native, 4: Asian, 5: Two or More Races

Race/Ethnicity of Householder[b]
1: Hispanic or Latino White,[1] 2: Hispanic or Latino Black or African American, 3: Hispanic or Latino Other, 4: Non-Hispanic or Latino White, 5: Non-Hispanic or Latino Black or African American, 6: Non-Hispanic or Latino Other

Segment-Combined Median Rent and Housing Value (Rent/Housing)[b,2]
1: First Quintile, 2: Second Quintile, 3: Third Quintile, 4: Fourth Quintile, 5: Fifth Quintile[1]

States[a,b,3]
Model Group 1: 1: Connecticut, 2: Maine, 3: Massachusetts,[1] 4: New Hampshire, 5: New Jersey, 6: New York, 7: Pennsylvania, 8: Rhode Island, 9: Vermont
Model Group 2: 1: Illinois, 2: Indiana, 3: Iowa, 4: Kansas, 5: Michigan, 6: Minnesota, 7: Missouri, 8: Nebraska, 9: North Dakota, 10: Ohio, 11: South Dakota, 12: Wisconsin[1]
Model Group 3: 1: Alabama, 2: Arkansas, 3: Delaware, 4: District of Columbia, 5: Florida, 6: Georgia, 7: Kentucky, 8: Louisiana, 9: Maryland, 10: Mississippi, 11: North Carolina,[1] 12: Oklahoma, 13: South Carolina, 14: Tennessee, 15: Texas, 16: Virginia, 17: West Virginia
Model Group 4: 1: Alaska, 2: Arizona,[1] 3: California, 4: Colorado, 5: Idaho, 6: Hawaii, 7: Montana, 8: Nevada, 9: New Mexico, 10: Oregon, 11: Utah, 12: Washington, 13: Wyoming

State/Region[b,3]
Model Group 1: 1: New York, 2: Pennsylvania, 3: Other[1]
Model Group 2: 1: Illinois, 2: Michigan, 3: Ohio, 4: Other[1]
Model Group 3: 1: Florida, 2: Texas, 3: Other[1]
Model Group 4: 1: California, 2: Other[1]

DU = dwelling unit; MSA = metropolitan statistical area; QDU = questionnaire dwelling unit.
[1] The reference level for this variable. This is the level against which effects of other factor levels are measured.
[2] Segment-Combined Median Rent and Housing Value is a composite measure based on rent, housing value, and percentage owner-occupied.
[3] The states or district assigned to a particular model is based on census regions.
[a] Counting variable. A count of all people in the household.
[b] Binary variable.

Exhibit 4.2 Definitions of Levels for Pair-Level Calibration Modeling Variables

Group Quarter Indicator
1: College Dorm, 2: Other Group Quarter, 3: Non-Group Quarter[1]

Household Size
1: DU with 2 People,[1] 2: DU with 3 People, 3: DU with ≥ 4 People

Pair Age (15 Levels)
1: 12–17 and 12–17,[1] 2: 12–17 and 18–25, 3: 12–17 and 26–34, 4: 12–17 and 35–49, 5: 12–17 and 50+, 6: 18–25 and 18–25, 7: 18–25 and 26–34, 8: 18–25 and 35–49, 9: 18–25 and 50+, 10: 26–34 and 26–34, 11: 26–34 and 35–49, 12: 26–34 and 50+, 13: 35–49 and 35–49, 14: 35–49 and 50+, 15: 50+ and 50+

Pair Age (6 Levels)
1: 12–17 and 12–17,[1] 2: 12–17 and 18–25, 3: 12–17 and 26+, 4: 18–25 and 18–25, 5: 18–25 and 26+, 6: 26+ and 26+

Pair Age (3 Levels)
1: 12–17 and 12–17,[1] 2: 12–17 and 18+, 3: 18+ and 18+

Pair Gender
1: Male and Female,[1] 2: Female and Female, 3: Male and Male

Pair Race/Ethnicity (10 Levels)
1: White and White,[1] 2: White and Black or African American, 3: White and Hispanic or Latino, 4: White and Other, 5: Black or African American and Black or African American, 6: Black or African American and Hispanic or Latino, 7: Black or African American and Other, 8: Hispanic or Latino and Hispanic or Latino, 9: Hispanic or Latino and Other, 10: Other and Other

Pair Race/Ethnicity (5 Levels)
1: Two or More Races Pair, 2: Hispanic or Latino Pair, 3: Black or African-American Pair, 4: White Pair,[1] 5: Other Pair

Pair Race/Ethnicity (4 Levels)
1: Two or More Races Pair or Other and Other, 2: Hispanic or Latino Pair, 3: Black or African-American Pair, 4: White Pair[1]

Percentage of Owner-Occupied Dwelling Units in Segment (% Owner-Occupied)
1: 50–100%,[1] 2: 10–<50%, 3: 0–<10%

Percentage of Segments That Are Black or African American
1: 50–100%, 2: 10–<50%, 3: 0–<10%[1]

Percentage of Segments That Are Hispanic or Latino
1: 50–100%, 2: 10–<50%, 3: 0–<10%[1]

Segment-Combined Median Rent and Housing Value (Rent/Housing)[2]
1: First Quintile, 2: Second Quintile, 3: Third Quintile, 4: Fourth Quintile, 5: Fifth Quintile[1]

Population Density
1: MSA 1,000,000 or More, 2: MSA Less than 1,000,000, 3: Non-MSA Urban, 4: Non-MSA Rural[1]

Quarter
1: Quarter 1, 2: Quarter 2, 3: Quarter 3, 4: Quarter 4[1]

Race/Ethnicity of Householder
1: Hispanic or Latino White,[1] 2: Hispanic or Latino Black or African American, 3: Hispanic or Latino Other, 4: Non-Hispanic or Latino White, 5: Non-Hispanic or Latino Black or African American, 6: Non-Hispanic or Latino Other

Exhibit 4.2 Definitions of Levels for Pair-Level Calibration Modeling Variables (continued)

State/Region

Model Group 1: 1: Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, Rhode Island, Vermont; 2: Alabama, Arkansas, Delaware, District of Columbia, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Virginia, West Virginia;[1] 3: New York; 4: Pennsylvania; 5: Florida; 6: Texas

Model Group 2: 1: Indiana, Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota, Wisconsin;[1] 2: Alaska, Arizona, Colorado, Idaho, Hawaii, Montana, Nevada, New Mexico, Oregon, Utah, Washington, Wyoming; 3: Michigan; 4: Illinois; 5: Ohio; 6: California

States[3]

Model Group 1: 1: Alabama, 2: Arkansas, 3: Connecticut, 4: Delaware, 5: District of Columbia, 6: Florida, 7: Georgia, 8: Kentucky, 9: Louisiana, 10: Maine, 11: Maryland,[1] 12: Massachusetts, 13: Mississippi, 14: New Hampshire, 15: New Jersey, 16: New York, 17: North Carolina, 18: Oklahoma, 19: Pennsylvania, 20: Rhode Island, 21: South Carolina, 22: Tennessee, 23: Texas, 24: Vermont, 25: Virginia, 26: West Virginia

Model Group 2: 1: Alaska, 2: Arizona,[1] 3: California, 4: Colorado, 5: Idaho, 6: Illinois, 7: Indiana, 8: Iowa, 9: Hawaii, 10: Kansas, 11: Michigan, 12: Minnesota, 13: Missouri, 14: Montana, 15: Nebraska, 16: Nevada, 17: New Mexico, 18: North Dakota, 19: Ohio, 20: Oregon, 21: South Dakota, 22: Utah, 23: Washington, 24: Wisconsin, 25: Wyoming

Pair Relationship Associated with Multiplicity

1: Parent-Child (12–14)*
2: Parent-Child (12–17)*
3: Parent-Child (12–20)*
4: Parent*-Child (12–14)
5: Parent*-Child (12–17)
6: Parent*-Child (12–20)
7: Sibling (12–14)-Sibling (15–17)*
8: Sibling (12–17)-Sibling (18–25)*
9: Spouse-Spouse/Partner-Partner
10: Spouse-Spouse/Partner-Partner with Children (Younger than 18)

DU = dwelling unit; MSA = metropolitan statistical area.

[1] The reference level for this variable. This is the level against which effects of other factor levels are measured.

[2] Segment-Combined Median Rent and Housing Value is a composite measure based on rent, housing value, and percentage owner-occupied.

[3] The states or district assigned to a particular model is based on combined census regions.

* The pair member focused on.

This page intentionally left blank

5. Definition of Extreme Weights

An important feature of the generalized exponential model (GEM) is the built-in provision of extreme value (ev) treatment. Sampling weights are often classified as extreme (high or low) if they fall outside the interval, median ± 3 × interquartile range (IQR). The interval is set for prespecified domains defined usually by design variables corresponding to deep stratification.[8] Similar to previous National Surveys on Drug Use and Health (NSDUHs), for the GEM modeling used in the 2017 NSDUH, a more conservative (narrower) interval was defined, median ± 2.5 × IQR. The narrower interval better prevents the adjusted weights from crossing the standard interval boundaries by treating weights near but not outside the commonly used boundaries (i.e., those that have the most potential to become extreme) as extreme as well.

Denote the interval boundaries (or critical values) for low and high extreme values by $b_{k(l)}$ and $b_{k(u)}$, respectively. For implementing ev control via GEM, the variable m_k was defined as the minimum of $b_{k(u)} / w_k$ and one for high extreme weights, and the maximum of $b_{k(l)} / w_k$ and one for low extreme weights, where w_k represents the sampling weight before adjustment, and $b_{k(u)}$ and $b_{k(l)}$ denote the critical values for the extreme weights. Note that under this definition, for high extreme weights, the more extreme the weight is, the smaller m_k will be, and, conversely, for low extreme weights, the more extreme the weight is, the bigger m_k will be. Nonextreme weights had a value of one for m_k. The upper and lower bounds for the adjustment factors were defined, respectively, as the product of m_k and the upper and lower boundary parameters of GEM. GEM allows inputs of up to three different upper and lower boundary parameters (L_1 and U_1, L_2 and U_2, L_3 and U_3) for high, non-, and low extreme weights. By applying a small upper boundary parameter for high extreme weights and a large lower boundary parameter for low extreme weights, the extreme weights can be controlled in the modeling process.

5.1 Questionnaire Dwelling Unit Extreme Weight Definition

For the questionnaire dwelling unit-level weight adjustment, extreme weights were defined using a nested hierarchy of six domains:

1. State;
2. State sampling region;
3. State by household type;

   Levels of household type indicate whether the household has members who are youths, young adults, or adults, where youth signifies 12- to 17-year-olds, young adult 18- to 25-year-olds, and adult 26 years or older.

[8] Deep stratification refers to the stratification that was used in the sample design. In the case of the 2017 survey, deep stratification refers to the cross-classification of state sampling region by age group.

a. Youth, Young Adult, Adult;
b. Youth, Young Adult;
c. Youth, Adult;
d. Young Adult, Adult;
e. Youth Only;
f. Young Adult Only; and
g. Adult Only.

4. State sampling region by household type;
5. State by household type by household size (1, 2, 3, 4+); and
6. State sampling region by household type by household size.

The hierarchy is used to satisfy the minimum of 30 observations for defining the boundaries for extreme values. If this sample size requirement is not met at the lower level, then the next level up in the hierarchy is used.

5.2 Person Pair Extreme Weight Definition

The pair selection probability is a function of the selection probability of each person in the pair given by formula (2.1) or (2.6), depending on the sum of the person selection probabilities within the household as discussed in Section 2.1. This probability can be very small if the selection probabilities of individual members are small. For example, consider a particular selected dwelling unit (DU) from the 2017 survey. This DU gave rise to a selected pair of respondents, one aged 60 and the other aged 51. The selection probability in this DU was 0.17305 for a respondent aged 50 or older. Using the formula (2.6) in Chapter 2, the pair selection probability was computed to be 0.000708147. Therefore, the inverse of the selection probability, the pair-level design weight, was 1,412.14. Thus, a small pair selection probability can create a high initial weight, which is the product of the screener dwelling unit weight and the person pair design-based weight.

As mentioned in the introduction, it turns out to be difficult to select suitable domains for defining extreme weights for pair-level data. However, as was done for the 1999–2016 surveys, the extreme weight definition was based on the following hierarchy of domains:

1. Pair age group[9] (with three age categories, 12 to 25, 26 to 49, and 50+) by number (0, 1, 2+) of people aged 12 to 25 in the household;
2. State cluster (with five levels [explained below]) by pair age group by number (0, 1, 2+) of people aged 12 to 25 in the household;
3. State cluster (with three levels [explained below]) by pair age group by number (0, 1, 2+) of people aged 12 to 25 in the household; and
4. State by pair age group by number of people aged 12 to 25 (0, 1, 2+) in the household.

[9] Pair age in this case should not be confused with the modeling term, which has a finer level breakdown.

The hierarchy was used to satisfy the minimum of 30 observations for defining the boundaries for extreme values. If this sample size requirement was not met at the lower level, then the next level up in the hierarchy was used.

We now briefly introduce the considerations behind the above definition for extreme weight domains. The sample design prespecified the person-level selection probability within state by five age groups (12 to 17, 18 to 25, 26 to 34, 35 to 49, 50+). Age groups 12 to 17 and 18 to 25 have a relatively similar selection probability, and the same is true for age groups 26 to 34 and 35 to 49. The 50+ group, however, has a quite different selection probability from the other groups. Furthermore, since the 12 to 17 and 18 to 25 age groups have large selection probabilities, they have a very high chance of being selected if the household has people in these age groups. Therefore, the number of people aged 12 to 25 in the household has a significant impact on the type of pair selected and the pair selection probability. Taking into consideration these design-related features, a suitable domain to define the pair-level extreme weight seems to be given by state by pair age group by number of people aged 12 to 25 in the household.

The hierarchy of domains mentioned above was used to satisfy the minimum of 30 observations. However, it was found that for many ev domains, the minimum sample size requirement was not met. To alleviate this problem, states were grouped into a small number of clusters, such as three or five. The assignment of states to clusters was determined by the clustering algorithm in PROC CLUSTER in SAS, where the clustering variable was defined as the average person-level weight (ANALWT) for each of the five age groups within each state. The choice of the average person-level weight for each group for each state was motivated from the objective of finding a single variable that would reflect the design-based difference in pair selection probabilities across states. Even with clustering of states, the ev domain sample size was insufficient in some cases, so the most general level of the hierarchy, the national level, was required. Furthermore, at the national level, we had to collapse some pair age categories in forming domains of reasonable sample size to define extreme weights. More specifically, for the national level, we collapsed all levels of number of people aged 12 to 25 for the pair age groups of 50+, 50+ and 26 to 49, 50+. In addition, levels 1 and 2+ of number of people aged 12 to 25 were combined for the pair age group of 26 to 49, 26 to 49.

This page intentionally left blank

6. Weight Calibration at Questionnaire Dwelling Unit and Pair Levels

The 2017 National Survey on Drug Use and Health (NSDUH) was based on probability sampling so that valid inferences can be made from survey findings about the target population. Probability sampling refers to sampling in which every unit on the frame is given a known, nonzero probability for inclusion in the survey. This is required for unbiased estimation of the population total. The assumption of nonzero inclusion probability for every pair of units in the frame also is required for unbiased variance estimation. The 2014–2017 NSDUH sample design plans slightly modified the 2005–2013 approach, such that the basic sampling plan involved five stages of selection across two phases of design: within Phase I, (1) the selection of census tracts and (2) census block groups within each state sampling (SS) region, (3) the selection of subareas or segments (comprising U.S. Census Bureau blocks) within SS regions; (4) the selection of dwelling units (DUs) within these subareas; and, finally, within Phase II, (5) the selection of eligible individuals within DUs. Specific details of the sample design and selection procedures for the sample and changes to the design for this year can be found in the 2017 NSDUH Methodological Resource Book (MRB) sample design report (Center for Behavioral Health Statistics and Quality [CBHSQ], 2018).

As part of the postsurvey data-processing activities, analysis weights that reflected the selection probabilities from various stages of the sample design were calculated for respondents. These sample weights were adjusted at the DU (screening sample), questionnaire dwelling unit (QDU), person, and paired respondent levels (the latter three all based around the questionnaire sample) to account for bias due to extreme values (ev), nonresponse (nr), and coverage.

The final sample weights for Phase I screener dwelling units (SDU) and Phase II QDU, person, and pair levels for the 2017 samples consisted of products of several factors, each representing either a probability of selection at some particular stage or some form of ev, nr, or poststratification (ps) calibration adjustment. In the following sections, we describe the QDU and pair weight components in greater detail. In summary, the first 11 factors were defined for all SDUs and reflected the fully adjusted SDU sample weight. The remaining components branched to reflect QDU and pair selection probabilities, as well as additional adjustments for ev, nr, and ps. Note that the final QDU and pair weights for the 2017 survey sample are the product of all weight components for each type of sample, illustrated in Exhibits 6.1 and 6.2.

For QDU data, generalized exponential modeling (GEM) calibration modeling was applied by partitioning the data into four groups of states: Northeast, South, Midwest, and West, based on census regions in the interest of computational feasibility. Previous experience showed that with current computing power, the large number of variables and records prevented any further reduction of modeling groups.

For pair data, GEM modeling was initially applied by partitioning the pair data into four groups based on census regions. However, there were not enough observations in each group to fit a comprehensive model to reduce bias. Alternatively, a single model was attempted for the whole pair data, but it was rejected as not practical due to computational limitations.

A compromise approach was adopted by combining census regions into two groups: Northeast with South and Midwest with West. This grouping proved both manageable and desirable as it assisted in bias reduction, ease of modeling, and workload reduction. Exhibit 6.3 provides more details of the data partition for GEM modeling. The resulting sample sizes of selected and respondent units for the pair and QDU data partitions are shown for the 2013–2017 surveys in Table 6.1.

It may be noted that for the pair data in the 1999, 2000, and 2001 surveys, the built-in ev control feature of GEM was not used until the final respondent pair ev adjustment step. The reason for this is that the definition for ev domain was not finalized before the pair data calibration process was begun. However, for the 2002–2017 survey pair data, the built-in ev control feature was used for each adjustment step.

Exhibit 6.1 Summary of 2017 NSDUH QDU Sample Weight Components

Phase I Screener Dwelling Unit Level

Design Weight Components	
#1	Inverse Probability of Selecting Census Tract
#2	Inverse Probability of Selecting Census Block Groups
#3	Inverse Probability of Selecting Segment
#4	Quarter Segment Weight Adjustment
#5	Subsegmentation Inflation Adjustment
#6	Inverse Probability of Selecting SDU
#7	Subsampling of Added SDU Adjustment
#8	SDU Release Adjustment

Weight Adjustment*	
#9	SDU Nonresponse Adjustment ***(res.sdu.nr)***
#10	SDU Poststratification Adjustment ***(res.sdu.ps)***
#11	SDU Extreme Value Adjustment ***(res.sdu.ev)***

Phase II Questionnaire Dwelling Unit Level

Design Weight Component	
#12	Inverse of Selection Probability of at Least One Person in the Dwelling Unit

Weight Adjustment*	
#13	Selecting QDU Poststratification to SDU-Based Control Totals ***(sel.qdu.ps)***
#14	Respondent QDU Nonresponse Adjustment ***(res.qdu.nr)***
#15	Respondent QDU Poststratification to SDU-Based Control Totals ***(res.qdu.ps)***
#16	Respondent QDU Extreme Value Adjustment ***(res.qdu.ev)***

QDU = questionnaire dwelling unit; SDU = screener dwelling unit.

* These adjustments use the generalized exponential model (GEM), which also involves pre- and post-processing in addition to running the GEM macro. See Exhibit 4.1 in the NSDUH Methodological Resource Book person-level sampling weight calibration report (Center for Behavioral Health Statistics and Quality, 2019b). For computational feasibility, all weight adjustments were done using the four model groups based on census regions defined in Exhibit 6.3.

Exhibit 6.2 Summary of 2017 NSDUH Person Pair Sample Weight Components

Phase I Screener Dwelling Unit Level

Design Weight Components	
#1	Inverse Probability of Selecting Census Tract
#2	Inverse Probability of Selecting Census Block Groups
#3	Inverse Probability of Selecting Segment
#4	Quarter Segment Weight Adjustment
#5	Subsegmentation Inflation Adjustment
#6	Inverse Probability of Selecting SDU
#7	Subsampling of Added SDU Adjustment
#8	SDU Release Adjustment

Weight Adjustment*	
#9	SDU Nonresponse Adjustment ***(res.sdu.nr)***
#10	SDU Poststratification Adjustment ***(res.sdu.ps)***
#11	SDU Extreme Value Adjustment ***(res.sdu.ev)***

Phase II Person Pair Level

Design Weight Component	
#12	Inverse of Selection Probability of a Person Pair in SDU

Weight Adjustment*	
#13	Selected Pair Poststratification to SDU-Based Control Totals ***(sel.pr.ps)***
#14	Respondent Pair Nonresponse Adjustment ***(res.pr.nr)***
#15	Respondent Pair Poststratification Adjustment to SDU-Based Control Totals ***(res.per.ps)***
#16	Respondent Pair Extreme Value Adjustment ***(res.per.ev)***

QDU = questionnaire dwelling unit; SDU = screener dwelling unit.

* These adjustments use the generalized exponential model (GEM), which also involves pre- and post-processing in addition to running the GEM macro. See Exhibit 4.1 in the NSDUH Methodological Resource Book person-level sampling weight calibration report (Center for Behavioral Health Statistics and Quality, 2019b). For computational feasibility, all weight adjustments were done using the four model groups based on census regions defined in Exhibit 6.3.

Exhibit 6.3 U.S. Census Bureau Regions/Model Groups

Model Group	Census Region
QDU	
1	**Northeast (9 States)** Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont
2	**Midwest (12 States)** Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, Wisconsin
3	**South (16 States and the District of Columbia)** Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, West Virginia
4	**West (13 States)** Alaska, Arizona, California, Colorado, Idaho, Hawaii, Montana, Nevada, New Mexico, Oregon, Utah, Washington, Wyoming
Pair	
1	**Northeast + South (25 States and the District of Columbia)** Alabama, Arkansas, Connecticut, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maine, Massachusetts, Maryland, Mississippi, New Hampshire, New Jersey, New York, North Carolina, Oklahoma, Pennsylvania, Rhode Island, South Carolina, Tennessee, Texas, Vermont, Virginia, West Virginia
2	**Midwest + West (25 States)** Alaska, Arizona, California, Colorado, Hawaii, Idaho, Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Montana, Nebraska, Nevada, New Mexico, North Dakota, Ohio, Oregon, South Dakota, Utah, Washington, Wisconsin, Wyoming

Table 6.1 Sample Size, by Model Group at QDU and Pair Levels

Model Group	2013		2014		2015		2016		2017	
	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs
QDU										
Northeast	12,791	9,954	12,950	9,664	13,519	9,777	13,414	9,552	14,037	9,915
South	18,766	15,073	21,448	16,680	21,887	16,708	22,287	16,810	22,628	16,901
Midwest	17,207	13,519	15,276	11,618	15,808	11,698	16,025	11,768	16,282	11,760
West	12,870	10,350	15,122	11,710	15,507	11,936	15,848	11,965	15,942	11,752
Total	61,634	48,896	64,796	49,672	66,721	50,119	67,574	50,095	68,889	50,328
Model Group	**2013**		**2014**		**2015**		**2016**		**2017**	
	Selected Pairs	Completed Pairs	Selected Pairs	Completed Pairs	Selected Pairs	Completed Pairs	Selected Pairs	Completed Pairs	Selected Pairs	Completed Pairs
Pair										
Northeast + South	13,535	9,416	13,969	9,436	14,502	9,309	14,543	9,182	15,072	9,284
Midwest + West	13,573	9,526	12,875	8,793	13,276	8,645	13,490	8,665	13,706	8,420
Total	27,108	18,942	26,844	18,229	27,778	17,954	28,033	17,847	28,778	17,704

QDU = questionnaire dwelling unit.

6.1 Phase I SDU-Level Weight Components

A total of 11 weight components for the SDU level correspond to selection probabilities and nr, ps, and ev adjustment factors. Note that this differs from previous National Household Surveys on Drug Abuse and NSDUHs in that beginning in 2014, a new design-based component was incorporated at the beginning of the process so that corresponding weight component numbers are incremented by one when compared to previous survey years with an otherwise similar weighting scheme. The first eight components in the Phase I sample weights reflect the probability of selecting the DUs. These components were derived from (1) the probability of selecting the census tract and (2) census block groups within each SS region, (3) the probability of selecting the geographic segment within each SS region, (4) a quarter segment weight adjustment, (5) a subsegmentation inflation factor, (6) the probability of selecting a DU from within each counted and listed sampled segment, (7) the probability of inclusion of added DUs, and (8) DU percent release adjustment. The three remaining weight components, #9 through #11, are GEM calibration adjustments accounting for (9) DU nonresponse at the screening level, (10) DU poststratification to census controls, and (11) DU-level ev adjustment, although in 2017, ev adjustment at this stage was deemed unnecessary, and thus Weight Component #11 was set to one for all respondent DUs. The person-level, QDU-level, and person pair-level weights use the product of the above 11 weight components as the common initial weight before further adjustments. For more detailed information on Weight Components #1 through #3 and #5 through #8, refer to the 2017 NSDUH MRB sample design report (CBHSQ, 2018), and for more detail on Weight Components #4 and #9 through #11, see the 2017 NSDUH MRB person-level sampling weight calibration report (CBHSQ, 2019b).

Note that from 2008 to 2010, there was an occasional second subsegmentation step when the initial partitioning of segments was insufficient because of out-of-date census counts or the

segment was still too large to list after the original subsegmentation. This second partitioning was not accounted for in the weighting over these survey years. A comparison was done to evaluate the effect of this omission, and it was determined that the missing second subsegmenting factor in the analysis weight had minimal impact on estimates. Therefore, weights for these years were not re-created with a correcting factor. Additional detail can be found in CBHSQ (2018).

Weight Component #2, an component reflecting the selection of one census block group from each selected census tract, was included beginning in 2014. This step was added to allow for possible transitioning to an address-based sampling design in the future. Additional changes to sample allocation and survey design are discussed in detail in CBHSQ (2018).

6.2 QDU Weight Components

6.2.1 QDU Weight Component #12: Inverse of Selection Probability of at Least One Person in the Dwelling Unit

The selection of a QDU from all completed SDUs is based on the outcome of a variant of Brewer's method, which may select zero, one, or two people. Any pair of survey-eligible residents within the dwelling unit had some known, nonzero chance of being selected for the survey. The value for Weight Component #12 is equal to the inverse of the probability that at least one person in the dwelling unit is selected (see Section 2.2 for details).

6.2.2 QDU Weight Component #13: Selected QDU Poststratification to SDU-Based Control Totals

This poststratification factor adjusts the weights for selected QDUs to the SDU-based control totals. The SDU-based control totals are obtained by using the calibrated SDU weights. This adjustment step provides more stable controls for the subsequent nonresponse adjustment (Weight Component #14). Exhibit 4.1 lists the initially proposed variables for GEM modeling. The predictor variables are either 0/1 indicators or counting variables representing the number of people who fall into a given demographic domain. The counting variables are derived from the screener demographic information. It may be noted that during screening, the only required demographic information was the age of each person rostered. Thus, other demographic information necessary for weight calibration, such as race/ethnicity and gender, may be missing for certain rostered eligible people, and so imputation was done to replace these missing data. For more details on the imputation of screener demographic information, see CBHSQ (2019b).

The details on the predictor variables retained in the model and model summary statistics can be found in Appendix C.

6.2.3 QDU Weight Component #14: Respondent QDU Nonresponse Adjustment

This nonresponse adjustment step accounts for the failure to obtain respondent person(s) from each and every selected QDU. The same set of initially proposed predictor variables were used as for the previous adjustment (#13).

See Appendix C for more details on the predictor variables retained in the model and model summary statistics.

6.2.4 QDU Weight Component #15: Respondent QDU Poststratification to SDU-Based Control Totals

This final poststratification for all respondent QDUs utilized the same set of initially proposed predictor variables as previous adjustments. The corresponding control totals were obtained from the SDU-level sample, as was done for Weight Component #13.

See Appendix C for more details on the predictor variables retained in the model and model summary statistics.

6.2.5 QDU Weight Component #16: Respondent QDU Extreme Value Adjustment

The extreme weight proportions for the final poststratified weights were acceptably low, eliminating the need for the extreme value adjustment. Weight Component #16 was set to one for each responding QDU. This adjustment has not been used since this design was implemented for the 1999 NSDUH but is entered as a placeholder in the event that it may be required. For details on extreme weight proportions at each adjustment step, please see Appendix E.

6.3 Pair-Level Weight Components

Exhibit 4.2 lists the initially proposed predictor variables for the following adjustment steps via GEM.

6.3.1 Pair Weight Component #12: Inverse of Selection Probability of a Person Pair in the Dwelling Unit

Selection of pairs of individuals from all eligible people residing within the dwelling unit is based on the outcome of a variant of Brewer's method, which may select zero, one, or two people. Any pair of survey-eligible residents within the DU has some known, nonzero chance of being selected for the survey. When two people are selected, a pair is formed. The pair selection probability is determined by either formula (2.1) or formula (2.6) in Chapter 2. This weight component is the inverse of the selection probability discussed above.

6.3.2 Pair Weight Component #13: Selected Pair Poststratification to SDU-Based Control Totals

Similar to QDU Weight Component #13, this step was motivated by the consideration that the larger sample of all possible pairs provides more stable control totals for the respondent pair nonresponse adjustment. The weights of selected pairs were poststratified to the control totals that derived from calibrated SDU weights of all possible pairs. The pair-level demographic variables for all selected pairs, such as pair age group, pair race/ethnicity, and so on, were derived from screener demographic information.

The details on the predictor variables retained in the model and model summary statistics can be found in Appendix H.

6.3.3 Pair Weight Component #14: Respondent Pair Nonresponse Adjustment

If both people in the selected pair completed interviews successfully, the pair then was considered a respondent pair. This adjustment step accounts for failure to obtain respondent pairs from all selected pairs. In this step, respondent pair weights were adjusted to the control totals based on the full sample of selected pairs. Because of the low response rate of person pairs, this step had a relatively large adjustment on the weights. The same set of proposed predictor variables was used as for Weight Component #13. Similar to Weight Component #13, the pair-level demographic variables for all selected pairs, such as pair age group, pair race/ethnicity, and so on, were derived from screener demographic information.

See Appendix H for more details on the predictor variables retained in the model and model summary statistics.

6.3.4 Pair Weight Component #15: Respondent Pair Poststratification to SDU-Based Control Totals

This final poststratification utilized the same set of initially proposed predictor variables as previous adjustment steps. In addition, 10 pair relationship domain-level indicator variables were added to the set of covariates. The control totals for GEM calibration were derived from the SDU sample of all possible pairs of eligible people, as was done for Weight Component #13. The calibration control totals for these 10 domains used household-level person counts and the final QDU weights. As mentioned in the introduction, use of these household-level count totals for pair relationship domains in GEM calibration provided Hajek-type weight adjustment in the interest of obtaining more stable estimates. In setting up calibration covariates, multiplicity factors were needed. These factors, as discussed in the introduction, are used in constructing estimates for person-level parameters based on pair-related drug behavior. The factors depend on the pair domains of interest. For a selected set of pair domains, multiplicity factors are provided along with the pair-level analysis weights. See Chapter 11 in the NSDUH MRB editing and imputation report (CBHSQ, 2019a) for more detail on the creation of and imputation of missing values in the pair relationship, multiplicity, and household-level person counts. See Chapter 4 for more detail on the use of multiplicities and household-level person counts in poststratification.

Unlike Weight Components #13 and #14, demographic covariates were based on data from the questionnaire instead of information pulled from the dwelling unit screener.

For more details on the predictor variables retained in the GEM model and model summary statistics, see Appendix H.

6.3.5 Pair Weight Component #16: Respondent Pair Extreme Weight Adjustment

We checked the extreme weight proportions for the weights up to Weight Component #15, using the extreme weight domains (see Section 5.2). The built-in extreme weight control feature of GEM implemented in previous adjustment steps successfully reduced the extreme weight proportions. To be consistent with previous years, the extreme weight adjustment via GEM was implemented, using the same final set of predictor variables kept in the model for Weight Component #15. This step was successful in further reducing the extreme weight proportion in all model groups. For details, see Appendix J.

This page intentionally left blank

7. Evaluation of Calibration Weights

During the weight calibration process, several criteria for quality control were implemented to assess model adequacy. In this chapter, we describe the individual procedures and a summary of their results. All tables referred to in this chapter can be found in Appendices D through G and I through L.

7.1 Response Rates

Table D.1 in Appendix D displays the final selected and responding questionnaire dwelling unit (QDU) sample sizes from the 2017 National Survey on Drug Use and Health for various national domains. This table also shows the weighted response rates.[10] Most domains reflect the overall 70.77 percent response rate, with most rates between 70 and 75 percent, although the highest response rate is 92.62 percent, from the Group category of the Group Quarters variable. The lowest response rate came from the 5th quintile of the Segment-Combined Median Rent and Housing value variable, with 67.26 percent, although that was likely influenced by a small sample for the category.

Table I.1 in Appendix I displays the final selected and responding pair-level sample sizes and weighted response rates from the 2017 survey for various national domains. Because of the nature of the pair data, the response rates were lower in all domains examined than at the QDU level, with an overall response rate of 54.9 percent. The response rates range from a low of 38.97 percent in the Pair Age Group category of 35–49, 50+ to a high of 74.61 percent from the Pair Age Group category of 12–17, 12–17. This extreme range of response rates is probably due to a combination of small sample sizes and response burden as a result of selection of pairs within households among various domains. Like at the QDU level, the top response rates are among the younger respondents (as measured by household type for the QDU data and pair age for the pair data). This pattern may be related to the relatively high response rates in the group level of the variable group quarters because it includes college dormitories.

7.2 Proportions of Extreme Values and Outwinsors

During the stages of modeling adjustments (i.e., nonresponse [nr] and poststratification [ps]), one major issue of concern when deciding the adequacy of a particular model was the extent of the resulting proportions of extreme value (ev) and outwinsor weights (see Sections 5.1 and 5.2 for these definitions). For each weight adjustment step, these proportions are computed before and after the step for various domains. Prior to adjustment, the product of all weight components is used to compute proportions of evs and outwinsors, while after the adjustment, the product includes the new adjustment factor. If the proportion of evs and outwinsors is deemed high (normally 3 percent of unweighted, 15 percent weighted, and 5 percent of outwinsor), a separate ev treatment step after ps could be performed. Although this threshold was not met in the 2017 data, this step has been implemented for pair-level weighting to reduce final

[10] Questionnaire dwelling unit response rates and pair response rates were computed using American Association for Public Opinion Research (AAPOR)'s Response Rate 2. See AAPOR's Standard Definitions (AAPOR, 2016) for more information.

ev and outwinsor proportions and to maintain year-to-year consistency. This was done for the pair-level weights. Details of this step are explained in Section 6.3.5. A separate ev treatment step was deemed unnecessary for the QDU-level weights.

Tables E.1 and E.2 and Tables J.1 through J.3 present percentages of evs at the QDU level and the pair level, respectively, for various domains. Unweighted percentages are the percentage of actual counts of units defined as evs relative to the total sample size. Weighted percentages reflect the percentage of total ev weights relative to the total sample weight, while outwinsor percentages represent the total amount of residual weight when the weights are trimmed to the critical values (used for ev definition) relative to the total sample weight. For evaluation purposes, the outwinsor percentage is considered the most important of the three percentages, as this gave a measure of the impact of winsorization (or trimming) of ev weights (if we performed this treatment). See Sections 5.1 and 5.2 for the domains that were used to define extreme values.

7.3 Slippage Rates

The slippage rate for a given domain is defined as the relative percentage difference between the sampling weights and the external control totals, both before and after ps. The control totals for QDU and person pair ps are derived from the screener dwelling unit weights, which were poststratified to U.S. Census Bureau population estimates (Center for Behavioral Health Statistics and Quality, 2019b). Table F.1 displays QDU national domain-specific weight sums for both before and after ps, as well as the desired totals to be met through ps. Table K.1 shows the same for the pair sample. These tables also show the relative percentage difference, or the amount of adjustment necessary (positive or negative) to meet the desired totals. The first relative difference is used explicitly during the ps modeling procedure to identify potential problems for convergence. Large differences in domains with relatively small sample sizes are indicative of potential large adjustment factors, which may cause problems in convergence while satisfying bound constraints. The reason is that adjustments required for one domain may have an adverse effect on another domain when a unit belongs to both.

As an example, consider that Table F.1, for the 2017 QDU domain household size of four or more, indicates a sample size of 10,617 with a total design-based weight of 14,848,214 and a census total of 14,848,202 with an initial slippage rate of 1.64 percent, which would imply a common weight adjustment approximately equal to 0.983832, if this were the only calibration control. Similarly, looking at pair data in Table K.1, the pair domain category of Pair Age Group 12–17, 26–34 has a sample size of 894, a design-based weight of 4,854,885, and a census total of 4,878,622, showing an initial slippage of -0.49 percent. The resultant required adjustment would be approximately equal to 1.004889, if this were the only control. However, in the generalized exponential model (GEM), all controls are simultaneously satisfied under a complex algorithm that allows for different adjustment factors for different units.

7.4 Weight Adjustment Summary Statistics

Tables G.1, G.2, and L.1 through L.3 display summary statistics on the product of weight components before and after all stages of adjustment for the QDU and person pair, respectively. The summary statistics include sample size (*n*), minimum (min), maximum (max), median

(med), 25th percentile (Q1), 75th percentile (Q3), and the unequal weighting effect (UWE). Note that in Tables L.2 and L.3, the sample size for pair age group, pair race/ethnicity, and pair gender are slightly different. This is because those variables were defined using screening demographic information in the nonresponse adjustment of respondent pairs, while in the poststratification of respondent pairs, they were defined from questionnaire demographic information. Because UWE is directly affected by weight adjustment factors and extreme weights, these values—along with the percentage of extreme weights as noted in Section 7.2—were used as guidelines for determining model adequacy.

7.5 Sensitivity Analysis of Drug Use Estimates

It is known that, in general, there is a trade-off between bias reduction and variance reduction. For instance, with GEM (for nr or ps), enlarging a simple model (such as the one with only main effects) has the potential of further reducing the bias. At the same time, this enlargement also may be associated with a corresponding increase in the variance of the estimate due to additional variability caused by estimating the model parameters. To check for possible overfitting of the GEM model, we conducted a sensitivity analysis for respondent QDU poststratification for the QDU weights, respondent pair poststratification, and extreme weight adjustment for the person pair weights. A simple baseline model was fitted with the same bounds and maximum number of iterations as was used for the chosen (more complex) final model. We then looked for substantial changes in point estimates and standard errors (SEs). For the QDU weights, some household-level characteristics were selected such as family income, number of youths in the household, whether the household had health insurance coverage, and number of elders living in the household. The estimates and SEs are displayed in Table 7.1. For the person pair weights, selected licit and illicit drug use prevalence rates of 12- to 17-year-olds were calculated from parent-child pairs, and estimates and SEs of the estimates based on pair weights are shown in Tables 7.2a to 7.2b.

As seen in Table 7.1, the estimates and their SEs for the two models (baseline and the final) are generally similar to each other for the QDU weights. However, among the person pair estimates and SEs, there are some differences, but they do not seem significant in general.

Because the sensitivity analyses for both QDU- and pair-level calibrated weights seem to indicate that adding more covariates does not introduce an undesirable degree of instability in the estimates or their SEs, the final, more complex GEM models were deemed reasonable.

Table 7.1 Estimates of Totals and SEs for Domains of Interest Based on QDU Sample: 2017

Domain	*n*	Baseline (B)[1]	Final (F)[2]	(B-F)/F% (Estimate)	(B-F)/F% (SE)
Households with Family Income					
$0–<$10,000	4,015	9,196,430 (291,591)	9,203,915 (292,318)	-0.08	-0.25
$10,000–<$20,000	5,545	14,876,068 (354,593)	14,866,553 (354,794)	0.06	-0.06
$20,000–<$30,000	5,258	13,581,002 (305,506)	13,583,627 (306,002)	-0.02	-0.16
$30,000–<$40,000	4,966	12,574,278 (292,983)	12,565,359 (293,131)	0.07	-0.05
$40,000–<$50,000	5,014	12,189,323 (282,780)	12,195,963 (283,239)	-0.05	-0.16
$50,000–<$75,000	7,834	19,627,173 (356,061)	19,620,140 (356,020)	0.04	0.01
$75,000–<$100,000	5,784	13,859,099 (287,627)	13,863,689 (287,821)	-0.03	-0.07
$100,000+	11,912	29,151,310 (551,958)	29,155,438 (551,907)	-0.01	0.01
Households with Number of Youths (<18)					
0	23,776	82,991,241 (945,425)	82,992,446 (945,427)	-0.00	-0.00
1	10,698	17,885,290 (235,622)	17,881,508 (235,714)	0.02	-0.04
2	9,330	14,852,033 (234,611)	14,851,992 (235,511)	0.00	-0.38
3	4,224	6,322,679 (136,843)	6,327,640 (137,725)	-0.08	-0.64
4+	2,300	3,003,441 (87,626)	3,001,097 (87,713)	0.08	-0.10
Households with Insurance Coverage					
Yes	45,581	114,219,168 (1,028,107)	114,220,068 (1,028,554)	-0.00	-0.04
No	4,747	10,835,516 (243,754)	10,834,615 (244,428)	0.01	-0.28
Households with Number of Older Adults (65+)					
0	42,761	90,294,013 (779,619)	90,294,282 (780,113)	-0.00	-0.06
1	4,817	22,159,843 (491,716)	22,162,040 (491,373)	-0.01	0.07
2	2,687	12,354,871 (332,603)	12,352,920 (332,621)	0.02	-0.01
3+	63	245,957 (41,418)	245,442 (41,290)	0.21	0.31

QDU = questionnaire dwelling unit; SE = standard error.
Note: Standard errors of prevalence estimates are provided in parentheses.
[1] Baseline refers to the weight obtained from using a main effects only model for the last step of calibration, res.qdu.ps, and a full model for preceding steps.
[2] Final refers to the weight obtained using a full model throughout all steps of calibration.

Table 7.2a Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Alcohol and Tobacco among Mother-Child (12 to 17) Pairs, by Mother Use: 2017

Drug	Mother User	*n*	Baseline[1]	Final[2]
Alcohol				
Lifetime	Yes	2,685	26.86 (1.48)	26.92 (1.47)
	No	439	13.77 (2.59)	14.31 (2.69)
	Overall	3,124	24.56 (1.33)	24.75 (1.34)
Past Year	Yes	2,259	21.68 (1.42)	21.89 (1.44)
	No	865	11.81 (1.66)	11.95 (1.70)
	Overall	3,124	18.62 (1.14)	18.84 (1.16)
Past Month	Yes	1,711	8.61 (1.11)	8.83 (1.16)
	No	1,413	6.59 (0.99)	6.73 (1.02)
	Overall	3,124	7.65 (0.78)	7.84 (0.81)
Cigarettes				
Lifetime	Yes	1,903	12.24 (1.24)	12.37 (1.27)
	No	1,221	5.43 (1.00)	5.58 (1.04)
	Overall	3,124	9.20 (0.81)	9.37 (0.83)
Past Year	Yes	711	11.38 (1.96)	11.45 (2.03)
	No	2,413	3.27 (0.51)	3.36 (0.53)
	Overall	3,124	4.74 (0.56)	4.84 (0.58)
Past Month	Yes	634	7.72 (1.90)	7.85 (1.98)
	No	2,490	1.07 (0.21)	1.05 (0.20)
	Overall	3,124	2.12 (0.37)	2.14 (0.38)

Note: Standard errors of prevalence estimates are provided in parentheses.
[1] Baseline refers to the weight obtained from using a main effects only model for the last two steps of calibration, res.pr.ps and res.pr.ev, and a full model for preceding steps.
[2] Final refers to the weight obtained using a full model throughout all steps of calibration.

Table 7.2b Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Alcohol and Tobacco among Father-Child (12 to 17) Pairs, by Father Use: 2017

Drug	Father User	*n*	Baseline[1]	Final[2]
Alcohol				
Lifetime	Yes	1,785	24.42 (1.81)	24.44 (1.78)
	No	140	16.60 (4.93)	16.33 (4.85)
	Overall	1,925	23.76 (1.72)	23.74 (1.70)
Past Year	Yes	1,515	21.11 (1.77)	21.14 (1.76)
	No	410	14.90 (3.17)	14.29 (2.94)
	Overall	1,925	19.64 (1.54)	19.52 (1.51)
Past Month	Yes	1,259	8.75 (1.35)	8.62 (1.35)
	No	666	6.50 (1.49)	6.62 (1.50)
	Overall	1,925	7.92 (0.99)	7.89 (1.00)
Cigarettes				
Lifetime	Yes	1,383	9.24 (1.35)	9.19 (1.28)
	No	542	2.20 (0.70)	2.18 (0.69)
	Overall	1,925	6.96 (0.96)	6.93 (0.91)
Past Year	Yes	494	7.23 (1.87)	7.30 (1.84)
	No	1,431	3.04 (0.69)	3.09 (0.70)
	Overall	1,925	3.92 (0.67)	3.98 (0.67)
Past Month	Yes	433	3.07 (1.45)	3.06 (1.39)
	No	1,492	1.29 (0.36)	1.34 (0.39)
	Overall	1,925	1.61 (0.40)	1.66 (0.41)

Note: Standard errors of prevalence estimates are provided in parentheses.
[1] Baseline refers to the weight obtained from using a main effects only model for the last two steps of calibration, res.pr.ps and res.pr.ev, and a full model for preceding steps.
[2] Final refers to the weight obtained using a full model throughout all steps of calibration.

Table 7.3a Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Any Illicit Drug or Marijuana among Mother-Child (12 to 17) Pairs, by Mother Use: 2017

Drug	Mother User	*n*	Baseline[1]	Final[2]
Any Illicit				
Lifetime	Yes	1,653	25.52 (1.94)	25.65 (1.92)
	No	1,471	17.32 (1.61)	17.51 (1.64)
	Overall	3,124	21.33 (1.26)	21.54 (1.27)
Past Year	Yes	444	24.39 (3.99)	23.90 (3.76)
	No	2,680	12.53 (1.02)	12.80 (1.05)
	Overall	3,124	13.98 (1.07)	14.17 (1.06)
Past Month	Yes	239	16.02 (4.03)	16.07 (4.04)
	No	2,885	5.72 (0.60)	5.84 (0.61)
	Overall	3,124	6.33 (0.59)	6.45 (0.60)
Marijuana				
Lifetime	Yes	1,513	18.41 (1.70)	18.67 (1.73)
	No	1,611	9.54 (1.26)	9.64 (1.24)
	Overall	3,124	13.41 (1.06)	13.64 (1.07)
Past Year	Yes	317	22.24 (3.98)	22.13 (3.99)
	No	2,807	9.19 (0.92)	9.37 (0.92)
	Overall	3,124	10.21 (0.93)	10.39 (0.93)
Past Month	Yes	193	15.05 (4.32)	15.09 (4.47)
	No	2,931	4.78 (0.55)	4.87 (0.56)
	Overall	3,124	5.22 (0.55)	5.32 (0.57)

Note: Standard errors of prevalence estimates are provided in parentheses.
[1] Baseline refers to the weight obtained from using a main effects only model for the last two steps of calibration, res.pr.ps and res.pr.ev, and a full model for preceding steps.
[2] Final refers to the weight obtained using a full model throughout all steps of calibration.

Table 7.3b Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Any Illicit Drug or Marijuana among Father-Child (12 to 17) Pairs, by Father Use: 2017

Drug	Father User	*n*	Baseline[1]	Final[2]
Any Illicit				
Lifetime	Yes	1,124	23.12 (2.16)	23.01 (2.14)
	No	801	14.65 (2.04)	14.78 (2.04)
	Overall	1,925	19.25 (1.54)	19.30 (1.53)
Past Year	Yes	325	21.16 (3.88)	21.54 (3.90)
	No	1,600	11.06 (1.36)	10.91 (1.33)
	Overall	1,925	12.60 (1.31)	12.57 (1.31)
Past Month	Yes	186	14.08 (4.37)	14.95 (4.73)
	No	1,739	4.69 (0.77)	4.64 (0.74)
	Overall	1,925	5.46 (0.85)	5.49 (0.86)
Marijuana				
Lifetime	Yes	1,036	16.37 (2.15)	16.18 (2.11)
	No	889	8.41 (1.64)	8.33 (1.59)
	Overall	1,925	12.30 (1.38)	12.21 (1.35)
Past Year	Yes	248	18.46 (4.51)	18.56 (4.47)
	No	1,677	8.99 (1.29)	8.89 (1.28)
	Overall	1,925	10.03 (1.24)	9.97 (1.23)
Past Month	Yes	160	11.08 (3.68)	11.80 (3.90)
	No	1,765	4.44 (0.87)	4.41 (0.88)
	Overall	1,925	4.88 (0.84)	4.90 (0.86)

Note: Standard errors of prevalence estimates are provided in parentheses.
[1] Baseline refers to the weight obtained from using a main effects only model for the last two steps of calibration, res.pr.ps and res.pr.ev, and a full model for preceding steps.
[2] Final refers to the weight obtained using a full model throughout all steps of calibration.

Table 7.4 Percentages of Youths (12 to 17) Living with a Parent Reporting Lifetime, Past Year, and Past Month Use of Alcohol and Tobacco among Parent-Child (12 to 17) Pairs, Asked Whether Their Parents Had Spoken to Them about the Dangers of Tobacco, Alcohol, or Drug Use within the Past 12 Months: 2017

Drug	Parent Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Alcohol				
Lifetime	Yes	2,849	24.98 (1.43)	25.04 (1.45)
	No	2,100	25.64 (1.88)	25.80 (1.85)
	Overall	4,949	25.28 (1.19)	25.38 (1.19)
Past Year	Yes	2,849	20.85 (1.35)	20.87 (1.37)
	No	2,100	17.29 (1.23)	17.50 (1.25)
	Overall	4,949	19.26 (0.95)	19.36 (0.96)
Past Month	Yes	2,849	9.09 (1.03)	9.14 (1.06)
	No	2,100	6.70 (0.77)	6.85 (0.78)
	Overall	4,949	8.02 (0.67)	8.11 (0.69)
Cigarettes				
Lifetime	Yes	2,849	8.39 (0.95)	8.48 (0.97)
	No	2,100	9.57 (1.00)	9.67 (1.00)
	Overall	4,949	8.92 (0.70)	9.01 (0.71)
Past Year	Yes	2,849	5.06 (0.80)	5.18 (0.84)
	No	2,100	4.36 (0.59)	4.44 (0.60)
	Overall	4,949	4.74 (0.51)	4.85 (0.54)
Past Month	Yes	2,849	2.29 (0.59)	2.36 (0.63)
	No	2,100	2.05 (0.35)	2.10 (0.37)
	Overall	4,949	2.18 (0.36)	2.24 (0.39)

Note: Standard errors of prevalence estimates are provided in parentheses.

[1] Baseline refers to the weight obtained from using a main effects only model for the last two steps of calibration, res.pr.ps and res.pr.ev, and a full model for preceding steps.

[2] Final refers to the weight obtained using a full model throughout all steps of calibration.

Table 7.5 Percentages of Youths (12 to 17) Living with a Parent Reporting Lifetime, Past Year, and Past Month Use of Any Illicit Drug and Marijuana among Parent-Child (12 to 17) Pairs, Asked Whether Their Parents Had Spoken to Them about the Dangers of Tobacco, Alcohol, or Drug Use within the Past 12 Months: 2017

Drug	Parent Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Any Illicit				
Lifetime	Yes	2,849	19.06 (1.28)	19.25 (1.31)
	No	2,100	24.92 (1.83)	24.94 (1.79)
	Overall	4,949	21.69 (1.11)	21.80 (1.11)
Past Year	Yes	2,849	12.93 (1.17)	13.07 (1.19)
	No	2,100	16.46 (1.66)	16.47 (1.60)
	Overall	4,949	14.51 (1.00)	14.59 (1.00)
Past Month	Yes	2,849	6.44 (0.87)	6.53 (0.89)
	No	2,100	6.55 (0.73)	6.62 (0.75)
	Overall	4,949	6.49 (0.57)	6.57 (0.59)
Marijuana				
Lifetime	Yes	2,849	12.48 (1.18)	12.64 (1.21)
	No	2,100	16.45 (1.64)	16.41 (1.59)
	Overall	4,949	14.26 (1.00)	14.33 (0.99)
Past Year	Yes	2,849	10.51 (1.14)	10.65 (1.17)
	No	2,100	12.20 (1.48)	12.14 (1.42)
	Overall	4,949	11.27 (0.92)	11.32 (0.91)
Past Month	Yes	2,849	5.78 (0.86)	5.87 (0.89)
	No	2,100	5.43 (0.68)	5.49 (0.69)
	Overall	4,949	5.62 (0.56)	5.70 (0.58)

Note: Standard errors of prevalence estimates are provided in parentheses.

[1] Baseline refers to the weight obtained from using a main effects only model for the last two steps of calibration, res.pr.ps and res.pr.ev, and a full model for preceding steps.

[2] Final refers to the weight obtained using a full model throughout all steps of calibration.

Table 7.6a Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Alcohol and Tobacco among Mother-Child (12 to 17) Pairs, for Mother in the Pair, Asked Whether She Had Spoken to Her Children about the Dangers of Tobacco, Alcohol, or Drug Use within the Past 12 Months: 2017

Drug	Mother Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Alcohol				
Lifetime	0 times	230	19.98 (4.21)	20.29 (4.28)
	1–2 times	561	17.19 (2.87)	17.35 (2.88)
	A few times	755	25.15 (2.85)	25.28 (2.74)
	Many times	1,431	28.12 (2.00)	28.22 (2.03)
	Overall	2,977	24.48 (1.36)	24.64 (1.36)
Past Year	0 times	230	15.17 (3.92)	15.20 (3.96)
	1–2 times	561	11.32 (2.39)	11.51 (2.39)
	A few times	755	16.39 (2.05)	16.75 (2.05)
	Many times	1,431	23.19 (1.87)	23.34 (1.91)
	Overall	2,977	18.51 (1.17)	18.73 (1.19)
Past Month	0 times	230	6.90 (3.12)	6.99 (3.19)
	1–2 times	561	2.17 (0.66)	2.27 (0.71)
	A few times	755	8.17 (1.68)	8.49 (1.68)
	Many times	1,431	9.98 (1.28)	10.17 (1.35)
	Overall	2,977	7.68 (0.81)	7.87 (0.84)
Cigarettes				
Lifetime	0 times	230	3.39 (1.29)	3.30 (1.22)
	1–2 times	561	3.26 (0.89)	3.18 (0.84)
	A few times	755	10.42 (1.87)	10.80 (1.91)
	Many times	1,431	12.26 (1.44)	12.45 (1.49)
	Overall	2,977	9.23 (0.84)	9.40 (0.87)
Past Year	0 times	230	1.38 (0.93)	1.18 (0.78)
	1–2 times	561	1.57 (0.58)	1.59 (0.58)
	A few times	755	3.45 (1.17)	3.65 (1.22)
	Many times	1,431	7.30 (1.04)	7.42 (1.09)
	Overall	2,977	4.75 (0.58)	4.86 (0.61)
Past Month	0 times	230	0.81 (0.76)	0.67 (0.61)
	1–2 times	561	1.13 (0.55)	1.16 (0.56)
	A few times	755	0.58 (0.37)	0.62 (0.39)
	Many times	1,431	3.47 (0.71)	3.48 (0.74)
	Overall	2,977	2.11 (0.38)	2.13 (0.39)

Note: Standard errors of prevalence estimates are provided in parentheses.

[1] Baseline refers to the weight obtained from using a main effects only model for the last two steps of calibration, res.pr.ps and res.pr.ev, and a full model for preceding steps.

[2] Final refers to the weight obtained using a full model throughout all steps of calibration.

Table 7.6b Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Alcohol and Tobacco among Father-Child (12 to 17) Pairs, for Father in the Pair, Asked Whether He Had Spoken to His Child about the Dangers of Tobacco, Alcohol, or Drug Use within the Past 12 Months: 2017

Drug	Father Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Alcohol				
Lifetime	0 times	229	16.26 (3.86)	15.28 (3.45)
	1–2 times	386	19.01 (2.90)	19.17 (2.95)
	A few times	573	23.01 (2.97)	22.64 (2.87)
	Many times	590	29.04 (3.25)	29.50 (3.25)
	Overall	1,778	23.14 (1.73)	23.09 (1.71)
Past Year	0 times	229	13.97 (3.64)	12.96 (3.24)
	1–2 times	386	16.56 (2.79)	16.76 (2.84)
	A few times	573	19.53 (2.81)	19.18 (2.71)
	Many times	590	21.73 (2.57)	21.81 (2.54)
	Overall	1,778	18.84 (1.51)	18.67 (1.49)
Past Month	0 times	229	6.03 (2.69)	4.64 (1.96)
	1–2 times	386	5.90 (1.77)	5.91 (1.84)
	A few times	573	8.52 (1.90)	8.68 (1.92)
	Many times	590	9.25 (1.63)	9.24 (1.63)
	Overall	1,778	7.86 (0.99)	7.74 (0.97)
Cigarettes				
Lifetime	0 times	229	4.86 (2.50)	4.66 (2.38)
	1–2 times	386	3.01 (0.94)	2.98 (0.93)
	A few times	573	6.95 (1.85)	6.91 (1.78)
	Many times	590	7.98 (1.59)	8.06 (1.57)
	Overall	1,778	6.17 (0.89)	6.15 (0.87)
Past Year	0 times	229	0.74 (0.39)	0.76 (0.39)
	1–2 times	386	1.06 (0.51)	1.00 (0.47)
	A few times	573	4.34 (1.44)	4.43 (1.42)
	Many times	590	5.49 (1.44)	5.40 (1.40)
	Overall	1,778	3.52 (0.68)	3.53 (0.67)
Past Month	0 times	229	0.43 (0.23)	0.44 (0.24)
	1–2 times	386	0.23 (0.13)	0.22 (0.13)
	A few times	573	1.06 (0.43)	1.12 (0.44)
	Many times	590	2.78 (1.08)	2.73 (1.04)
	Overall	1,778	1.34 (0.37)	1.35 (0.37)

Note: Standard errors of prevalence estimates are provided in parentheses.

[1] Baseline refers to the weight obtained from using a main effects only model for the last two steps of calibration, res.pr.ps and res.pr.ev, and a full model for preceding steps.

[2] Final refers to the weight obtained using a full model throughout all steps of calibration.

Table 7.7a Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Any Illicit Drug and Marijuana among Mother-Child (12 to 17) Pairs, for Mother in the Pair, Asked Whether She Had Spoken to Her Child about the Dangers of Tobacco, Alcohol, or Drug Use within the Past 12 Months: 2017

Drug	Mother Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Any Illicit				
Lifetime	0 times	230	16.99 (4.06)	17.14 (4.14)
	1–2 times	561	14.33 (2.23)	14.27 (2.23)
	A few times	755	23.24 (3.01)	23.21 (2.90)
	Many times	1,431	24.39 (1.88)	24.77 (1.92)
	Overall	2,977	21.41 (1.31)	21.60 (1.32)
Past Year	0 times	230	12.47 (3.57)	12.66 (3.64)
	1–2 times	561	8.34 (1.75)	8.57 (1.82)
	A few times	755	16.57 (2.91)	16.46 (2.75)
	Many times	1,431	15.54 (1.41)	15.86 (1.45)
	Overall	2,977	14.00 (1.10)	14.20 (1.10)
Past Month	0 times	230	3.98 (1.47)	4.11 (1.51)
	1–2 times	561	3.54 (1.01)	3.58 (1.01)
	A few times	755	5.26 (1.34)	5.44 (1.32)
	Many times	1,431	8.34 (1.05)	8.46 (1.08)
	Overall	2,977	6.28 (0.61)	6.41 (0.62)
Marijuana				
Lifetime	0 times	230	7.82 (3.08)	7.88 (3.13)
	1–2 times	561	7.48 (1.75)	7.78 (1.83)
	A few times	755	13.98 (2.62)	14.04 (2.50)
	Many times	1,431	16.73 (1.57)	16.99 (1.60)
	Overall	2,977	13.44 (1.11)	13.66 (1.11)
Past Year	0 times	230	7.42 (3.08)	7.49 (3.13)
	1–2 times	561	4.77 (1.21)	5.01 (1.29)
	A few times	755	10.99 (2.51)	10.97 (2.35)
	Many times	1,431	12.77 (1.31)	13.01 (1.35)
	Overall	2,977	10.25 (0.96)	10.43 (0.96)
Past Month	0 times	230	3.01 (1.35)	3.09 (1.39)
	1–2 times	561	2.68 (0.92)	2.78 (0.93)
	A few times	755	4.48 (1.31)	4.60 (1.28)
	Many times	1,431	6.92 (0.96)	7.01 (0.98)
	Overall	2,977	5.16 (0.57)	5.27 (0.59)

Note: Standard errors of prevalence estimates are provided in parentheses.

[1] Baseline refers to the weight obtained from using a main effects only model for the last two steps of calibration, res.pr.ps and res.pr.ev, and a full model for preceding steps.

[2] Final refers to the weight obtained using a full model throughout all steps of calibration.

Table 7.7b Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Any Illicit Drug and Marijuana among Father-Child (12 to 17) Pairs, for Father in the Pair, Asked Whether He Had Spoken to His Child about the Dangers of Tobacco, Alcohol, or Drug Use within the Past 12 Months: 2017

Drug	Father Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Any Illicit				
Lifetime	0 times	229	13.45 (3.36)	13.86 (3.57)
	1–2 times	386	11.67 (2.12)	11.63 (2.12)
	A few times	573	22.37 (3.14)	22.58 (3.12)
	Many times	590	20.36 (2.31)	20.38 (2.30)
	Overall	1,778	18.27 (1.47)	18.43 (1.48)
Past Year	0 times	229	10.41 (3.20)	10.91 (3.43)
	1–2 times	386	6.49 (1.64)	6.44 (1.63)
	A few times	573	14.80 (2.68)	14.74 (2.64)
	Many times	590	13.35 (1.93)	13.15 (1.90)
	Overall	1,778	12.01 (1.25)	11.99 (1.25)
Past Month	0 times	229	4.95 (2.49)	4.98 (2.44)
	1–2 times	386	1.56 (0.68)	1.52 (0.66)
	A few times	573	6.92 (1.73)	6.98 (1.71)
	Many times	590	5.28 (1.29)	5.23 (1.26)
	Overall	1,778	5.02 (0.81)	5.02 (0.80)
Marijuana				
Lifetime	0 times	229	8.21 (3.11)	8.83 (3.37)
	1–2 times	386	6.19 (1.71)	6.02 (1.67)
	A few times	573	14.27 (2.72)	14.37 (2.69)
	Many times	590	12.27 (1.91)	12.10 (1.87)
	Overall	1,778	11.12 (1.26)	11.17 (1.26)
Past Year	0 times	229	7.95 (3.11)	8.57 (3.37)
	1–2 times	386	4.30 (1.50)	4.07 (1.44)
	A few times	573	11.74 (2.49)	11.79 (2.46)
	Many times	590	10.53 (1.79)	10.25 (1.74)
	Overall	1,778	9.28 (1.17)	9.26 (1.17)
Past Month	0 times	229	4.05 (2.45)	4.11 (2.40)
	1–2 times	386	1.15 (0.63)	1.10 (0.60)
	A few times	573	6.24 (1.73)	6.26 (1.71)
	Many times	590	4.68 (1.26)	4.64 (1.23)
	Overall	1,778	4.39 (0.80)	4.39 (0.79)

Note: Standard errors of prevalence estimates are provided in parentheses.

[1] Baseline refers to the weight obtained from using a main effects only model for the last two steps of calibration, res.pr.ps and res.pr.ev, and a full model for preceding steps.

[2] Final refers to the weight obtained using a full model throughout all steps of calibration.

References

American Association for Public Opinion Research. (2016.) *Standard definitions: Final dispositions of case codes and outcome rates for surveys (9th edition)*. AAPOR. Retrieved from https://www.aapor.org/Education-Resources/For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview.aspx

Center for Behavioral Health Statistics and Quality. (2017). *How to prepare and analyze pair data in the National Survey on Drug Use and Health.* Rockville, MD: Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (2018). *2017 National Survey on Drug Use and Health: Methodological resource book (Section 2, Sample design report)*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (2019a). *2017 National Survey on Drug Use and Health: Methodological resource book (Section 10, Editing and imputation report*). Rockville, MD: Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (2019b). *2017 National Survey on Drug Use and Health: Methodological resource book (Section 11, Person-level sampling weight calibration)*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Chen, P., Dai, L., Gordek, H., Grau, E., Shi, W., & Westlake, M. (2006). Person-level sampling weight calibration. In *2004 National Survey on Drug Use and Health: Methodological resource book* (Section 12, prepared for the Substance Abuse and Mental Health Services Administration, Contract No. 283-03-9028, Deliverable No. 26, RTI/0208726.174.003). Research Triangle Park, NC: RTI International.

Chen, P., Penne, M. A., & Singh, A. C. (2000). Experience with the generalized exponential model (GEM) for weight calibration for NHSDA. In *Proceedings of the 2000 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Indianapolis, IN* (pp. 604–607). Alexandria, VA: American Statistical Association. Retrieved from https://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. Retrieved from https://www.amstat.org/ASA/Membership/Sections-and-Interest-Groups.aspx

Deville, J. C., & Särndal, C. E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association, 87*(418), 376–382.

Folsom, R. E., & Singh, A. C. (2000). The generalized exponential model for sampling weight calibration for extreme values, nonresponse, and poststratification. In *Proceedings of the 2000 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Indianapolis, IN* (pp. 598–603). Alexandria, VA: American Statistical Association. Retrieved from https://www.amstat.org/ASA/Membership/Sections-and-Interest-Groups.aspx

Folsom, R. E., & Witt, M. B. (1994). Testing a new attrition nonresponse adjustment method for SIPP. In *Proceedings of the 1994 Joint Statistical Meetings, American Statistical Association, Social Statistics Section, Toronto, Ontario, Canada* (pp. 428–433). Alexandria, VA: American Statistical Association.

Hajek, J. (1971). Comment on D. Basu's paper, "An essay on the logical foundations of survey sampling, part one" [pp. 203–234]. In V. P. Godambe, & D. A. Sprott (Eds.), *Foundations of statistical inference: A symposium* (p. 236). Toronto, Ontario, & Montreal, Quebec, Canada: Holt, Rinehart and Winston of Canada.

Penne, M., Chen, P., & Singh, A. C. (2001). Person-pair sampling weight calibration using the generalized exponential model for the National Household Survey on Drug Abuse. In *Proceedings of the 2001 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Atlanta, GA* [CD-ROM]. Alexandria, VA: American Statistical Association. Retrieved from https://www.amstat.org/ASA/Membership/Sections-and-Interest-Groups.aspx

Singh, A. C., & Mohl, C. A. (1996). Understanding calibration estimators in survey sampling. *Survey Methodology, 22*, 107–115.

List of Contributors

This methodological report was prepared by the Substance Abuse and Mental Health Services Administration (SAMHSA), Center for Behavioral Health Statistics and Quality (CBHSQ), and RTI International (a registered trademark and a trade name of Research Triangle Institute). Work by RTI was performed under Contract No. HHSS283201300001C. Peter Tice served as the Government Project Officer, and David Hunter served as the RTI Project Director.

Significant contributors to this report at SAMHSA include Art Hughes (retired from SAMHSA) and Matthew Williams (formerly with SAMHSA).

Significant contributors to this report at RTI include Debbie Bond, Patrick Chen, Claudia Clark, Ralph Folsom (retired from RTI), Harper Gordek, Margaret Johnson, Jeff Laufenberg, and Matthew Westlake.

This page intentionally left blank

Appendix A: Technical Details about the Generalized Exponential Model

This page intentionally left blank

Appendix A: Technical Details about the Generalized Exponential Model

A.1 Distance Function

Let $\Delta(w,d)$ denote the distance between the initial weights $d = \{d_k : k \in s\}$ and the adjusted weights *w*, with *k* being the k^{th} unit in the sample and *s* being the sample selected. The distance function minimized under the generalized exponential model (GEM), subject to calibration constraints, is given by

$$\Delta(w,d) = \sum_{k\in s} \frac{d_k}{A_k}\left\{(a_k - \ell_k)\log\frac{a_k - \ell_k}{c_k - \ell_k} + (u_k - a_k)\log\frac{u_k - a_k}{u_k - c_k}\right\}, \tag{A.1.1}$$

where $a_k = w_k / d_k, A_k = (u_k - \ell_k)/[(u_k - c_k)(c_k - \ell_k)]$ and ℓ_k, c_k, and u_k are prescribed real numbers. Let T_x denote the *p*-vector of control totals corresponding to predictor variables ($x_1, ..., x_p$). Then, the calibration constraints for the above minimization problem are

$$\sum_{k\in s} x_k d_k a_k = T_x. \tag{A.1.2}$$

The solution for the above minimization problem, if it exists, is given by a GEM with model parameters λ; that is,

$$a_k(\lambda) = \frac{\ell_k(u_k - c_k) + u_k(c_k - \ell_k)\exp\{A_k x_k'\lambda\}}{(u_k - c_k) + (c_k - \ell_k)\exp\{A_k x_k'\lambda\}}. \tag{A.1.3}$$

Note that the number of parameters in the GEM should be $\leq n$, where *n* is the size of the sample *s*. This is also the dimension of vectors *d* and *w*. It follows from equation A.1.3 that

$$\ell_k < a_k < u_k, k = 1,\ldots,n. \tag{A.1.4}$$

The weight adjustment factor achieved by the usual raking ratio algorithm (Singh & Mohl, 1996) can also be derived as a special case of the GEM, noting that for $\ell_k = 0, u_k = \infty, c_k = 1,$ and $k = 1,\ldots,n$, we have

$$\Delta(w,d) = \sum_{k\in s} d_k a_k \log a_k - \sum_{k\in s} d_k (a_k - 1) \tag{A.1.5}$$

and $a_k(\lambda) = \exp(x_k'\lambda)$.

The logit model of Deville and Särndal (1992) is also a special case of the GEM, by setting $\ell_k = \ell,\ u_k = u,$ and $c_k = 1$ for all *k*. The new method was introduced by Folsom and Singh (2000).

A.2 GEM Adjustments for Extreme Value Treatment, Nonresponse, and Poststratification

By choosing the user-specified parameters ℓ_k, c_k, and u_k appropriately, the unified GEM formula (A.1.3) can be justified for all three types of adjustment: extreme value treatment, nonresponse, and poststratification. For extreme value treatment via winsorization, denote the winsorized weights by $\{b_k\}$, where $b_k = d_k$ if d_k is not an extreme weight, and $b_k = \text{med}\{d_k\} \pm 3 * \text{IQR}$ if d_k is an extreme weight, where IQR denotes the interquartile range, and the median and quartiles for the weights are defined with respect to a suitable design-based stratum.

For the nonresponse adjustment, the sample is first divided into two parts: the nonextreme weight subsample and the extreme weight subsample. For nonextreme weights, the following are set: $\ell_2 = 1,\ c_2 = \rho^{-1},\ u_2 = u > \rho^{-1},$ where ρ is the overall response propensity. For extreme weights with high weights, $\ell_k = \ell_1 m_k,\ c_k = \rho^{-1} m_k,$ and $u_k = u_1 m_k,$ where $m_k = b_k / d_k$ and $1 \le \ell_1 < \rho^{-1} = c_1 < u_1$ are prescribed numbers. Similarly, for extreme weights with low weights, $\ell_k = \ell_3 m_k,\ c_k = \rho^{-1} m_k,\ u_k = u_3 m_k,$ and $1 \le \ell_3 < \rho^{-1} = c_3 < u_3$.

For the poststratification adjustment, the following weights are set: for nonextreme weights, $\ell_k = \ell_2,\ c_k = c_2 = 1,$ and $u_k = u_2;$ for high extreme weights, $\ell_k = \ell_1 m_k,\ c_k = m_k,$ and $u_k = u_1 m_k;$ and similarly, for low extreme weights, $\ell_k = \ell_3 m_k,\ c_k = m_k,$ and $u_k = u_3 m_k.$ The extreme value adjustment is identical to poststratification, except for tighter bounds on extreme weights resulting from the final poststratification.

Notice that the GEM allows the flexibility of specifying different bounds for different subsamples. In addition, the lower bound (in the case of nonresponse adjustments) can be made to equal one by choosing the center $c_k > 1$.

A.3 Newton-Raphson Steps

Let *X* denote the $n \times p$ matrix of predictor values, and for the v^{th} iteration,

$$\Gamma_{\phi v} = \text{diag}\left(d_k \phi_k^{(v)}\right),\ \phi_k^{(o)} = 1,$$

where $\phi_k^{(v)} = \left[\left(u_k - a_k^{(v)}\right)\left(a_k^{(v)} - \ell_k\right)\right] \Big/ \left[\left(u_k - c_k\right)\left(c_k - \ell_k\right)\right].$

Then, for the Newton-Raphson iteration v, the value of the p-vector λ is adjusted as

$$\lambda^{(v)} = \lambda^{(v-1)} + \left(X' \Gamma_{\phi,v-1} X \right)^{-1} \left(T_x - \hat{T}_x^{(v-1)} \right),$$

where $\lambda^{(0)} = 0$, and $\hat{T}_x$ is calculated by using equation A.1.2, in which a_k is calculated by plugging the current λ into equation A.1.3.

The convergence criterion is based on the Euclidean distance $\left\| T_x - \hat{T}_x^{(v)} \right\|$, which is defined as $\sqrt{\left(T_x - \hat{T}_x^{(v)} \right)' \left(T_x - \hat{T}_x^{(v)} \right)}$. At each iteration, it is checked to determine whether it is decreasing. If it is not, a half step is used in the iteration increment for λ.

A.4 Scaled Constrained Exponential Model

In National Household Surveys on Drug Abuse (NHSDAs)[1] prior to 1999, constrained exponential models (CEMs) were used for poststratification, and scaled CEMs were used for nonresponse adjustments. The CEM refers to the logit model of Deville and Särndal (1992), in which lower and upper bounds do not vary with k; that is, $\ell_k = \ell$, $u_k = u$, and $c_k = c = 1$, such that $\ell < 1 < u$. Thus, the CEM is a special case of the GEM. For the nonresponse adjustment, Folsom and Witt (1994) modified the CEM estimating equations by a scaling factor (ρ^{-1}, the inverse of the overall response propensity), such that $1 < \rho^{-1} a_k < \rho^{-1} u$. This implies that choosing ℓ in the CEM as ρ ensures that the scaled adjustment factor for nonresponse is at least one.

[1] The National Household Survey on Drug Abuse (NHSDA) was renamed the National Survey on Drug Use and Health (NSDUH) in the 2002 survey year.

This page intentionally left blank

Appendix B: Derivation of Poststratification Control Totals

This page intentionally left blank

Appendix B: Derivation of Poststratification Control Totals

Unlike the person-level poststratification adjustment, the control totals for questionnaire dwelling unit (QDU)-level and person pair-level weight calibration could not be derived from the U.S. Census Bureau directly. Estimates of the number of households and person pairs were not available at the domains that we wanted to control, and person pair population estimates were not available even at a national level. However, by taking advantage of the two-phase design of the National Survey on Drug Use and Health (NSDUH), the screener dwelling unit (SDU) sample weights could be poststratified to census population estimates. The calibrated SDU weights then could be used as stable control totals for the QDU- and person pair-level sample weights. In addition to the SDU weights, the person pair-level weights were calibrated to a second set of controls derived from the questionnaire, called household-level person counts. These controls were applied to pairs that were members of the 10 selected pair domains given below.

1. Parent-child pairs, child aged 12 to 14, target population is parents whose children aged 12 to 14 live with them;
2. Parent-child pairs, child aged 12 to 14, target population is children aged 12 to 14 living with their parents;
3. Parent-child pairs, child aged 12 to 17, target population is parents whose children aged 12 to 17 live with them;
4. Parent-child pairs, child aged 12 to 17, target population is children aged 12 to 17 living with their parents;
5. Parent-child pairs, child aged 12 to 20, target population is parents whose children aged 12 to 20 live with them;
6. Parent-child pairs, child aged 12 to 20, target population is children aged 12 to 20 living with their parents;
7. Sibling-sibling pairs, older sibling aged 15 to 17, younger sibling aged 12 to 14, target population is siblings aged 15 to 17 whose siblings are aged 12 to 14;
8. Sibling-sibling pairs, older sibling aged 18 to 25, younger sibling aged 12 to 17, target population is siblings aged 18 to 25 whose siblings are aged 12 to 17;
9. Spouse-spouse and partner-partner pairs; and
10. Spouse-spouse and partner-partner pairs with children younger than the age of 18 living in the household.

B.1 Derivation of QDU-Level Poststratification Controls

The derivation of QDU-level poststratification controls was not directly possible. Instead, it had to be based on work done for the person-level calibration. At the person level, weights were calibrated to the control totals that we wished to reach. These weights then were altered in order to conform to use with QDU-level data.

B.1.1 Person Level

B.1.1.1 Receiving and Deriving Person-Level Poststratification Control Totals

Civilian, noninstitutionalized population estimates for ages 12 or older were provided by the Population Estimates Branch of the U.S. Census Bureau. We received two files, one at the national level and the other at the state level, each containing estimates of the population broken down by levels of month (12), Hispanicity (2), race (6), gender (2), and age (11).

The breakdown received from the census did not match the levels of the domains that we wanted to control. To account for this, we collapsed levels. From this altered data, we created datasets with model group-specific control totals. Observations in these datasets corresponded to a breakdown by quarter (4), Hispanicity (2), race (5), gender (2), age (11), and number of states[1] in the model group (number of states varied according to which census region was represented in the model group).

B.1.1.2 Adjusting SDU Data to the Control Totals

In the person-level weighting, the SDU weights were poststratified to meet control totals based on the population estimates received from the census. For NSDUH weighting, GEM was utilized to calibrate sample weights to multiple control totals. In doing so, each SDU received an adjustment factor, which, when multiplied by the initial weight, produced a final weight. The sum of all final weights corresponded to the civilian, noninstitutionalized population estimate for ages 12 or older, and the sum of all final weights in a domain corresponded to the control total for that domain. Note that there were a number of controls being calibrated to for each SDU, depending upon the domains to which the SDU belonged. The adjusted SDU weight reflected the civilian, noninstitutionalized population estimates for ages 12 or older and could be utilized as a basis for constructing controls at the QDU and person pair levels.

B.1.2 QDU Level

B.1.2.1 Deriving QDU-Level Poststratification Control Totals from Adjusted SDU Weights

Since there were no controls for QDU-level poststratification available directly, we used the adjusted SDU weights. For these weights to be applicable at the QDU level, the SDU-level data had to be restructured by sorting and summing over the domains to be used in the QDU-level calibration. This provided a dataset where the summed weight, which still added up to the proper population, was available for every domain to be utilized in the QDU calibration and thus could be used as a control total.

[1] The District of Columbia is included among states.

B.1.2.2 Adjusting QDU-Level Data to the Control Totals

As was done for the SDU data, the QDU-level data were adjusted via calibration in GEM of sample weights to multiple control totals. Each QDU received an adjustment factor, similar to that described for the SDU weight in B.1.1.2. The controls utilized in this calibration were based on the SDU weight as described in B.1.2.1 above. The adjusted weight was representative of the civilian, noninstitutionalized population estimates for ages 12 or older for all domains controlled within the modeling.

B.2 Derivation of Person Pair-Level Poststratification Controls

B.2.1 Deriving Person Pair-Level Poststratification Control Totals from Adjusted SDU Weights and Household-Level Person Counts

Analogous to the QDU weights, some of the person pair controls were based on the SDU weights. However, two sets of control totals were utilized in the modeling, with one set based on the SDU weights and the other set based on the questionnaire roster.

For most pair data domains—those other than the 10 pair domains based on relationship—the control totals for the poststratification adjustments were obtained from SDU data and were based on the number of possible pairs within SDUs. In order to obtain these pair counts belonging to various sociodemographic domains, the screener roster information was used to calculate all possible pairs within SDUs. For example, consider an SDU with two people aged 12 to 17 and three people aged 26 to 34. From this household composition, one can construct one pair of people aged 12 to 17, three pairs of people aged 26 to 34, and six pairs of people aged 12 to 17 and 26 to 34. It follows that the total number of possible pairs in this SDU is 10, from which the number of pairs belonging to the domain of interest can be obtained.

On the other hand, for the 10 selected pair domains based on relationship, the control totals for the poststratification adjustments were obtained from the questionnaire roster. This involved calibrating the pair weights to the number of people in households belonging to each domain of interest. These controls were obtained from the larger sample of singles and pairs (i.e., one or two people selected from dwelling units) and were calculated at the QDU (household) level. The pair weights were adjusted by the appropriate multiplicity. See Chapter 11 in the NSDUH Methodological Resource Book editing and imputation report (Center for Behavioral Health Statistics and Quality, 2019a) for details on the multiplicity counts and household-level control totals, which are referred to as household-level person counts.

B.2.2 Adjusting Person-Pair Level Data to the Control Totals

Like the SDU- and QDU-level data, the person pair-level data was adjusted via GEM. The use of two different types of controls required a minor modification to the GEM macro so that both sets of controls might be addressed simultaneously. Similar to the SDU- and QDU-level poststratification steps, each pair received an adjustment factor, which, when multiplied by the initial weight, produced a final weight. The sum of all final weights corresponded to the civilian, noninstitutionalized population estimate for ages 12 or older, and the sum of all final weights in a domain corresponded to the control total for that domain.

This page intentionally left blank

Appendix C: GEM Modeling Summary for the Questionnaire Dwelling Unit Weights

This page intentionally left blank

Appendix C: GEM Modeling Summary for the Questionnaire Dwelling Unit Weights

This appendix summarizes each questionnaire dwelling unit (QDU) model group throughout all stages of weight calibration modeling. Unlike much of the other information presented in this report, this appendix provides a model-specific overview of weight calibration, as opposed to a state- or domain-specific one.

For 2017, modeling involved taking four model groups through three adjustment steps: (1) selected dwelling unit poststratification, (2) respondent dwelling unit nonresponse adjustment, and (3) respondent dwelling unit poststratification. After the final poststratification, the adjusted sampling weights were reasonably distributed and did not require the additional treatment of the extreme value step.

Model-specific summary statistics are shown in Tables C.1a through C.4b. Included in these tables, for each stage of modeling, are the number of factor effects included; the high, low, and nonextreme weight bounds set to provide the upper and lower limits for the generalized exponential model (GEM) macro; weighted, unweighted, and winsorized weight proportions; the unequal weighting effect (UWE); and weight distributions. The UWE provides an approximate partial measure of variance and provides a summary of how much impact a particular stage of modeling has on the distribution of the new product of weights. For more details on bounds, see Section 4.1. At each stage in the modeling, these summary statistics were calculated and utilized to help evaluate the quality of the current weight component under the model chosen.

Occurrences of small sample sizes and exact linear combinations in the realized data led to situations whereby inclusion of all originally proposed levels of covariates in the model was not possible. The text and exhibits in Sections C.1 through C.4 summarize the decisions made with regard to final covariates included in each model. For a list of the proposed initial covariates considered at each stage of modeling, see Exhibit C.2, and for the list of realized final model covariates, see Exhibits C.1.1 through C.4.3. The following sections establish a series of guidelines to assist in their interpretation.

C.1 Final Model Explanatory Variables

For brevity, numeric abbreviations for factor levels are established in Exhibit 4.1 (included here as Exhibit C.1 for easy reference) in Chapter 4. There, a complete list is provided of all variables and associated levels used at any stage of modeling. Note that not all factors or levels were present in all stages of modeling, and the initial set of variables was the same across model groups but may change over stages of modeling. The initial candidates are found in any of the proposed variables columns for a particular stage of weight adjustment. Exhibits C.1.1 through C.4.3 provide lists of the proposed and realized covariates.

To help understand what effects were controlled for at each stage of the modeling, it was useful to create cross-classification tables as shown in Section C.3. Sections C.2 and C.3 explain how to use various exhibits for selected model variables to construct these tables.

Exhibit C.1 Definitions of Levels for QDU-Level Calibration Modeling Variables

Age[a]
1: 12–17, 2: 18–25, 3: 26–34, 4: 35–49, 5: 50+[1]

Gender[a]
1: Male, 2: Female[1]

Group Quarter Indicator[b]
1: College Dorm, 2: Other Group Quarter, 3: Non-Group Quarter[1]

Hispanicity[a]
1: Hispanic or Latino, 2: Non-Hispanic or Latino[1]

Household Size[a]
Continuous Variable Count of Individuals Rostered with DU

Household Type (Ages of People Rostered within DU)[b]
1: 12–17, 18–25, 26+; 2: 12–17, 18–25; 3: 12–17, 26+; 4: 18–25, 26+; 5: 12–17; 6: 18–25; 7: 26+[1]

Percentage of Owner-Occupied Dwelling Units in Segment (% Owner-Occupied)[b]
1: 50–100%,[1] 2: 10–<50%, 3: 0–<10%

Percentage of Segments That Are Black or African American[b]
1: 50–100%, 2: 10–<50%, 3: 0–<10%[1]

Percentage of Segments That Are Hispanic or Latino[b]
1: 50–100%, 2: 10–<50%, 3: 0–<10%[1]

Population Density[b]
1: MSA 1,000,000 or More, 2: MSA Less than 1,000,000, 3: Non-MSA Urban, 4: Non-MSA Rural[1]

Quarter[a,b]
1: Quarter 1, 2: Quarter 2, 3: Quarter 3, 4: Quarter 4[1]

Race (3 Levels)[a]
1: White[1], 2: Black or African American, 3: Other

Race (5 Levels)[a]
1: White,[1] 2: Black or African American, 3: American Indian or Alaska Native, 4: Asian, 5: Two or More Races

Race/Ethnicity of Householder[b]
1: Hispanic or Latino White,[1] 2: Hispanic or Latino Black or African American, 3: Hispanic or Latino Other, 4: Non-Hispanic or Latino White, 5: Non-Hispanic or Latino Black or African American, 6: Non-Hispanic or Latino Other

Relation to Householder[b]
1: Householder or Spouse, 2: Child, 3: Other Relative, 4: Nonrelative[1]

Segment-Combined Median Rent and Housing Value (Rent/Housing)[b,2]
1: First Quintile, 2: Second Quintile, 3: Third Quintile, 4: Fourth Quintile, 5: Fifth Quintile[1]

State[a,b,3]
Model Group 1: 1: Connecticut, 2: Maine, 3: Massachusetts,[1] 4: New Hampshire, 5: New Jersey, 6: New York, 7: Pennsylvania, 8: Rhode Island, 9: Vermont
Model Group 2: 1: Illinois, 2: Indiana, 3: Iowa, 4: Kansas, 5: Michigan, 6: Minnesota, 7: Missouri, 8: Nebraska, 9: North Dakota, 10: Ohio, 11: South Dakota, 12: Wisconsin[1]
Model Group 3: 1: Alabama, 2: Arkansas, 3: Delaware, 4: District of Columbia, 5: Florida, 6: Georgia, 7: Kentucky, 8: Louisiana, 9: Maryland, 10: Mississippi, 11: North Carolina,[1] 12: Oklahoma, 13: South Carolina, 14: Tennessee, 15: Texas, 16: Virginia, 17: West Virginia
Model Group 4: 1: Alaska, 2: Arizona,[1] 3: California, 4: Colorado, 5: Idaho, 6: Hawaii, 7: Montana, 8: Nevada, 9: New Mexico, 10: Oregon, 11: Utah, 12: Washington, 13: Wyoming

State/Region[b,3]
Model Group 1: 1: New York, 2: Pennsylvania, 3: Other[1]
Model Group 2: 1: Illinois, 2: Michigan, 3: Ohio, 4: Other[1]
Model Group 3: 1: Florida, 2: Texas, 3: Other[1]
Model Group 4: 1: California, 2: Other[1]

DU = dwelling unit; MSA = metropolitan statistical area; QDU = questionnaire dwelling unit.

[1] The reference level for this variable. This is the level against which effects of other factor levels are measured.

[2] Segment-Combined Median Rent and Housing Value is a composite measure based on rent, housing value, and percentage owner-occupied.

[3] The states or district assigned to a particular model is based on census regions.

[a] Counting variable. A count of all people in the household.

[b] Binary variable.

C.2 Glossary of Terms Used in the Description of the Variables in the Final Model

This glossary provides a list of general terms. Certain other terms are sometimes used within a particular section.

All levels present. All effects and all levels of the factor under consideration are in the model.

Coll. *(levels)*. Collapse these factor effects together. Factor effects that have been collapsed with others manifest themselves jointly in the model.

Conv. If the model is not convergent, dropping or collapsing of variables is performed.

Do the same for *(effects)*. Repeat the previous step for all effect levels listed.

Drop all levels. All factor effects are completely removed from the model for all levels and any combinations involving this factor.

Drop *level(s)*. Collapse these factor effects into the reference set. The factor effects comprising the dropped levels are manifested jointly with either some or all of the factor effects in the reference set.

Drop *level(s)*; sing. During the modeling process, the factor effects listed are removed from the model due to singularity.

Drop *level(s); zero cnts.* During the modeling process, the factor effects listed are removed from the model due to zero sample.

Drop or collapse using*. The asterisk is used as a wild card character to indicate all levels of the factor for that effect.

Factor effect. The factor effect represents the effects of levels considered for one factor, two factors, and higher order factors.

Hier. One or more of the factor effects in a higher order interaction is collapsed or dropped in an interaction at a lower order and the hierarchical effect carries up, either eliminating or combining factors of higher order interactions with that effect.

Reference/reference set. Factor effects composed of reference levels are not explicitly listed in the set of model variables. However, these effects manifest themselves either separately or in combination with other factors depending on the presence of other factors in the model.

C.3 How to Interpret Collapsing and Dropping of Factor Effects

To help visualize what effects are directly controlled for in our model, one can construct the table that reflects the collapsing scheme employed. The following is a complex example from the 2004 person-level modeling (Chen et al., 2006).

1. Locate the Factor Effect—Model 9 Person Nonresponse Adjustment.

Three-Factor Effects	Comments
State × Age × Race (3 Levels)	Coll. (2,1,2) & (2,1,3); hier. Repeat for all levels of age in state (2); hier. Coll. (1,4,2) & (1,4,3); conv. Drop (3,4,2); sing. Drop (3,*,*); conv. Coll. (5,1,2) & (5,1,3); conv. Repeat for all levels of age in state (5).

2. Determine the initial range of possible levels for the variables by referring to the variable definitions. See Exhibits C.1 and H.1 for QDU- and pair-level variable definitions. In addition, the columns "Levels," "Proposed," and "Final" will provide counts of all factor effects, all explicitly proposed factors, and all explicitly controlled factors, but these are not necessary for construction of the cross-classification table. The following example is based upon person-level variables, but the process is the same.

State (for the model group in question, in this case, Model Group 9)
Model Group 9: 1: Alaska, 2: Hawaii, 3: Oregon, 4: Washington,[1] 5: California

Age
1: 12 to 17, 2: 18 to 25, 3: 26 to 34, 4: 35 to 49, 5: 50+[1]

Race (3 Levels)
1: White,[1] 2: Black or African American, 3: Other

3. Construct the cross-classification table.

For example, the initial proposed set of covariates in Race (4 Levels) is defined this way:

Race (4 Levels)	White	Black or African American	Asian	American Indian or Alaska Native

Shading indicates the reference-level set.

[1] This is the reference level for this variable. This is the level against which effects of other factor levels are measured.

This is the cross-classification table for the initial proposed set of covariates in State × Race (4 Levels):

State × Race (4 Levels)	White	Black or African American	Asian	American Indian or Alaska Native
AK				
HI				
OR				
WA				
CA				

Shading indicates the reference-level set.

The cross-classification table of interest for the initial proposed set of covariates in State × Age × Race (3 Levels) is as follows:

State × Age × Race (3 Levels)	White	Black or African American	Other
AK × 12–17			
18–25			
26–34			
35–49			
50+			
HI × 12–17			
18–25			
26–34			
35–49			
50+			
OR × 12–17			
18–25			
26–34			
35–49			
50+			
WA × 12–17			
18–25			
26–34			
35–49			
50+			
CA × 12–17			
18–25			
26–34			
35–49			
50+			

Shading indicates the reference-level set.

The number of respondents in the class State × Age × Race (3 Levels) at this stage of modeling would appear within each cell of the table. Construction of the other cross-classification tables follows the same logic and is only necessary to the point of providing understanding of the final table.

4. Use the information under the "Comments" column definition to determine the combination of factors controlled.

One-Factor Effects	Comments
State	All levels present.
Race (4 Levels)	All levels present.
Age	All levels present.

Two-Factor Effects	Comments
State × Age	All levels present.
State × Race (4 Levels)	Coll. (1,3) & (1,4). Do the same for all other states except (2). Coll. (2,2), (2,3), & (2,4).
Age × Race (3 Levels)	All levels present.

The reason for the hier. instruction in the three-factor effect directions is the State × Race (4 Levels) interaction. It indicates a need to maintain the collapsing scheme when setting up any three-factor crosses involving State × Race. Following these directions, the resulting two-factor table we would then have to work with is as follows:

State × Race (4 Levels)	White	Black or African American	Asian	American Indian or Alaska Native
AK				
HI				
OR				
WA				
CA				

Shading indicates the reference-level set.

Returning to our instructions, we see that several other factor crosses have been affected by modeling:

Three-Factor Effects	Comments
State × Age × Race (3 Levels)	Coll. (2,1,2) & (2,1,3); hier. Repeat for all levels of age in state (2); hier. Coll. (1,4,2) & (1,4,3); conv. Drop (3,4,2); sing. Drop (3,*,*); conv. Coll. (5,1,2) & (5,1,3); conv. Repeat for all levels of age in state (5).

Construct the complete table, and then begin combining blocks as directed. The unshaded cells represent the factors directly controlled for by the model. The shaded cells represent the composite reference set, whose values may be obtained by utilizing the marginal sums, although when changes to the initially proposed set occur, it can make certain reference cell counts indistinguishable.

After following the directions, the resulting post-modeling cross-classification table should appear as follows:

State × Age × Race (3 Levels)	White	Black or African American	Other
AK × 12–17			
18–25			
26–34			
35–49			
50+			
HI × 12–17			
18–25			
26–34			
35–49			
50+			
OR × 12–17			
18–25			
26–34			
35–49			
50+			
WA × 12–17			
18–25			
26–34			
35–49			
50+			
CA × 12–17			
18–25			
26–34			
35–49			
50+			

Shading indicates the reference-level set.

Exhibit C.2 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights

Variables	Binary	Counting	Level	Proposed
One-Factor Effects				
Intercept	Yes		1	1
Population Density	Yes		4	3
Group Quarter	Yes		3	2
Race/Ethnicity of Householder	Yes		6	5
Rent/Housing	Yes		5	4
Segment % Black or African American	Yes		3	2
Segment % Hispanic or Latino	Yes		3	2
Segment % Owner-Occupied	Yes		3	2
Household Type	Yes		7	6
State	Yes	Yes	Model-specific	
Quarter	Yes	Yes	4	3
Age Group		Yes	5	4
Race		Yes	5	4
Hispanicity		Yes	2	1
Gender		Yes	2	1
Household Size		Yes	1	1
Two-Factor Effects				
Age × Race (3 Levels)		Yes	5 × 3	8
Age × Hispanicity		Yes	5 × 2	4
Age × Gender		Yes	5 × 2	4
Race (3 Levels) × Hispanicity		Yes	3 × 2	2
Race (3 Levels) × Gender		Yes	3 × 2	2
Hispanicity × Gender		Yes	2 × 2	1
State × Age		Yes	Model-specific	
State × Race (5 Levels)		Yes	Model-specific	
State × Gender		Yes	Model-specific	
State × Hispanicity		Yes	Model-specific	
% Black or African American × % Owner-Occupied	Yes		3 × 5	8
% Black or African American × Rent/Housing	Yes		3 × 5	8
% Hispanicity × % Owner-Occupied	Yes		3 × 3	4
% Hispanicity × Rent/Housing	Yes		3 × 5	8
% Owner × Rent/Housing	Yes		3 × 5	8
Three-Factor Effects				
Race (3 Levels) × Age × Gender		Yes	8	8
State/Region × Age × Gender		Yes	Model-specific	
State/Region × Age × Hispanicity		Yes	Model-specific	
State/Region × Age × Race (3 Levels)		Yes	Model-specific	
State/Region × Hispanicity × Gender		Yes	Model-specific	
State/Region × Race (3 Levels) × Hispanicity		Yes	Model-specific	
State/Region × Race (3 Levels) × Gender		Yes	Model-specific	

Appendix C.1: Model Group 1: Northeast

(Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont)

This page intentionally left blank

Table C.1a 2017 QDU Weight GEM Modeling Summary (Model Group 1: Northeast)

Modeling Step[1]	Extreme Weight Proportions			UWE[2]	# Covariates[3]	Bounds[4]	
	% Unweighted	% Weighted	% Outwinsor			Nominal	Realized
sel.qdu.ps	2.22	5.42	1.49	2.3017	243	(0.76, 2.70)	(0.77, 2.70)
	1.67	4.16	0.88	2.3230	242	(0.38, 2.70)	(0.40, 2.25)
						(0.90, 1.08)	(0.90, 1.08)
res.qdu.nr	1.53	3.42	0.77	2.3118	243	(1.00, 2.30)	(1.00, 2.30)
	1.44	4.35	0.70	2.5159	241	(1.00, 3.25)	(1.00, 3.24)
						(1.40, 1.73)	(1.40, 1.73)
res.qdu.ps	1.44	4.35	0.70	2.5159	243	(0.97, 1.60)	(0.99, 1.60)
	1.34	3.98	0.43	2.5020	241	(0.86, 1.60)	(0.93, 1.12)
						N/A	N/A

GEM = generalized exponential model; QDU = questionnaire dwelling unit.

[1] For a key to modeling abbreviations, see Chapter 6, Exhibit 6.1.

[2] Unequal weighting effect (UWE) defined as $1 + \left[(n-1)/n\right]*CV^2$, where CV = coefficient of variation of weights.

[3] Number of proposed covariates on top line and number finalized after modeling.

[4] There are six sets of bounds for each modeling step. Nominal bounds are used in defining maximum/minimum values for the GEM adjustment factors. The realized bound is the actual adjustment produced by the modeling. The set of three bounds listed for each step correspond to the high extreme values, the nonextreme values, and the low extreme values.

Table C.1b 2017 Distribution of Weight Adjustment Factors and Weight Products (Model Group 1: Northeast)

	SDU Weight	QDU Design Weight		sel.qdu.ps[1]		res.qdu.nr[1]		res.qdu.ps[1]	
Statistics	**1–11**	**duwght12**	**1–12**	**duwght13**	**1–13**	**duwght14**	**1–14**	**duwght15**	**1–15**
Minimum	15	1.00	15	0.37	9	0.40	9	0.73	9
1%	59	1.00	78	0.71	73	1.00	87	0.96	87
5%	112	1.00	153	0.82	149	1.04	184	0.98	183
10%	159	1.00	209	0.87	208	1.10	265	0.99	265
25%	279	1.00	532	0.94	515	1.24	630	1.00	629
Median	810	1.28	1,065	1.00	1,062	1.38	1,370	1.00	1,371
75%	1,048	2.05	1,780	1.06	1,808	1.54	2,597	1.00	2,598
90%	1,317	4.59	3,392	1.13	3,498	1.74	5,163	1.01	5,184
95%	1,576	6.96	4,961	1.20	5,142	1.89	7,785	1.02	7,780
99%	2,418	9.24	8,868	1.41	9,115	2.32	13,314	1.04	13,375
Maximum	6,350	14.33	35,590	2.37	30,620	3.24	42,823	1.25	37,426
n	14,037	-	14,037	-	14,037	-	9,915	-	9,915
Mean	781	2.09	1,552	1.01	1,570	1.41	2,223	1.00	2,223
Max/Mean	8	-	23	-	20	-	19	-	17

QDU = questionnaire dwelling unit; SDU = screener dwelling unit.

[1] For a key to modeling abbreviations, see Chapter 6, Exhibit 6.1.

Model Group 1 Overview

Selected Questionnaire Dwelling Unit-Level Poststratification

Of the 243 proposed effects, 242 were kept in the model. Main effect Household Type category 12–17 had zero sample and was combined with 18–25.

Respondent Questionnaire Dwelling Unit-Level Nonresponse

Of the 243 proposed effects, 241 were kept in the model. Main effect Household Type category 12–17 had zero sample and was combined with 18–25. New Hampshire race categories American Indian or Alaska Native and Asian were combined in the State × Race interaction.

Respondent Questionnaire Dwelling Unit-Level Poststratification

This step used the same set of 241 effects as the respondent questionnaire dwelling unit-level nonresponse.

Exhibit C.1.1 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (sel.qdu.ps) Model Group 1: Northeast

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		**60**	**59**	
Intercept	1	1	1	All levels present.
Group Quarter	3	2	2	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Household Type	7	6	5	Coll. (5) & (6); sing.
Household Size	1	1	1	All levels present.
Rent/Housing	5	4	4	All levels present.
Population Density	4	3	3	All levels present.
% Black or African American	3	2	2	All levels present.
% Hispanic or Latino	35	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
State (Count)	9	8	8	All levels present.
State (Binary)	9	8	8	All levels present.
Quarter (Count)	4	3	3	All levels present.
Quarter (Binary)	4	3	3	All levels present.
Age Group	5	4	4	All levels present.
Race	5	4	4	All levels present.
Hispanicity	2	1	1	All levels present.
Gender	2	1	1	All levels present.
Two-Factor Effects		**133**	**133**	
Age × Race (3 Levels)	5 × 3	8	8	All levels present.
Age × Hispanicity	5 × 2	4	4	All levels present.
Age × Gender	5 × 2	4	4	All levels present.
Race (3 Levels) × Hispanicity	3 × 2	2	2	All levels present.
Race (3 Levels) × Gender	3 × 2	2	2	All levels present.
Hispanicity × Gender	2 × 2	1	1	All levels present.
State × Age	9 × 5	32	32	All levels present.
State × Race	9 × 5	32	32	All levels present.
State × Gender	9 × 2	8	8	All levels present.
State × Hispanicity	9 × 2	8	8	All levels present.
% Black or African American × % Owner-Occupied	3 × 3	4	4	All levels present.
% Black or African American × Rent/Housing	3 × 5	8	8	All levels present.
% Hispanicity × % Owner-Occupied	3 × 3	4	4	All levels present.
% Hispanicity × Rent/Housing	3 × 5	8	8	All levels present.
% Owner-Occupied × Rent/Housing	3 × 5	8	8	All levels present.
Three-Factor Effects		**50**	**50**	
Race (3 Levels) × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Hispanicity	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Race (3 Levels)	3 × 5 × 3	16	16	All levels present.
State/Region × Hispanicity × Gender	3 × 2 × 2	2	2	All levels present.
State/Region × Race (3 Levels) × Hispanicity	3 × 3 × 2	4	4	All levels present.
State/Region × Race (3 Levels) × Gender	3 × 3 × 2	4	4	All levels present.
Total		**243**	**242**	

Exhibit C.1.2 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 1: Northeast

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		**60**	**59**	
Intercept	1	1	1	All levels present.
Group Quarter	3	2	2	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Household Type	7	6	5	Coll. (5) & (6); sing.
Household Size	1	1	1	All levels present.
Rent/Housing	5	4	4	All levels present.
Population Density	4	3	3	All levels present.
% Black or African American	3	2	2	All levels present.
% Hispanic or Latino	35	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
State (Count)	9	8	8	All levels present.
State (Binary)	9	8	8	All levels present.
Quarter (Count)	4	3	3	All levels present.
Quarter (Binary)	4	3	3	All levels present.
Age Group	5	4	4	All levels present.
Race	5	4	4	All levels present.
Hispanicity	2	1	1	All levels present.
Gender	2	1	1	All levels present.
Two-Factor Effects		**133**	**132**	
Age × Race (3 Levels)	5 × 3	8	8	All levels present.
Age × Hispanicity	5 × 2	4	4	All levels present.
Age × Gender	5 × 2	4	4	All levels present.
Race (3 Levels) × Hispanicity	3 × 2	2	2	All levels present.
Race (3 Levels) × Gender	3 × 2	2	2	All levels present.
Hispanicity × Gender	2 × 2	1	1	All levels present.
State × Age	9 × 5	32	32	All levels present.
State × Race	9 × 5	32	31	Coll. (4,3) & (4,4); conv.
State × Gender	9 × 2	8	8	All levels present.
State × Hispanicity	9 × 2	8	8	All levels present.
% Black or African American × % Owner-Occupied	3 × 3	4	4	All levels present.
% Black or African American × Rent/Housing	3 × 5	8	8	All levels present.
% Hispanicity × % Owner-Occupied	3 × 3	4	4	All levels present.
% Hispanicity × Rent/Housing	3 × 5	8	8	All levels present.
% Owner-Occupied × Rent/Housing	3 × 5	8	8	All levels present.
Three-Factor Effects		**50**	**50**	
Race (3 Levels) × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Hispanicity	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Race (3 Levels)	3 × 5 × 3	16	16	All levels present.
State/Region × Hispanicity × Gender	3 × 2 × 2	2	2	All levels present.
State/Region × Race (3 Levels) × Hispanicity	3 × 3 × 2	4	4	All levels present.
State/Region × Race (3 Levels) × Gender	3 × 3 × 2	4	4	All levels present.
Total		**243**	**241**	

Exhibit C.1.3 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 1: Northeast

This step used the same set of effects as the respondent questionnaire dwelling unit-level nonresponse.

Appendix C.2: Model Group 2: Midwest

(Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, Wisconsin)

This page intentionally left blank

Table C.2a 2017 QDU Weight GEM Modeling Summary (Model Group 2: Midwest)

Modeling Step[1]	Extreme Weight Proportions			UWE[2]	# Covariates[3]	Bounds[4]	
	% Unweighted	% Weighted	% Outwinsor			Nominal	Realized
sel.qdu.ps	1.90	2.92	0.66	1.9500	300	(0.63, 1.10)	(0.63, 1.10)
	1.57	2.16	0.21	1.9428	298	(0.58, 1.89)	(0.58, 1.89)
						(0.90, 1.78)	(0.90, 1.78)
res.qdu.nr	1.29	1.58	0.14	1.9903	300	(1.00, 2.10)	(1.00, 2.10)
	1.10	1.59	0.28	2.0484	299	(1.00, 3.79)	(1.00, 3.79)
						(1.40, 1.58)	(1.40, 1.58)
res.qdu.ps	1.10	1.59	0.28	2.0484	300	(0.20, 1.60)	(0.96, 1.60)
	1.11	1.53	0.19	2.0483	299	(0.20, 5.00)	(0.37, 2.03)
						(0.90, 5.00)	(0.90, 0.98)

GEM = generalized exponential model; QDU = questionnaire dwelling unit.

[1] For a key to modeling abbreviations, see Chapter 6, Exhibit 6.1.

[2] Unequal weighting effect (UWE) defined as $1 + \left[(n-1)/n\right] * CV^2$, where CV = coefficient of variation of weights.

[3] Number of proposed covariates on top line and number finalized after modeling.

[4] There are six sets of bounds for each modeling step. Nominal bounds are used in defining maximum/minimum values for the GEM adjustment factors. The realized bound is the actual adjustment produced by the modeling. The set of three bounds listed for each step correspond to the high extreme values, the nonextreme values, and the low extreme values.

Table C.2b 2017 Distribution of Weight Adjustment Factors and Weight Products (Model Group 2: Midwest)

	SDU Weight	QDU Design Weight		sel.qdu.ps[1]		res.qdu.nr[1]		res.qdu.ps[1]	
Statistics	**1–11**	**duwght12**	**1–12**	**duwght13**	**1–13**	**duwght14**	**1–14**	**duwght15**	**1–15**
Minimum	39	1.00	44	0.32	46	0.69	65	0.37	64
1%	108	1.00	113	0.67	113	1.00	142	0.90	142
5%	151	1.00	219	0.82	213	1.06	255	0.98	253
10%	200	1.00	342	0.88	332	1.13	415	0.99	413
25%	608	1.00	762	0.94	761	1.25	965	1.00	964
Median	857	1.27	1,155	1.00	1,167	1.37	1,529	1.00	1,534
75%	1,087	2.08	1,958	1.06	1,961	1.49	2,729	1.00	2,731
90%	1,364	4.16	3,617	1.13	3,579	1.61	5,267	1.01	5,238
95%	1,556	5.51	5,066	1.19	5,112	1.73	7,325	1.02	7,352
99%	2,253	7.99	8,219	1.40	8,060	2.12	11,773	1.07	11,766
Maximum	5,178	10.04	18,980	1.90	16,930	3.79	24,564	2.03	24,680
n	16,282	-	16,282	-	16,282	-	11,760	-	11,760
Mean	860	1.97	1,673	1.01	1,671	1.39	2,313	1.00	2,313
Max/Mean	6	-	11	-	10	-	11	-	11

QDU = questionnaire dwelling unit; SDU = screener dwelling unit.

[1] For a key to modeling abbreviations, see Chapter 6, Exhibit 6.1.

Model Group 2 Overview

Selected Questionnaire Dwelling Unit-Level Poststratification

Out of 300 proposed effects, 298 were kept in the model. Main effect Household Type category 12–17 had zero sample and was combined with 18–25. Three-factor effect State/Region × Race × Hispanicity was modified by collapsing Race levels Black or African American and Other for Ohio.

Respondent Questionnaire Dwelling Unit-Level Nonresponse

Out of 300 proposed effects, 299 were kept in the model. Main effect Household Type category 12–17 had zero sample and was combined with 18–25.

Respondent Questionnaire Dwelling Unit-Level Poststratification

This step used the same set of 299 effects as the respondent questionnaire dwelling unit-level nonresponse.

Exhibit C.2.1 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (sel.qdu.ps) Model Group 2: Midwest

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		**66**	**65**	
Intercept	1	1	1	All levels present.
Group Quarter	3	2	2	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Household Type	7	6	5	Coll. (5) & (6); sing.
Household Size	1	1	1	All levels present.
Rent/Housing	5	4	4	All levels present.
Population Density	4	3	3	All levels present.
% Black or African American	3	2	2	All levels present.
% Hispanic or Latino	35	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
State (Count)	9	11	11	All levels present.
State (Binary)	9	11	11	All levels present.
Quarter (Count)	4	3	3	All levels present.
Quarter (Binary)	4	3	3	All levels present.
Age Group	5	4	4	All levels present.
Race	5	4	4	All levels present.
Hispanicity	2	1	1	All levels present.
Gender	2	1	1	All levels present.
Two-Factor Effects		**163**	**163**	
Age × Race (3 Levels)	5 × 3	8	8	All levels present.
Age × Hispanicity	5 × 2	4	4	All levels present.
Age × Gender	5 × 2	4	4	All levels present.
Race (3 Levels) × Hispanicity	3 × 2	2	2	All levels present.
Race (3 Levels) × Gender	3 × 2	2	2	All levels present.
Hispanicity × Gender	2 × 2	1	1	All levels present.
State × Age	12 × 5	44	44	All levels present.
State × Race	12 × 5	44	44	All levels present.
State × Gender	12 × 2	11	11	All levels present.
State × Hispanicity	12 × 2	11	11	All levels present.
% Black or African American × % Owner-Occupied	3 × 3	4	4	All levels present.
% Black or African American × Rent/Housing	3 × 5	8	8	All levels present.
% Hispanicity or Latino × % Owner-Occupied	3 × 3	4	4	All levels present.
% Hispanicity or Latino × Rent/Housing	3 × 5	8	8	All levels present.
% Owner-Occupied × Rent/Housing	3 × 5	8	8	All levels present.
Three-Factor Effects		**71**	**70**	
Race (3 Levels) × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Gender	4 × 5 × 2	12	12	All levels present.
State/Region × Age × Hispanicity	4 × 5 × 2	12	12	All levels present.
State/Region × Age × Race (3 Levels)	4 × 5 × 3	24	24	All levels present.
State/Region × Hispanicity × Gender	4 × 2 × 2	3	3	All levels present.
State/Region × Race (3 Levels) × Hispanicity	4 × 3 × 2	6	5	Coll. (3,2,1) & (3,3,1); conv.
State/Region × Race (3 Levels) × Gender	4 × 3 × 2	6	6	All levels present.
Total		**300**	**298**	

Exhibit C.2.2 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 2: Midwest

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		**66**	**65**	
Intercept	1	1	1	All levels present.
Group Quarter	3	2	2	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Household Type	7	6	5	Coll. (5) & (6); sing.
Household Size	1	1	1	All levels present.
Rent/Housing	5	4	4	All levels present.
Population Density	4	3	3	All levels present.
% Black or African American	3	2	2	All levels present.
% Hispanic or Latino	35	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
State (Count)	9	11	11	All levels present.
State (Binary)	9	11	11	All levels present.
Quarter (Count)	4	3	3	All levels present.
Quarter (Binary)	4	3	3	All levels present.
Age Group	5	4	4	All levels present.
Race	5	4	4	All levels present.
Hispanicity	2	1	1	All levels present.
Gender	2	1	1	All levels present.
Two-Factor Effects		**163**	**163**	
Age × Race (3 Levels)	5 × 3	8	8	All levels present.
Age × Hispanicity	5 × 2	4	4	All levels present.
Age × Gender	5 × 2	4	4	All levels present.
Race (3 Levels) × Hispanicity	3 × 2	2	2	All levels present.
Race (3 Levels) × Gender	3 × 2	2	2	All levels present.
Hispanicity × Gender	2 × 2	1	1	All levels present.
State × Age	12 × 5	44	44	All levels present.
State × Race	12 × 5	44	44	All levels present.
State × Gender	12 × 2	11	11	All levels present.
State × Hispanicity	12 × 2	11	11	All levels present.
% Black or African American × % Owner-Occupied	3 × 3	4	4	All levels present.
% Black or African American × Rent/Housing	3 × 5	8	8	All levels present.
% Hispanicity × % Owner-Occupied	3 × 3	4	4	All levels present.
% Hispanicity × Rent/Housing	3 × 5	8	8	All levels present.
% Owner-Occupied × Rent/Housing	3 × 5	8	8	All levels present.
Three-Factor Effects		**71**	**71**	
Race (3 Levels) × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Gender	4 × 5 × 2	12	12	All levels present.
State/Region × Age × Hispanicity	4 × 5 × 2	12	12	All levels present.
State/Region × Age × Race (3 Levels)	4 × 5 × 3	24	24	All levels present.
State/Region × Hispanicity × Gender	4 × 2 × 2	3	3	All levels present.
State/Region × Race (3 Levels) × Hispanicity	4 × 3 × 2	6	6	All levels present.
State/Region × Race (3 Levels) × Gender	4 × 3 × 2	6	6	All levels present.
Total		**300**	**299**	

Exhibit C.2.3 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 2: Midwest

This step used the same set of effects as the respondent questionnaire dwelling unit-level nonresponse.

Appendix C.3: Model Group 3: South

(Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, West Virginia)

This page intentionally left blank

Table C.3a 2017 QDU Weight GEM Modeling Summary (Model Group 3: South)

Modeling Step[1]	Extreme Weight Proportions			UWE[2]	# Covariates[3]	Bounds[4]	
	% Unweighted	% Weighted	% Outwinsor			Nominal	Realized
sel.qdu.ps	0.99	1.77	0.40	1.9032	339	(0.79, 2.00)	(0.80, 2.00)
	0.93	1.72	0.28	1.9182	338	(0.57, 2.00)	(0.57, 1.91)
						(0.90, 1.22)	(0.90, 1.22)
res.qdu.nr	0.96	1.61	0.30	1.9313	339	(1.00, 1.40)	(1.00, 1.40)
	0.83	1.46	0.21	2.0423	339	(1.00, 4.65)	(1.00, 4.62)
						(1.30, 1.47)	(1.30, 1.47)
res.qdu.ps	0.83	1.46	0.21	2.0423	339	(0.98, 2.09)	(0.99, 2.09)
	0.76	1.44	0.13	2.0403	339	(0.42, 2.09)	(0.44, 1.26)
						(0.98, 1.35)	(0.99, 1.34)

GEM = generalized exponential model; QDU = questionnaire dwelling unit.

[1] For a key to modeling abbreviations, see Chapter 6, Exhibit 6.1.

[2] Unequal weighting effect (UWE) defined as $1 + \left[(n-1)/n\right]*CV^2$, where CV = coefficient of variation of weights.

[3] Number of proposed covariates on top line and number finalized after modeling.

[4] There are six sets of bounds for each modeling step. Nominal bounds are used in defining maximum/minimum values for the GEM adjustment factors. The realized bound is the actual adjustment produced by the modeling. The set of three bounds listed for each step correspond to the high extreme values, the nonextreme values, and the low extreme values.

Table C.3b 2017 Distribution of Weight Adjustment Factors and Weight Products (Model Group 3: South)

	SDU Weight	QDU Design Weight		sel.qdu.ps[1]		res.qdu.nr[1]		res.qdu.ps[1]	
Statistics	**1–11**	**duwght12**	**1–12**	**duwght13**	**1–13**	**duwght14**	**1–14**	**duwght15**	**1–15**
Minimum	7	1.00	7	0.34	12	0.62	17	0.44	17
1%	48	1.00	75	0.78	70	1.00	82	0.98	83
5%	135	1.00	235	0.85	237	1.07	312	0.99	312
10%	277	1.00	427	0.89	425	1.11	549	0.99	551
25%	708	1.00	907	0.94	907	1.20	1,128	1.00	1,129
Median	1,038	1.29	1,495	1.00	1,498	1.32	1,911	1.00	1,912
75%	1,432	2.23	2,521	1.07	2,542	1.45	3,310	1.00	3,310
90%	1,887	4.22	4,510	1.14	4,569	1.59	6,387	1.01	6,397
95%	2,190	5.49	6,247	1.19	6,250	1.68	8,872	1.01	8,867
99%	2,884	8.82	10,175	1.33	10,115	1.97	14,352	1.03	14,316
Maximum	9,287	11.06	25,865	3.89	28,822	4.62	28,607	1.78	28,599
n	22,628	-	22,628	-	22,628	-	16,901	-	16,901
Mean	1,098	2.02	2,093	1.01	2,108	1.34	2,822	1.00	2,822
Max/Mean	8	-	12	-	14	-	10	-	10

QDU = questionnaire dwelling unit; SDU = screener dwelling unit.

[1] For a key to modeling abbreviations, see Chapter 6, Exhibit 6.1.

Model Group 3 Overview

Selected Questionnaire Dwelling Unit-Level Poststratification

Out of 339 proposed effects, 338 were kept in the model. The American Indian or Alaska Native and Asian Race categories were combined because of small sample sizes for West Virginia.

Respondent Questionnaire Dwelling Unit-Level Nonresponse

All 339 proposed effects were kept in the model.

Respondent Questionnaire Dwelling Unit-Level Poststratification

This step used the same set of 339 effects as the respondent questionnaire dwelling unit-level nonresponse.

Exhibit C.3.1 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (sel.qdu.ps) Model Group 3: South

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		**76**	**76**	
Intercept	1	1	1	All levels present.
Group Quarter	3	2	2	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Household Type	7	6	6	All levels present.
Household Size	1	1	1	All levels present.
Rent/Housing	5	4	4	All levels present.
Population Density	4	3	3	All levels present.
% Black or African American	3	2	2	All levels present.
% Hispanic or Latino	35	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
State (Count)	17	16	16	All levels present.
State (Binary)	17	16	16	All levels present.
Quarter (Count)	4	3	3	All levels present.
Quarter (Binary)	4	3	3	All levels present.
Age Group	5	4	4	All levels present.
Race	5	4	4	All levels present.
Hispanicity	2	1	1	All levels present.
Gender	2	1	1	All levels present.
Two-Factor Effects		**213**	**212**	
Age × Race (3 Levels)	5 × 3	8	8	All levels present.
Age × Hispanicity	5 × 2	4	4	All levels present.
Age × Gender	5 × 2	4	4	All levels present.
Race (3 Levels) × Hispanicity	3 × 2	2	2	All levels present.
Race (3 Levels) × Gender	3 × 2	2	2	All levels present.
Hispanicity × Gender	2 × 2	1	1	All levels present.
State × Age	17 × 5	64	64	All levels present.
State × Race	17 × 5	64	63	Coll. (17,3) & (17,4); conv.
State × Gender	17 × 2	16	16	All levels present.
State × Hispanicity	17 × 2	16	16	All levels present.
% Black or African American × % Owner-Occupied	3 × 3	4	4	All levels present.
% Black or African American × Rent/Housing	3 × 5	8	8	All levels present.
% Hispanicity × % Owner-Occupied	3 × 3	4	4	All levels present.
% Hispanicity × Rent/Housing	3 × 5	8	8	All levels present.
% Owner-Occupied × Rent/Housing	3 × 5	8	8	All levels present.
Three-Factor Effects		**50**	**50**	
Race (3 Levels) × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Hispanicity	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Race (3 Levels)	3 × 5 × 3	16	16	All levels present.
State/Region × Hispanicity × Gender	3 × 2 × 2	2	2	All levels present.
State/Region × Race (3 Levels) × Hispanicity	3 × 3 × 2	4	4	All levels present.
State/Region × Race (3 Levels) × Gender	3 × 3 × 2	4	4	All levels present.
Total		**339**	**338**	

Exhibit C.3.2 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 3: South

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		**76**	**76**	
Intercept	1	1	1	All levels present.
Group Quarter	3	2	2	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Household Type	7	6	6	All levels present.
Household Size	1	1	1	All levels present.
Rent/Housing	5	4	4	All levels present.
Population Density	4	3	3	All levels present.
% Black or African American	3	2	2	All levels present.
% Hispanic or Latino	35	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
State (Count)	17	16	16	All levels present.
State (Binary)	17	16	16	All levels present.
Quarter (Count)	4	3	3	All levels present.
Quarter (Binary)	4	3	3	All levels present.
Age Group	5	4	4	All levels present.
Race	5	4	4	All levels present.
Hispanicity	2	1	1	All levels present.
Gender	2	1	1	All levels present.
Two-Factor Effects		**213**	**213**	
Age × Race (3 Levels)	5 × 3	8	8	All levels present.
Age × Hispanicity	5 × 2	4	4	All levels present.
Age × Gender	5 × 2	4	4	All levels present.
Race (3 Levels) × Hispanicity	3 × 2	2	2	All levels present.
Race (3 Levels) × Gender	3 × 2	2	2	All levels present.
Hispanicity × Gender	2 × 2	1	1	All levels present.
State × Age	17 × 5	64	64	All levels present.
State × Race	17 × 5	64	64	All levels present.
State × Gender	17 × 2	16	16	All levels present.
State × Hispanicity	17 × 2	16	16	All levels present.
% Black or African American × % Owner-Occupied	3 × 3	4	4	All levels present.
% Black or African American × Rent/Housing	3 × 5	8	8	All levels present.
% Hispanicity × % Owner-Occupied	3 × 3	4	4	All levels present.
% Hispanicity × Rent/Housing	3 × 5	8	8	All levels present.
% Owner-Occupied × Rent/Housing	3 × 5	8	8	All levels present.
Three-Factor Effects		**50**	**50**	
Race (3 Levels) × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Hispanicity	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Race (3 Levels)	3 × 5 × 3	16	16	All levels present.
State/Region × Hispanicity × Gender	3 × 2 × 2	2	2	All levels present.
State/Region × Race (3 Levels) × Hispanicity	3 × 3 × 2	4	4	All levels present.
State/Region × Race (3 Levels) × Gender	3 × 3 × 2	4	4	All levels present.
Total		**339**	**339**	

Exhibit C.3.3 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 3: South

This step used the same set of effects as the respondent questionnaire dwelling unit-level nonresponse.

Appendix C.4: Model Group 4: West

(Alaska, Arizona, California, Colorado, Idaho, Hawaii, Montana, Nevada, New Mexico, Oregon, Utah, Washington, Wyoming)

This page intentionally left blank

Table C.4a 2017 QDU Weight GEM Modeling Summary (Model Group 4: West)

Modeling Step[1]	Extreme Weight Proportions			UWE[2]	# Covariates[3]	Bounds[4]	
	% Unweighted	% Weighted	% Outwinsor			Nominal	Realized
sel.qdu.ps	1.54	3.12	0.64	2.1769	270	(0.74, 1.80)	(0.75, 1.80)
	1.69	3.58	0.65	2.2225	265	(0.42, 2.19)	(0.51, 2.19)
						(0.90, 1.19)	(0.90, 1.18)
res.qdu.nr	1.70	3.72	0.66	2.3379	270	(1.00, 2.10)	(1.00, 2.10)
	1.15	3.16	0.39	2.4443	264	(1.00, 2.45)	(1.00, 2.45)
						(1.40, 1.53)	(1.40, 1.53)
res.qdu.ps	1.15	3.16	0.39	2.4443	270	(0.99, 1.52)	(0.99, 1.51)
	1.11	2.92	0.25	2.4387	265	(0.89, 1.52)	(0.91, 1.12)
						(0.94, 1.01)	(0.94, 1.00)

GEM = generalized exponential model; QDU = questionnaire dwelling unit.

[1] For a key to modeling abbreviations, see Chapter 6, Exhibit 6.1.

[2] Unequal weighting effect (UWE) defined as $1 + \left[(n-1)/n\right]*CV^2$, where CV = coefficient of variation of weights.

[3] Number of proposed covariates on top line and number finalized after modeling.

[4] There are six sets of bounds for each modeling step. Nominal bounds are used in defining maximum/minimum values for the GEM adjustment factors. The realized bound is the actual adjustment produced by the modeling. The set of three bounds listed for each step correspond to the high extreme values, the nonextreme values, and the low extreme values.

Table C.4b 2017 Distribution of Weight Adjustment Factors and Weight Products (Model Group 4: West)

	SDU Weight	QDU Design Weight		sel.qdu.ps[1]		res.qdu.nr[1]		res.qdu.ps[1]	
Statistics	**1–11**	**duwght12**	**1–12**	**duwght13**	**1–13**	**duwght14**	**1–14**	**duwght15**	**1–15**
Minimum	21	1.00	21	0.38	19	0.75	21	0.76	22
1%	84	1.00	92	0.69	90	1.00	103	0.96	103
5%	117	1.00	146	0.82	143	1.07	169	0.99	170
10%	143	1.00	216	0.87	213	1.12	256	0.99	255
25%	311	1.00	480	0.93	476	1.21	566	1.00	566
Median	869	1.27	1,269	1.00	1,258	1.34	1,510	1.00	1,509
75%	1,494	1.93	2,201	1.08	2,243	1.48	3,033	1.00	3,029
90%	1,837	3.50	3,928	1.16	3,932	1.62	5,483	1.01	5,473
95%	2,114	5.42	5,353	1.23	5,408	1.71	8,045	1.01	8,032
99%	2,967	7.95	9,563	1.42	9,605	1.94	14,320	1.04	14,330
Maximum	8,503	10.84	20,816	2.19	24,390	2.45	43,700	2.22	42,582
n	15,942	-	15,942	-	15,942	-	11,752	-	11,752
Mean	970	1.88	1,749	1.01	1,764	1.36	2,314	1.00	2,393
Max/Mean	9	-	12	-	14	-	18	-	18

QDU = questionnaire dwelling unit; SDU = screener dwelling unit.

[1] For a key to modeling abbreviations, see Chapter 6, Exhibit 6.1.

Model Group 4 Overview

Selected Questionnaire Dwelling Unit-Level Poststratification

Out of 270 proposed effects, 265 were kept in the model. All main effects were maintained in full. Two-factor effects were modified for percent Black or African American × Rent/Housing, combining 50–100 percent and 10–<50 percent for levels 1, 3, and 4 of Rent/Housing. Also combined were 50–100 percent and 10–<50 percent Black or African American × 0–<10 percent and 10–<50 percent Owner-Occupied.

Respondent Questionnaire Dwelling Unit-Level Nonresponse

This step used the same set of effects as the selected questionnaire dwelling unit-level poststratification, except that main effect Group Quarter combined categories College Dorm and Other Group Quarters, resulting in 264 effects in the final model.

Respondent Questionnaire Dwelling Unit-Level Poststratification

This step used the same set of 265 effects as the selected questionnaire dwelling unit-level poststratification.

Exhibit C.4.1 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (sel.qdu.ps) Model Group 4: West

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		**68**	**68**	
Intercept	1	1	1	All levels present.
Group Quarter	3	2	2	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Household Type	7	6	6	All levels present.
Household Size	1	1	1	All levels present.
Rent/Housing	5	4	4	All levels present.
Population Density	4	3	3	All levels present.
% Black or African American	3	2	2	All levels present.
% Hispanic or Latino	35	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
State (Count)	13	12	12	All levels present.
State (Binary)	13	12	12	All levels present.
Quarter (Count)	4	3	3	All levels present.
Quarter (Binary)	4	3	3	All levels present.
Age Group	5	4	4	All levels present.
Race	5	4	4	All levels present.
Hispanicity	2	1	1	All levels present.
Gender	2	1	1	All levels present.
Two-Factor Effects		**173**	**168**	
Age × Race (3 Levels)	5 × 3	8	8	All levels present.
Age × Hispanicity	5 × 2	4	4	All levels present.
Age × Gender	5 × 2	4	4	All levels present.
Race (3 Levels) × Hispanicity	3 × 2	2	2	All levels present.
Race (3 Levels) × Gender	3 × 2	2	2	All levels present.
Hispanicity × Gender	2 × 2	1	1	All levels present.
State × Age	13 × 5	48	48	All levels present.
State × Race	13 × 5	48	48	All levels present.
State × Gender	13 × 2	12	12	All levels present.
State × Hispanicity	13 × 2	12	12	All levels present.
% Black or African American × % Owner-Occupied	3 × 3	4	2	Coll. (1,2) & (2,2), (1,3) & (2,3); sing.
% Black or African American × Rent/Housing	3 × 5	8	5	Coll. (1,1) & (2,1), (1,4) & (2,4); zero. Coll. (1,3) & (2,3); sing.
% Hispanicity × % Owner-Occupied	3 × 3	4	4	All levels present.
% Hispanicity × Rent/Housing	3 × 5	8	8	All levels present.
% Owner-Occupied × Rent/Housing	3 × 5	8	8	All levels present.
Three-Factor Effects		**29**	**29**	
Race (3 Levels) × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Gender	2 × 5 × 2	4	4	All levels present.
State/Region × Age × Hispanicity	2 × 5 × 2	4	4	All levels present.
State/Region × Age × Race (3 Levels)	2 × 5 × 3	8	8	All levels present.
State/Region × Hispanicity × Gender	2 × 2 × 2	1	1	All levels present.
State/Region × Race (3 Levels) × Hispanicity	2 × 3 × 2	2	2	All levels present.
State/Region × Race (3 Levels) × Gender	2 × 3 × 2	2	2	All levels present.
Total		**270**	**265**	

Exhibit C.4.2 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 4: West

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		**68**	**67**	
Intercept	1	1	1	All levels present.
Group Quarter	3	2	1	Coll. (1) & (2); conv.
Race/Ethnicity of Householder	6	5	5	All levels present.
Household Type	7	6	6	All levels present.
Household Size	1	1	1	All levels present.
Rent/Housing	5	4	4	All levels present.
Population Density	4	3	3	All levels present.
% Black or African American	3	2	2	All levels present.
% Hispanic or Latino	35	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
State (Count)	13	12	12	All levels present.
State (Binary)	13	12	12	All levels present.
Quarter (Count)	4	3	3	All levels present.
Quarter (Binary)	4	3	3	All levels present.
Age Group	5	4	4	All levels present.
Race	5	4	4	All levels present.
Hispanicity	2	1	1	All levels present.
Gender	2	1	1	All levels present.
Two-Factor Effects		**173**	**168**	
Age × Race (3 Levels)	5 × 3	8	8	All levels present.
Age × Hispanicity	5 × 2	4	4	All levels present.
Age × Gender	5 × 2	4	4	All levels present.
Race (3 Levels) × Hispanicity	3 × 2	2	2	All levels present.
Race (3 Levels) × Gender	3 × 2	2	2	All levels present.
Hispanicity × Gender	2 × 2	1	1	All levels present.
State × Age	13 × 5	48	48	All levels present.
State × Race	13 × 5	48	48	All levels present.
State × Gender	13 × 2	12	12	All levels present.
State × Hispanicity	13 × 2	12	12	All levels present.
% Black or African American × % Owner-Occupied	3 × 3	4	2	Coll. (1,2) & (2,2), (1,3) & (2,3); sing.
% Black or African American × Rent/Housing	3 × 5	8	5	Coll. (1,1) & (2,1), (1,4) & (2,4); zero. Coll. (1,3) & (2,3); sing.
% Hispanicity × % Owner-Occupied	3 × 3	4	4	All levels present.
% Hispanicity × Rent/Housing	3 × 5	8	8	All levels present.
% Owner-Occupied × Rent/Housing	3 × 5	8	8	All levels present.
Three-Factor Effects		**29**	**29**	
Race (3 Levels) × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Gender	2 × 5 × 2	4	4	All levels present.
State/Region × Age × Hispanicity	2 × 5 × 2	4	4	All levels present.
State/Region × Age × Race (3 Levels)	2 × 5 × 3	8	8	All levels present.
State/Region × Hispanicity × Gender	2 × 2 × 2	1	1	All levels present.
State/Region × Race (3 Levels) × Hispanicity	2 × 3 × 2	2	2	All levels present.
State/Region × Race (3 Levels) × Gender	2 × 3 × 2	2	2	All levels present.
Total		**270**	**264**	

Exhibit C.4.3 Covariates for 2017 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 4: West

This step used the same set of covariates as the selected questionnaire dwelling unit-level poststratification.

Appendix D: Evaluation of Calibration Weights: Questionnaire Dwelling Unit-Level Response Rates

This page intentionally left blank

Table D.1 2017 NSDUH QDU-Level Response Rates

Domain	Selected QDU	Respondent QDU	% Interview Response Rate[1]
Total	68,889	50,328	70.77
Census Region			
Northeast	14,037	9,915	67.78
South	22,628	16,901	72.48
Midwest	16,282	11,760	70.53
West	15,942	11,752	70.44
Quarter			
Quarter 1	16,000	11,574	70.43
Quarter 2	17,868	13,117	71.13
Quarter 3	17,303	12,705	71.17
Quarter 4	17,718	12,932	70.32
Household Type			
12–17, 18–25, 26+	5,095	4,077	79.98
12–17, 18–25	73	55	73.60
12–17, 26+	15,284	11,909	77.67
18–25, 26+	11,623	8,450	72.18
12–17	15	12	90.81
18–25	5,487	4,222	76.61
26+	31,312	21,603	68.52
Race/Ethnicity of Householder			
Hispanic or Latino White	9,301	6,961	71.80
Hispanic or Latino Black or African American	227	170	71.52
Hispanic or Latino Other	688	520	75.04
Non-Hispanic or Latino White	45,004	32,241	69.66
Non-Hispanic or Latino Black or African American	7,837	6,176	76.42
Non-Hispanic or Latino Other	5,832	4,260	68.82
% Hispanic or Latino in Segment			
50–100%	5,106	3,827	72.19
10–<50%	16,930	12,394	71.09
<10%	46,853	34,107	70.44
% Black or African American in Segment			
50–100%	4,785	3,758	75.35
10–<50%	13,178	9,806	72.15
<10%	50,926	36,764	69.90
% Owner-Occupied DUs in Segment			
50–100%	50,545	36,589	70.04
10–<50%	14,586	10,851	71.97
<10%	3,758	2,888	75.59
Combined Median Rent/Housing Value			
1st Quintile	11,337	8,642	73.41
2nd Quintile	15,538	11,674	73.22
3rd Quintile	16,164	11,843	71.30
4th Quintile	14,570	10,375	68.99
5th Quintile	11,280	7,794	67.26
Population Density			
Large MSA	29,928	21,260	68.95
Medium to Small MSA	33,344	24,896	72.88
Non-MSA, Urban	1,952	1,499	73.76
Non-MSA, Rural	3,665	2,673	71.63
Group Quarters			
Group	517	483	92.62
Non-Group	68,372	49,845	70.66
Household Size			
One	9,307	6,817	71.94
Two	28,830	20,436	68.69
Three	16,922	12,458	71.27
Four or More	13,830	10,617	75.52

DU = dwelling unit; MSA = metropolitan statistical area; QDU = questionnaire dwelling unit; SDU = screener dwelling unit.

[1] The weight used for calculating the response rate includes SDU- and QDU-level design weights, SDU nonresponse and poststratification adjustments, and selected QDU poststratification adjustment. This weight is the product of WT1*...*WT11*DUWT12*DUWT13.

This page intentionally left blank

Appendix E: Evaluation of Calibration Weights: Questionnaire Dwelling Unit-Level Proportions of Extreme Values and Outwinsors

This page intentionally left blank

Table E.1 2017 NSDUH Selected QDU-Level Proportions of Extreme Values and Outwinsors

Domain	*n*	SDU-Level Weights[1] (SDUWT: WT1*...*WT11)			Before sel.qdu.ps[1] (SDUWT*DUWT12)			After sel.qdu.ps[1] (SDUWT*DUWT12*DUWT13)		
		% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]
Total	68,889	1.99	4.58	1.17	1.58	2.97	0.71	1.41	2.67	0.45
Census Region										
Northeast	14,037	2.16	6.54	1.93	2.22	5.42	1.49	1.67	4.16	0.88
South	22,628	1.41	3.09	0.73	0.99	1.77	0.40	0.93	1.72	0.28
Midwest	16,282	2.67	5.47	1.33	1.90	2.92	0.66	1.57	2.16	0.21
West	15,942	1.96	4.79	1.19	1.54	3.12	0.64	1.69	3.58	0.65
Quarter										
Quarter 1	16,000	2.69	5.90	1.44	2.13	3.46	0.77	1.73	2.77	0.44
Quarter 2	17,868	1.89	4.68	1.22	1.44	3.14	0.70	1.36	3.09	0.52
Quarter 3	17,303	1.95	4.27	1.09	1.51	2.93	0.74	1.36	2.75	0.47
Quarter 4	17,718	1.50	3.47	0.93	1.30	2.35	0.62	1.21	2.05	0.38
Household Type										
12–17, 18–25, 26+	5,095	1.86	4.80	1.16	1.86	4.80	1.16	1.71	4.28	0.85
12–17, 18–25	73	2.74	2.22	0.16	0.00	0.00	0.00	0.00	0.00	0.00
12–17, 26+	15,284	2.19	5.28	1.37	2.19	5.29	1.38	2.07	4.89	0.97
18–25, 26+	11,623	2.18	5.33	1.57	2.09	5.09	1.43	1.58	4.13	0.92
12–17	15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18–25	5,487	2.06	4.43	0.91	2.02	4.12	0.82	2.21	4.47	0.58
26+	31,312	1.83	3.95	0.96	0.98	2.08	0.45	0.83	1.87	0.27
Race/Ethnicity of Householder										
Hispanic or Latino White	9,301	1.25	2.69	0.73	1.04	2.30	0.51	1.06	2.15	0.38
Hispanic or Latino Black or African American	227	55.07	80.80	34.67	44.05	60.19	26.06	52.42	65.85	21.41
Hispanic or Latino Other	688	19.91	43.98	14.09	12.94	25.26	7.58	9.88	25.15	5.17
Non-Hispanic or Latino White	45,004	0.90	1.88	0.35	0.90	1.69	0.29	0.77	1.35	0.13
Non-Hispanic or Latino Black or African American	7,837	4.24	8.19	1.70	2.87	4.41	1.02	2.18	3.56	0.51
Non-Hispanic or Latino Other	5,832	4.39	8.64	1.82	3.02	4.76	0.99	2.85	4.60	0.71

Table E.1 2017 NSDUH Selected QDU-Level Proportions of Extreme Values and Outwinsors (continued)

Domain	*n*	SDU-Level Weights[1] (SDUWT: WT1*...*WT11)			Before sel.qdu.ps[1] (SDUWT*DUWT12)			After sel.qdu.ps[1] (SDUWT*DUWT12*DUWT13)		
		% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]
% Hispanic or Latino in Segment										
50–100%	5,106	1.86	4.87	1.65	1.49	3.06	0.93	1.63	3.56	0.97
10–<50%	16,930	2.38	5.62	1.62	1.88	3.96	1.08	1.74	3.85	0.76
<10%	46,853	1.86	4.03	0.87	1.49	2.52	0.51	1.27	2.03	0.25
% Black or African American in Segment										
50–100%	4,785	3.55	8.30	2.31	2.88	5.78	1.90	2.51	4.61	0.89
10–<50%	13,178	3.04	6.53	1.66	2.00	3.39	0.81	1.93	3.27	0.57
<10%	50,926	1.57	3.59	0.90	1.35	2.55	0.55	1.17	2.29	0.38
% Owner-Occupied DUs in Segment										
50–100%	50,545	1.41	3.08	0.74	1.25	2.18	0.48	1.04	1.82	0.32
10–<50%	14,586	3.32	8.07	2.23	2.41	5.10	1.28	2.10	4.65	0.78
<10%	3,758	4.68	9.40	2.26	2.79	4.89	1.34	3.73	6.05	0.93
Combined Median Rent/Housing Value										
1st Quintile	11,337	1.55	3.57	0.96	1.24	2.13	0.53	1.13	1.91	0.37
2nd Quintile	15,538	1.81	3.73	0.92	1.39	2.01	0.48	1.10	1.91	0.38
3rd Quintile	16,164	2.21	5.62	1.59	1.73	3.35	0.88	1.60	3.15	0.51
4th Quintile	14,570	1.95	4.44	1.07	1.62	3.46	0.73	1.41	3.06	0.51
5th Quintile	11,280	2.41	5.17	1.20	1.92	3.71	0.87	1.84	3.07	0.47
Population Density										
Large MSA[1]	29,928	2.50	5.59	1.44	2.09	3.97	0.96	1.84	3.52	0.63
Medium to Small MSA[1]	33,344	1.67	3.60	0.90	1.22	1.97	0.45	1.14	1.80	0.27
Non-MSA,[1] Urban	1,952	1.08	1.94	0.51	1.02	1.16	0.27	0.77	1.26	0.28
Non-MSA,[1] Rural	3,665	1.26	1.50	0.30	1.09	0.92	0.13	0.71	0.90	0.09
Group Quarters										
Group	517	3.29	6.88	1.12	2.90	3.48	0.30	2.13	3.65	0.66
Non-Group	68,372	1.98	4.57	1.17	1.57	2.96	0.71	1.40	2.66	0.45
Household Size										
One	9,307	1.75	4.14	0.98	0.56	0.94	0.20	0.80	1.43	0.17
Two	28,830	1.85	4.00	1.01	1.43	2.98	0.71	1.24	2.32	0.37
Three	16,922	1.94	4.61	1.26	1.83	4.36	1.06	1.60	3.94	0.69
Four or More	13,830	2.51	5.92	1.47	2.28	5.62	1.35	1.95	5.10	1.11

[1] DU = dwelling unit, MSA = metropolitan statistical area, ps = poststratification adjustment, QDU = questionnaire dwelling unit, SDU = screener dwelling unit, Sel = selected.
[2] Weighted extreme value proportion: $100*\sum_k w_{ek}/\sum_k w_k$, where w_{ek} denotes the weight for extreme values, and w_k denotes the weight for both extreme values and nonextreme values.
[3] Outwinsor weight proportion: $100*\sum_k(w_{ek} - b_k)/\sum_k w_k$, where b_k denotes the winsorized weight.

Table E.2 2017 NSDUH Respondent QDU-Level Proportions of Extreme Values and Outwinsors

Domain	*n*	Before res.qdu.nr[1] (SDUWT*DUWT12*DUWT13)			After res.qdu.nr[1] (SDUWT*DUWT12*...*DUWT14)			Final Weight: After res.qdu.ps[1] (SDUWT*DUWT12*...*DUWT15)		
		% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]
Total	50,328	1.32	2.38	0.43	1.09	2.38	0.35	1.04	2.24	0.22
Census Region										
Northeast	9,915	1.53	3.42	0.77	1.44	4.35	0.70	1.34	3.98	0.43
South	16,901	0.96	1.61	0.30	0.83	1.46	0.21	0.76	1.44	0.13
Midwest	11,760	1.29	1.58	0.14	1.10	1.59	0.28	1.11	1.53	0.19
West	11,752	1.70	3.72	0.66	1.15	3.16	0.39	1.11	2.92	0.25
Quarter										
Quarter 1	11,574	1.64	2.54	0.44	1.26	2.21	0.33	1.21	2.23	0.20
Quarter 2	13,117	1.27	2.75	0.50	1.06	2.81	0.40	1.02	2.66	0.24
Quarter 3	12,705	1.33	2.54	0.40	1.01	2.41	0.36	0.97	2.16	0.23
Quarter 4	12,932	1.08	1.69	0.37	1.04	2.09	0.32	0.98	1.90	0.22
Household Type										
12–17, 18–25, 26+	4,077	1.82	4.62	0.84	1.59	4.70	0.84	1.67	4.86	0.68
12–17, 18–25	55	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12–17, 26+	11,909	2.07	4.88	0.95	1.55	4.02	0.74	1.52	3.81	0.53
18–25, 26+	8,450	1.81	4.50	1.04	1.49	4.01	0.56	1.41	3.87	0.42
12–17	12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18–25	4,222	2.25	4.56	0.58	1.56	3.94	0.49	1.23	3.04	0.19
26+	21,603	0.45	1.27	0.19	0.50	1.62	0.22	0.48	1.53	0.12
Race/Ethnicity of Householder										
Hispanic or Latino White	6,961	1.01	1.90	0.42	0.88	2.32	0.34	0.86	2.25	0.18
Hispanic or Latino Black or African American	170	52.35	67.42	19.84	49.41	62.38	11.27	52.94	64.23	9.83
Hispanic or Latino Other	520	10.19	22.71	5.19	8.85	19.81	3.62	9.62	22.01	3.06
Non-Hispanic or Latino White	32,241	0.64	1.04	0.10	0.50	1.05	0.10	0.38	0.80	0.03
Non-Hispanic or Latino Black or African American	6,176	2.06	3.15	0.52	1.25	2.44	0.42	1.18	2.24	0.20
Non-Hispanic or Latino Other	4,260	2.84	4.30	0.56	2.79	6.05	0.98	3.03	6.38	0.71
% Hispanic or Latino in Segment										
50–100%	3,827	1.70	3.93	1.08	1.52	3.78	0.54	1.54	3.86	0.50
10–<50%	12,394	1.65	3.43	0.70	1.50	3.92	0.59	1.52	3.81	0.38
<10%	34,107	1.16	1.72	0.22	0.89	1.52	0.22	0.81	1.35	0.12

Table E.2 2017 NSDUH Respondent QDU-Level Proportions of Extreme Values and Outwinsors (continued)

Domain	*n*	Before res.qdu.nr[1] (SDUWT*DUWT12*DUWT13)			After res.qdu.nr[1] (SDUWT*DUWT12*...*DUWT14)			Final Weight: After res.qdu.ps[1] (SDUWT*DUWT12*...*DUWT15)		
		% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]
% Black or African American in Segment										
50–100%	3,758	2.13	3.55	0.73	1.81	4.10	0.79	1.78	3.99	0.51
10–<50%	9,806	1.74	3.06	0.56	1.42	2.61	0.38	1.40	2.68	0.27
<10%	36,764	1.13	2.05	0.35	0.93	2.14	0.30	0.87	1.93	0.18
% Owner-Occupied DUs in Segment										
50–100%	36,589	0.98	1.60	0.31	0.87	1.64	0.25	0.81	1.54	0.16
10–<50%	10,851	1.88	4.04	0.66	1.52	4.33	0.63	1.57	4.14	0.40
<10%	2,888	3.57	5.66	0.99	2.18	4.37	0.65	2.01	4.02	0.38
Combined Median Rent/Housing Value										
1st Quintile	8,642	1.13	1.92	0.41	0.72	1.23	0.22	0.72	1.33	0.17
2nd Quintile	11,674	1.11	1.77	0.37	0.80	1.41	0.23	0.82	1.43	0.17
3rd Quintile	11,843	1.43	2.85	0.50	1.04	2.45	0.41	0.91	2.24	0.31
4th Quintile	10,375	1.27	2.61	0.44	1.39	3.27	0.44	1.31	3.03	0.22
5th Quintile	7,794	1.77	2.66	0.41	1.62	3.27	0.43	1.57	2.98	0.21
Population Density										
Large MSA[1]	21,260	1.73	3.22	0.60	1.67	3.50	0.49	1.59	3.22	0.30
Medium to Small MSA[1]	24,896	1.11	1.59	0.25	0.70	1.12	0.19	0.66	1.12	0.14
Non-MSA,[1] Urban	1,499	0.60	1.27	0.34	0.33	0.38	0.06	0.33	0.69	0.05
Non-MSA,[1] Rural	2,673	0.49	0.50	0.08	0.56	1.48	0.26	0.56	1.44	0.22
Group Quarters										
Group	483	2.07	3.16	0.59	0.21	0.06	0.00	0.21	0.09	0.03
Non-Group	49,845	1.32	2.38	0.43	1.10	2.39	0.35	1.05	2.25	0.22
Household Size										
One	6,817	0.81	1.29	0.14	0.73	1.70	0.18	0.62	1.56	0.10
Two	20,436	0.92	1.62	0.29	0.79	1.98	0.32	0.72	1.75	0.16
Three	12,458	1.62	3.89	0.73	1.19	2.88	0.48	1.14	2.85	0.35
Four or More	10,617	2.09	5.50	1.14	1.78	4.79	0.71	1.81	4.85	0.58

[1] DU = dwelling unit, MSA = metropolitan statistical area, nr = nonresponse adjustment, ps = poststratification adjustment, QDU = questionnaire dwelling unit, res = Respondent, SDU = screener dwelling unit.

[2] Weighted extreme value proportion: $100*\sum_k w_{ek}/\sum_k w_k$, where w_{ek} denotes the weight for extreme values, and w_k denotes the weight for both extreme values and nonextreme values.

[3] Outwinsor weight proportion: $100*\sum_k(w_{ek} - b_k)/\sum_k w_k$, where b_k denotes the winsorized weight.

Appendix F: Evaluation of Calibration Weights: Questionnaire Dwelling Unit-Level Slippage Rates

This page intentionally left blank

Table F.1 2017 NSDUH QDU-Level Slippage Rates

Domain	*n*	Initial Total (*I*)[1]	Final Total (*F*)[2]	Control from SDU Weights (*C*)	(*I* - *C*)/*C*%	(*F* - *C*)/*C*%
Total	50,328	125,054,683	125,054,683	125,054,683	0.00	0.00
Census Region						
Northeast	9,915	22,036,954	22,036,954	22,036,954	0.00	-0.00
South	16,901	47,691,168	47,691,168	47,691,168	0.00	-0.00
Midwest	11,760	27,199,079	27,199,079	27,199,079	0.00	-0.00
West	11,752	28,127,482	28,127,482	28,127,482	-0.00	0.00
Quarter						
Quarter 1	11,574	31,140,954	31,140,954	31,140,954	0.00	0.00
Quarter 2	13,117	31,269,089	31,269,089	31,269,089	0.00	0.00
Quarter 3	12,705	31,354,775	31,354,775	31,354,775	0.00	0.00
Quarter 4	12,932	31,289,866	31,289,866	31,289,866	0.00	-0.00
Household Type						
12–17, 18–25, 26+	4,077	5,065,701	5,065,701	5,065,701	-0.00	0.00
12–17, 18–25	55	62,484	62,484	62,484	0.00	0.00
12–17, 26+	11,909	13,687,226	13,687,226	13,687,226	0.00	-0.00
18–25, 26+	8,450	13,861,116	13,861,116	13,861,116	0.00	0.00
12–17	12	10,635	10,635	10,635	0.00	0.00
18–25	4,222	5,685,693	5,685,693	5,685,693	0.00	0.00
26+	21,603	86,681,829	86,681,829	86,681,829	0.00	0.00
Race/Ethnicity of Householder						
Hispanic or Latino White	6,961	15,406,870	15,406,870	15,406,870	0.00	0.00
Hispanic or Latino Black or African American	170	908,664	908,664	908,664	-0.00	0.00
Hispanic or Latino Other	520	1,269,330	1,269,330	1,269,330	0.00	0.00
Non-Hispanic or Latino White	32,241	82,936,396	82,936,396	82,936,396	0.00	0.00
Non-Hispanic or Latino Black or African American	6,176	15,453,034	15,453,034	15,453,034	0.00	0.00
Non-Hispanic or Latino Other	4,260	9,080,390	9,080,390	9,080,390	0.00	-0.00
% Hispanic or Latino in Segment						
50–100%	3,827	10,101,214	10,101,214	10,101,214	0.00	0.00
10–<50%	12,394	35,109,146	35,109,146	35,109,146	0.00	0.00
<10%	34,107	79,844,323	79,844,323	79,844,323	0.00	0.00
% Black or African American in Segment						
50–100%	3,758	9,192,365	9,192,365	9,192,365	0.00	0.00
10–<50%	9,806	26,030,656	26,030,656	26,030,656	0.00	0.00
<10%	36,764	89,831,662	89,831,662	89,831,662	0.00	0.00
% Owner-Occupied DUs in Segment						
50–100%	36,589	90,961,006	90,961,006	90,961,006	0.00	0.00
10–<50%	10,851	27,226,082	27,226,082	27,226,082	0.00	0.00
<10%	2,888	6,867,596	6,867,596	6,867,596	0.00	0.00

(continued)

Table F.1 2017 NSDUH QDU-Level Slippage Rates (continued)

Domain	*n*	Initial Total (*I*)[1]	Final Total (*F*)[2]	Control from SDU Weights (*C*)	(*I* - *C*)/*C*%	(*F* - *C*)/*C*%
Combined Median Rent/Housing Value						
1st Quintile	8,642	18,690,173	18,690,172	18,690,172	0.00	0.00
2nd Quintile	11,674	26,824,257	26,824,257	26,824,257	0.00	0.00
3rd Quintile	11,843	28,887,312	28,887,312	28,887,312	0.00	-0.00
4th Quintile	10,375	27,270,541	27,270,541	27,270,541	0.00	0.00
5th Quintile	7,794	23,382,401	23,382,401	23,382,401	0.00	0.00
Population Density						
Large MSA	21,260	66,435,635	66,435,635	66,435,635	0.00	0.00
Medium to Small MSA	24,896	50,882,590	50,882,590	50,882,590	0.00	-0.00
Non-MSA, Urban	1,499	2,856,022	2,856,022	2,856,022	-0.00	0.00
Non-MSA, Rural	2,673	4,880,436	4,880,436	4,880,436	0.00	0.00
Group Quarters						
Group	483	613,170	613,170	613,170	0.00	0.00
Non-Group	49,845	124,441,514	124,441,513	124,441,513	0.00	0.00
Household Size						
One	6,817	33,751,642	33,756,387	33,073,327	2.05	2.07
Two	20,436	56,800,297	56,779,229	57,711,968	-1.58	-1.62
Three	12,458	19,654,530	19,670,865	19,661,247	-0.03	0.05
Four or More	10,617	14,848,214	14,848,202	14,608,142	1.64	1.64

DU = dwelling unit, MSA = metropolitan statistical area, QDU = questionnaire dwelling unit, SDU = screener dwelling unit.

[1] WT1*...*WT11*DUWT12*...*DUWT14 (before QDU poststratification and QDU extreme value adjustment).

[2] WT1*...*WT11*DUWT12*...*DUWT16 (after QDU poststratification and QDU extreme value adjustment).

Appendix G: Evaluation of Calibration Weights: Questionnaire Dwelling Unit-Level Weight Summary Statistics

This page intentionally left blank

Table G.1 2017 NSDUH Selected QDU-Level Weight Summary Statistics

Domain	*n*	SDU-Level Weights[1] (SDUWT: WT1*...*WT11)						Before sel.qdu.ps[1] (SDUWT*DUWT12)						After sel.qdu.ps[1] (SDUWT*DUWT12*DUWT13)					
		Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
Total	68,889	7	515	904	1,275	9,287	1.43	7	702	1,264	2,170	35,590	2.06	9	697	1,267	2,190	30,620	2.08
Census Region																			
Northeast	14,037	15	279	810	1,048	6,350	1.47	15	532	1,065	1,780	35,590	2.30	9	515	1,062	1,808	30,620	2.32
South	22,628	7	708	1,038	1,432	9,287	1.34	7	907	1,495	2,521	25,865	1.90	12	907	1,498	2,542	28,822	1.92
Midwest	16,282	39	608	857	1,087	5,178	1.28	44	762	1,155	1,958	18,980	1.95	46	761	1,167	1,961	16,930	1.94
West	15,942	21	311	869	1,494	8,503	1.57	21	480	1,269	2,201	20,816	2.18	19	476	1,258	2,243	24,390	2.22
Quarter																			
Quarter 1	16,000	22	565	973	1,365	9,287	1.42	22	781	1,374	2,358	25,865	2.01	17	780	1,386	2,364	24,411	2.02
Quarter 2	17,868	10	517	855	1,221	7,498	1.43	10	676	1,220	2,053	21,317	2.07	9	679	1,225	2,083	24,172	2.11
Quarter 3	17,303	7	477	897	1,293	7,030	1.44	7	661	1,237	2,165	35,590	2.12	12	648	1,250	2,181	30,620	2.14
Quarter 4	17,718	12	533	890	1,237	7,781	1.39	12	715	1,241	2,125	19,295	2.03	13	705	1,235	2,123	24,390	2.03
Household Type																			
12–17, 18–25, 26+	5,095	10	549	940	1,333	7,475	1.41	10	549	940	1,333	7,475	1.41	9	533	937	1,339	5,351	1.41
12–17, 18–25	73	70	420	756	1,233	2,626	1.46	70	420	756	1,233	2,626	1.46	61	458	762	1,214	3,385	1.53
12–17, 26+	15,284	7	447	846	1,205	7,049	1.46	7	448	848	1,209	7,055	1.46	12	440	849	1,224	7,375	1.46
18–25, 26+	11,623	7	607	972	1,369	9,287	1.41	13	679	1,121	1,581	11,229	1.41	21	670	1,131	1,592	8,404	1.40
12–17	15	82	182	796	1,092	1,538	1.46	82	182	804	1,103	1,554	1.46	63	147	819	1,138	1,336	1.44
18–25	5,487	12	414	869	1,224	6,188	1.46	12	460	999	1,433	7,976	1.45	17	450	974	1,418	6,614	1.46
26+	31,312	20	529	911	1,269	7,141	1.40	50	1,193	2,115	3,644	35,590	1.71	34	1,191	2,132	3,658	30,620	1.73
Race/Ethnicity of Householder																			
Hispanic or Latino White	9,301	24	656	1,029	1,436	8,503	1.30	24	782	1,292	1,904	18,534	1.79	19	774	1,310	1,959	19,264	1.85
Hispanic or Latino Black or African American	227	27	781	1,997	3,662	9,287	1.63	27	998	2,382	4,882	35,590	2.30	59	1,256	2,797	4,805	30,620	2.29
Hispanic or Latino Other	688	7	222	746	1,781	8,679	2.17	7	303	1,018	2,426	14,641	2.52	12	296	1,004	2,354	18,532	2.61
Non-Hispanic or Latino White	45,004	16	452	874	1,189	7,781	1.39	16	686	1,238	2,201	25,865	2.10	13	685	1,241	2,217	24,411	2.11
Non-Hispanic or Latino Black or African American	7,837	32	730	1,039	1,400	6,643	1.32	32	909	1,433	2,351	19,295	1.89	20	885	1,422	2,360	17,124	1.91
Non-Hispanic or Latino Other	5,832	12	292	791	1,386	6,327	1.64	12	444	1,119	2,091	29,140	2.07	9	441	1,103	2,084	21,907	2.08
% Hispanic or Latino in Segment																			
50–100%	5,106	46	780	1,235	1,588	8,679	1.27	46	962	1,495	2,222	15,194	1.72	51	967	1,532	2,327	28,822	1.79
10–<50%	16,930	7	668	1,072	1,528	8,503	1.38	7	856	1,503	2,526	35,590	1.94	12	849	1,495	2,542	30,620	1.96
<10%	46,853	12	375	843	1,137	9,287	1.42	12	612	1,154	2,014	29,140	2.14	9	611	1,157	2,028	24,411	2.16

(continued)

Table G.1 2017 NSDUH Selected QDU-Level Weight Summary Statistics (continued)

Domain	*n*	SDU-Level Weights[1] (SDUWT: WT1*...*WT11)						Before sel.qdu.ps[1] (SDUWT*DUWT12)						After sel.qdu.ps[1] (SDUWT*DUWT12*DUWT13)					
		Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
% Black or African American in Segment																			
50–100%	4,785	7	715	999	1,312	9,287	1.36	7	878	1,363	2,260	29,140	1.99	12	855	1,363	2,299	21,907	1.95
10–<50%	13,178	12	698	997	1,372	8,022	1.35	12	869	1,395	2,328	17,265	1.89	17	868	1,407	2,359	24,390	1.91
<10%	50,926	12	419	869	1,243	8,679	1.45	12	634	1,216	2,117	35,590	2.12	9	630	1,214	2,129	30,620	2.14
% Owner-Occupied DUs[1] in Segment																			
50–100%	50,545	7	507	889	1,239	9,287	1.40	7	701	1,246	2,148	29,140	2.04	9	698	1,251	2,164	24,411	2.05
10–<50%	14,586	12	550	957	1,352	8,679	1.46	12	725	1,324	2,221	35,590	2.11	17	714	1,321	2,257	30,620	2.14
<10%	3,758	32	485	955	1,413	8,503	1.52	32	623	1,251	2,177	16,288	2.07	21	618	1,275	2,300	19,393	2.13
Combined Median Rent/Housing Value																			
1st Quintile	11,337	27	374	809	1,113	6,643	1.42	27	580	1,120	1,921	21,317	2.14	17	579	1,129	1,945	18,528	2.16
2nd Quintile	15,538	15	477	868	1,209	7,498	1.40	15	655	1,198	2,043	18,974	2.13	9	636	1,187	2,032	21,763	2.16
3rd Quintile	16,164	12	508	900	1,258	9,287	1.46	12	680	1,247	2,109	25,865	2.04	13	685	1,250	2,144	28,822	2.07
4th Quintile	14,570	7	574	931	1,339	6,441	1.41	7	749	1,312	2,214	29,140	2.05	12	737	1,318	2,272	24,172	2.05
5th Quintile	11,280	10	698	1,056	1,458	8,503	1.38	10	865	1,469	2,533	35,590	1.92	17	882	1,499	2,557	30,620	1.93
Population Density																			
Large MSA[1]	29,928	7	827	1,108	1,514	9,287	1.29	7	1,019	1,590	2,609	35,590	1.83	12	1,020	1,602	2,651	30,620	1.85
Medium to Small MSA[1]	33,344	15	312	742	1,075	7,141	1.47	15	488	1,028	1,797	25,865	2.23	9	485	1,022	1,799	24,411	2.24
Non-MSA,[1] Urban	1,952	24	297	740	998	4,814	1.45	24	470	989	1,720	14,409	2.28	22	454	966	1,680	13,460	2.28
Non-MSA,[1] Rural	3,665	27	193	552	931	3,617	1.57	31	302	834	1,584	15,298	2.43	18	310	827	1,624	15,023	2.41
Group Quarters																			
Group	517	59	196	669	1,082	4,209	1.72	77	310	853	1,420	15,146	2.93	69	293	777	1,447	12,960	2.65
Non-Group	68,372	7	518	905	1,276	9,287	1.42	7	706	1,268	2,174	35,590	2.05	9	701	1,271	2,194	30,620	2.07
Household Size																			
One	9,307	21	455	871	1,197	5,829	1.40	77	1,178	2,520	5,155	35,590	1.80	76	1,161	2,510	5,174	30,620	1.82
Two	28,830	12	513	899	1,247	9,287	1.41	14	828	1,535	2,623	21,317	1.69	19	829	1,535	2,645	18,528	1.70
Three	16,922	12	523	912	1,299	8,022	1.43	12	575	1,012	1,533	10,637	1.58	19	574	1,018	1,538	16,951	1.59
Four or More	13,830	7	541	933	1,361	8,679	1.46	7	552	960	1,415	10,110	1.50	9	533	947	1,428	11,412	1.53

[1] DU = dwelling unit, MSA = metropolitan statistical area, ps = poststratification adjustment, QDU = questionnaire dwelling unit, SDU = screener dwelling unit, sel = selected.

[2] Q1 and Q3 refer to the first and third quartile of the weight distribution.

[3] Unequal weighting effect (UWE) is defined as $1 + [(n - 1)/n]*CV^2$, where CV = coefficient of variation of weights.

Table G.2 2017 NSDUH Respondent QDU-Level Weight Summary Statistics

Domain	*n*	Before res.qdu.nr[1] (SDUWT*DUWT12*DUWT13)						After res.qdu.nr[1] (SDUWT*DUWT12*...*DUWT14)						Final Weight: After res.qdu.ps[1] (SDUWT*DUWT12*...*DUWT15)					
		Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
Total	50,328	9	668	1,224	2,103	28,822	2.11	9	851	1,631	2,979	43,700	2.23	9	850	1,633	2,976	42,582	2.22
Census Region																			
Northeast	9,915	9	478	1,013	1,729	18,263	2.31	9	630	1,370	2,597	42,823	2.52	9	629	1,371	2,598	37,426	2.50
South	16,901	12	893	1,463	2,436	28,822	1.93	17	1,128	1,911	3,310	28,607	2.04	17	1,129	1,912	3,310	28,599	2.04
Midwest	11,760	46	735	1,126	1,899	16,930	1.99	65	965	1,529	2,729	24,564	2.05	64	964	1,534	2,731	24,680	2.05
West	11,752	19	442	1,171	2,111	24,390	2.34	22	566	1,510	3,033	43,700	2.44	22	566	1,509	3,029	42,582	2.44
Quarter																			
Quarter 1	11,574	17	754	1,342	2,305	18,300	2.03	21	954	1,812	3,256	27,800	2.15	20	954	1,817	3,258	27,926	2.15
Quarter 2	13,117	9	658	1,184	2,004	24,172	2.15	9	831	1,557	2,837	42,823	2.27	9	831	1,562	2,847	37,426	2.26
Quarter 3	12,705	12	619	1,202	2,095	28,822	2.19	18	805	1,586	2,922	31,366	2.28	18	805	1,590	2,922	30,790	2.27
Quarter 4	12,932	13	672	1,195	2,028	24,390	2.07	17	845	1,604	2,860	43,700	2.19	17	843	1,605	2,864	42,582	2.19
Household Type																			
12–17, 18–25, 26+	4,077	9	524	938	1,334	5,351	1.42	9	616	1,178	1,673	6,972	1.44	9	615	1,177	1,672	6,468	1.44
12–17, 18–25	55	61	439	760	1,204	3,385	1.55	61	509	1,063	1,620	4,813	1.62	60	503	1,066	1,617	4,912	1.63
12–17, 26+	11,909	12	438	846	1,223	7,375	1.46	17	538	1,080	1,557	10,963	1.49	17	536	1,082	1,559	8,305	1.48
18–25, 26+	8,450	21	647	1,121	1,586	8,404	1.42	22	822	1,526	2,232	8,873	1.44	22	817	1,526	2,234	9,732	1.44
12–17	12	137	322	845	1,226	1,336	1.31	148	334	974	1,318	1,495	1.29	146	337	971	1,318	1,494	1.29
18–25	4,222	17	432	962	1,415	6,614	1.48	18	552	1,234	1,882	6,989	1.49	17	559	1,237	1,885	6,336	1.49
26+	21,603	34	1,153	2,102	3,618	28,822	1.75	36	1,617	3,038	5,369	43,700	1.77	36	1,616	3,038	5,371	42,582	1.76
Race/Ethnicity of Householder																			
Hispanic or Latino White	6,961	34	741	1,274	1,887	19,264	1.85	35	925	1,652	2,630	30,226	2.04	36	924	1,654	2,638	28,688	2.04
Hispanic or Latino Black or African American	170	59	1,273	2,665	4,411	28,822	2.28	79	2,255	3,804	5,759	43,700	2.27	81	2,238	3,700	5,729	42,582	2.25
Hispanic or Latino Other	520	12	299	1,013	2,354	13,389	2.52	17	383	1,242	3,109	20,067	2.71	17	384	1,259	3,010	20,856	2.69
Non-Hispanic or Latino White	32,241	13	653	1,192	2,135	24,172	2.17	17	855	1,639	3,080	30,796	2.23	17	854	1,642	3,079	31,117	2.23
Non-Hispanic or Latino Black or African American	6,176	20	877	1,396	2,264	17,124	1.91	20	1,042	1,706	2,940	26,632	2.09	20	1,043	1,710	2,942	26,605	2.08
Non-Hispanic or Latino Other	4,260	9	403	1,032	1,955	18,299	2.08	9	511	1,328	2,803	42,823	2.39	9	508	1,327	2,809	37,426	2.37
% Hispanic or Latino in Segment																			
50–100%	3,827	51	919	1,491	2,256	28,822	1.80	55	1,179	1,926	3,157	27,811	1.91	58	1,179	1,923	3,156	28,543	1.91
10–<50%	12,394	12	812	1,444	2,445	24,390	2.01	17	1,021	1,908	3,471	43,700	2.13	17	1,022	1,911	3,466	42,582	2.12
<10%	34,107	9	586	1,121	1,948	24,172	2.19	9	761	1,499	2,760	31,366	2.29	9	762	1,502	2,759	31,117	2.29

(continued)

Table G.2 2017 NSDUH Respondent QDU-Level Weight Summary Statistics (continued)

Domain	*n*	Before res.qdu.nr[1] (SDUWT*DUWT12*DUWT13)						After res.qdu.nr[1] (SDUWT*DUWT12*...*DUWT14)						Final Weight: After res.qdu.ps[1] (SDUWT*DUWT12*...*DUWT15)					
		Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
% Black or African American in Segment																			
50–100%	3,758	12	846	1,335	2,213	17,164	1.91	17	1,011	1,658	2,888	42,823	2.15	17	1,014	1,658	2,890	37,426	2.13
10–<50%	9,806	17	847	1,361	2,275	24,390	1.94	20	1,060	1,795	3,136	43,700	2.09	20	1,060	1,799	3,140	42,582	2.09
<10%	36,764	9	598	1,164	2,037	28,822	2.19	9	781	1,575	2,940	31,366	2.27	9	780	1,576	2,940	31,117	2.27
% Owner-Occupied DUs[1] in Segment																			
50–100%	36,589	9	671	1,207	2,075	20,450	2.09	9	869	1,632	2,974	31,366	2.18	9	867	1,634	2,969	30,790	2.18
10–<50%	10,851	17	682	1,277	2,162	28,822	2.18	18	840	1,647	2,999	43,700	2.34	17	838	1,652	2,995	42,582	2.33
<10%	2,888	21	595	1,237	2,242	19,393	2.20	21	699	1,524	2,962	26,632	2.37	20	696	1,532	2,964	26,605	2.37
Combined Median Rent/Housing Value																			
1st Quintile	8,642	17	566	1,097	1,877	18,528	2.16	17	733	1,414	2,507	21,879	2.26	17	733	1,416	2,507	21,823	2.26
2nd Quintile	11,674	9	624	1,162	1,964	21,763	2.20	9	785	1,497	2,679	31,366	2.32	9	784	1,498	2,683	30,790	2.32
3rd Quintile	11,843	13	659	1,215	2,064	28,822	2.11	17	830	1,612	2,892	43,700	2.24	17	831	1,614	2,895	42,582	2.24
4th Quintile	10,375	12	711	1,271	2,183	24,172	2.08	17	936	1,757	3,158	30,796	2.19	17	936	1,761	3,161	31,117	2.18
5th Quintile	7,794	17	838	1,440	2,475	18,299	1.97	18	1,127	2,071	3,718	29,339	2.02	17	1,124	2,076	3,706	29,630	2.02
Population Density																			
Large MSA[1]	21,260	12	989	1,549	2,533	28,822	1.88	17	1,311	2,154	3,730	43,700	1.98	17	1,312	2,156	3,727	42,582	1.97
Medium to Small MSA[1]	24,896	9	474	999	1,746	21,763	2.26	9	617	1,306	2,417	28,607	2.34	9	616	1,306	2,415	28,599	2.34
Non-MSA,[1] Urban	1,499	22	436	950	1,615	12,880	2.34	30	553	1,220	2,147	21,392	2.50	30	552	1,218	2,139	21,473	2.50
Non-MSA,[1] Rural	2,673	18	302	819	1,597	15,023	2.44	18	411	1,101	2,188	31,366	2.63	18	410	1,104	2,182	30,790	2.62
Group Quarters																			
Group	483	69	292	777	1,458	12,960	2.61	69	305	825	1,592	15,087	2.72	63	305	792	1,604	15,037	2.70
Non-Group	49,845	9	674	1,229	2,107	28,822	2.11	9	860	1,640	2,991	43,700	2.22	9	860	1,643	2,988	42,582	2.21
Household Size																			
One	6,817	76	1,128	2,441	5,048	28,822	1.84	79	1,508	3,369	7,239	43,700	1.89	80	1,516	3,375	7,237	42,582	1.89
Two	20,436	19	787	1,480	2,531	18,528	1.73	23	1,023	2,058	3,669	22,572	1.82	23	1,024	2,058	3,666	19,841	1.82
Three	12,458	19	554	993	1,501	16,951	1.59	20	698	1,310	2,054	18,975	1.71	19	699	1,311	2,057	18,993	1.71
Four or More	10,617	9	519	938	1,407	8,576	1.52	9	638	1,208	1,856	13,375	1.62	9	637	1,209	1,859	13,504	1.62

[1] DU = dwelling unit, MSA = metropolitan statistical area, nr = nonresponse adjustment, ps = poststratification adjustment, QDU = questionnaire dwelling unit, res = respondent, SDU = screener dwelling unit, sel = selected.

[2] Q1 and Q3 refer to the first and third quartile of the weight distribution.

[3] Unequal weighting effect (UWE) is defined as $1 + [(n - 1)/n]*CV^2$, where CV = coefficient of variation of weights.

Appendix H: GEM Modeling Summary for the Pair Weights

This page intentionally left blank

Appendix H: GEM Modeling Summary for the Pair Weights

This appendix summarizes each model group throughout all stages of weight calibration modeling. Unlike much of the other information presented in this report, this section provides a model-specific overview of weight calibration, as opposed to a domain-specific one.

For 2017, modeling involved taking two model groups through four adjustment steps: (1) selected pair poststratification, (2) pair nonresponse adjustment, (3) responding pair poststratification, and (4) responding pair extreme value adjustment.

Model-specific summary statistics are shown in Tables H.1a through H.2b. Included in these tables, for each stage of modeling, are the number of factor effects included in the final model; the high, low, and nonextreme weight bounds set to provide the upper and lower limits for the generalized exponential model (GEM) macro; the weighted, unweighted, and winsorized weight proportions; the unequal weighting effect (UWE); and weight distributions. The UWE provides an approximate partial measure of variance and provides a summary of how much impact a particular stage of modeling has on the distribution of the new product of weights. At each stage in the modeling, these summary statistics were calculated and utilized to help evaluate the quality of the weight component under the model chosen.

Occurrences of small sample sizes and exact linear combinations in the realized data led to situations whereby modeling inclusion of all originally proposed levels of covariates in the model was not possible. The text and exhibits in Sections H.1 and H.2 summarize the decisions made with regard to final covariates included in each model. For the list of proposed initial covariates considered at each stage of modeling, see Exhibit H.2. For the list of realized final model covariates, see Exhibits H.1.1 to H.2.4. For guidelines on interpreting these exhibits, see Appendix C.

Final Model Explanatory Variables

For brevity, numeric abbreviations for factor levels are established in Exhibit 4.2 (included here as Exhibit H.1 for easy reference). A complete list of all variables and associated levels used at any stage of modeling is provided. Note that not all factors or levels are present in all stages of modeling, and the initial set of variables is the same across model groups but may change for an adjustment step of modeling. The initial candidates are found in any of the proposed variable columns for a particular stage of weight adjustment.

Exhibit H.1 Definitions of Levels for Pair-Level Calibration Modeling Variables

Group Quarter Indicator

1: College Dorm, 2: Other Group Quarter, 3: Non-Group Quarter[1]

Household Size

2: DU with 2 People,[1] 3: DU with 3 People, 4: DU with ≥ 4 People

Pair Age (15 Levels)

1: 12–17 and 12–17,[1] 2: 12–17 and 18–25, 3: 12–17 and 26–34, 4: 12–17 and 35–49, 5: 12–17 and 50+, 6: 18–25 and 18–25, 7: 18–25 and 26–34, 8: 18–25 and 35–49, 9: 18–25 and 50+, 10: 26–34 and 26–34, 11: 26–34 and 35–49, 12: 26–34 and 50+, 13: 35–49 and 35–49, 14: 35–49 and 50+, 15: 50+ and 50+

Pair Age (6 Levels)

1: 12–17 and 12–17,[1] 2: 12–17 and 18–25, 3: 12–17 and 26+, 4: 18–25 and 18–25, 5: 18–25 and 26+, 6: 26+ and 26+

Pair Age (3 Levels)

1: 12–17 and 12–17,[1] 2: 12–17 and 18+, 3: 18+ and 18+

Pair Gender

1: Male and Female,[1] 2: Female and Female, 3: Male and Male

Pair Race/Ethnicity (10 Levels)

1: White and White,[1] 2: White and Black or African American, 3: White and Hispanic or Latino, 4: White and Other, 5: Black or African American and Black or African American, 6: Black or African American and Hispanic or Latino, 7: Black or African American and Other, 8: Hispanic or Latino and Hispanic or Latino, 9: Hispanic or Latino and Other, 10: Other and Other

Pair Race/Ethnicity (5 Levels)

1: Two or More Races Pair, 2: Hispanic or Latino Pair, 3: Black or African American Pair, 4: White Pair,[1] 5: Other Pair

Pair Race/Ethnicity (4 Levels)

1: Two or More Races Pair or Other and Other, 2: Hispanic or Latino Pair, 3: Black or African American Pair, 4: White Pair[1]

Percentage of Owner-Occupied Dwelling Units in Segment (% Owner-Occupied)

1: 50–100%,[1] 2: 10–<50%, 3: 0–<10%

Percentage of Segments That Are Black or African American

1: 50–100%, 2: 10–<50%, 3: 0–<10%[1]

Percentage of Segments That Are Hispanic or Latino

1: 50–100%, 2: 10–<50%, 3: 0–<10%[1]

Segment-Combined Median Rent and Housing Value (Rent/Housing)[2]

1: First Quintile, 2: Second Quintile, 3: Third Quintile, 4: Fourth Quintile, 5: Fifth Quintile[1]

Population Density

1: MSA 1,000,000 or More, 2: MSA Less than 1,000,000, 3: Non-MSA Urban, 4: Non-MSA Rural[1]

Quarter

1: Quarter 1, 2: Quarter 2, 3: Quarter 3, 4: Quarter 4[1]

Race/Ethnicity of Householder

1: Hispanic or Latino White,[1] 2: Hispanic or Latino Black or African American, 3: Hispanic or Latino Other, 4: Non-Hispanic or Latino White, 5: Non-Hispanic or Latino Black or African American, 6: Non-Hispanic or Latino Other

State/Region

Model Group 1: 1: Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, Rhode Island, Vermont; 2: Alabama, Arkansas, Delaware, District of Columbia, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Virginia, West Virginia;[1] 3: New York; 4: Pennsylvania; 5: Florida; 6: Texas

Model Group 2: 1: Indiana, Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota, Wisconsin;[1] 2: Alaska, Arizona, Colorado, Idaho, Hawaii, Montana, Nevada, New Mexico, Oregon, Utah, Washington, Wyoming; 3: Michigan; 4: Illinois; 5: Ohio; 6: California

Exhibit H.1 Definitions of Levels for Pair-Level Calibration Modeling Variables (continued)

States[3]

Model Group 1: 1: Alabama, 2: Arkansas, 3: Connecticut, 4: Delaware, 5: District of Columbia, 6: Florida, 7: Georgia, 8: Kentucky, 9: Louisiana, 10: Maine, 11: Maryland,[1] 12: Massachusetts, 13: Mississippi, 14: New Hampshire, 15: New Jersey, 16: New York, 17: North Carolina, 18: Oklahoma, 19: Pennsylvania, 20: Rhode Island, 21: South Carolina, 22: Tennessee, 23: Texas, 24: Vermont, 25: Virginia, 26: West Virginia

Model Group 2: 1: Alaska, 2: Arizona,[1] 3: California, 4: Colorado, 5: Idaho, 6: Illinois, 7: Indiana, 8: Iowa, 9: Hawaii, 10: Kansas, 11: Michigan, 12: Minnesota, 13: Missouri, 14: Montana, 15: Nebraska, 16: Nevada, 17: New Mexico, 18: North Dakota, 19: Ohio, 20: Oregon, 21: South Dakota, 22: Utah, 23: Washington, 24: Wisconsin, 25: Wyoming

Pair Relationship Associated with Multiplicity

1: Parent-Child (12–14)*
2: Parent-Child (12–17)*
3: Parent-Child (12–20)*
4: Parent*-Child (12–14)
5: Parent*-Child (12–17)
6: Parent*-Child (12–20)
7: Sibling (12–14)-Sibling (15–17)*
8: Sibling (12–17)-Sibling (18–25)*
9: Spouse-Spouse/Partner-Partner
10: Spouse-Spouse/Partner-Partner with Children (Younger than 18)

DU = dwelling unit, MSA = metropolitan statistical area.

[1] The reference level for this variable. This is the level against which effects of other factor levels are measured.

[2] Segment-Combined Median Rent and Housing Value is a composite measure based on rent, housing value, and percentage owner-occupied.

[3] The states or district assigned to a particular model is based on combined census regions.

* The pair member focused on.

Exhibit H.2 Covariates for 2017 NSDUH Pair Weights

Variables	Level	Proposed
One-Factor Effects		
Intercept	1	1
State	Model-specific	
Quarter	4	3
Population Density	3	2
Group Quarter	3	2
Household Size	3	2
Pair Age	15	14
Pair Gender	4	2
Pair Race/Ethnicity	10	9
Race/Ethnicity of Householder	6	5
Rent/Housing	5	4
Segment % Black or African American	3	2
Segment % Hispanic or Latino	3	2
% Owner-Occupied	3	2
Pair Relationship[1,2]	10	10
Two-Factor Effects		
Pair Race/Ethnicity (5 Levels) × Pair Age (6 Levels)	5 × 6	20
Pair Race/Ethnicity (5 Levels) × Pair Gender	5 × 3	8
Pair Gender × Pair Age (6 Levels)	3 × 6	10
State/Region × Pair Race/Ethnicity (5 Levels)	Model-specific	
State/Region × Pair Age (6 Levels)	Model-specific	
State/Region × Pair Gender	Model-specific	
Rent/Housing × % Black or African American	5 × 3	8
Rent/Housing × % Hispanic or Latino	5 × 3	8
Rent/Housing × % Owner-Occupied	5 × 3	8
% Owner-Occupied × % Black or African American	3 × 3	4
% Owner-Occupied × % Hispanic or Latino	3 × 3	4
Three-Factor Effects		
Pair Race/Ethnicity (4 Levels) × Pair Gender × Pair Age (3 Levels)	4 × 3 × 3	12

[1] Pair Relationship variables are included in only the respondent pair poststratification and respondent pair extreme value adjustment steps.

[2] Note that Pair Relationship variables are single category indicators; as such, they do not require a reference level.

Appendix H.1: Model Group 1: Northeast and South

(Alabama, Arkansas, Connecticut, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maine, Massachusetts, Maryland, Mississippi, New Hampshire, New Jersey, New York, North Carolina, Oklahoma, Pennsylvania, Rhode Island, South Carolina, Tennessee, Texas, Vermont, Virginia, West Virginia)

This page intentionally left blank

Table H.1a 2017 Pair Weight GEM Modeling Summary (Model Group 1: Northeast and South)

Modeling Step[1]	Extreme Weight Proportions			UWE[2]	# Covariates[3]	Bounds[4]	
	% Unweighted	% Weighted	% Winsorized			Nominal	Realized
sel.pr.ps	3.92	21.33	9.70	10.7984	213	(0.39, 1.30)	(0.39, 1.30)
	1.31	3.07	0.47	4.8448	204	(0.20, 3.31)	(0.20, 3.31)
						N/A	N/A
res.pr.nr	1.62	4.96	0.66	4.6874	213	(1.00, 2.00)	(1.00, 2.00
	2.39	8.66	1.79	6.4694	213	(1.00, 5.00)	(1.00, 5.00)
						N/A	N/A
res.pr.ps	2.53	10.06	2.20	6.4694	223	(0.31, 1.10)	(031, 1.10)
	1.27	4.03	0.62	6.2385	214	(0.22, 2.32)	(0.22, 2.32)
						N/A	N/A
res.pr.ev	1.27	4.03	0.62	6.2385	223	(0.95, 1.33)	(0.95, 1.33)
	0.71	2.49	0.18	6.2116	214	(0.92, 1.36)	(0.92, 1.36)
						N/A	N/A

GEM = generalized exponential model; N/A = not applicable.

[1] For a key to modeling abbreviations, see Chapter 6, Exhibit 6.1.

[2] Unequal weighting effect (UWE) defined as $1 + \left[(n-1)/n\right]*CV^2$, where CV = coefficient of variation of weights.

[3] Number of proposed covariates on top line and number finalized after modeling.

[4] Nominal bounds are used in defining maximum/minimum values for the GEM adjustment factors. The realized bound is the actual adjustment produced by the modeling. The first set of bounds listed is for high extreme values, the second is for nonextreme values, and the third is for low extreme values.

Table H.1b 2017 Distribution of Weight Adjustment Factors and Weight Products (Model Group 1: Northeast and South)

	SDU Weight	Pair Selection Probability		sel.pr.ps[1]		res.pr.nr[1]		res.pr.ps[1]		res.pr.ev[1]	
Statistics	**1–11**	**pairwt12**	**1–12**	**pairwt13**	**1–13**	**pairwt14**	**1–14**	**pairwt15**	**1–15**	**pairwt16**	**1–16**
Minimum	7	1.02	11	0.07	10	0.55	10	0.09	17	0.58	17
1%	62	1.16	179	0.25	136	1.00	160	0.28	131	0.91	129
5%	139	1.38	457	0.36	380	1.03	477	0.47	419	0.95	413
10%	210	1.54	829	0.58	677	1.06	857	0.62	748	0.97	742
25%	610	2.47	1,753	0.82	1,629	1.16	2,117	0.83	1,912	0.99	1,903
Median	953	3.88	3,615	1.02	3,692	1.37	5,115	1.02	4,911	1.00	4,929
75%	1,322	8.25	8,104	1.27	8,303	1.74	12,053	1.18	11,981	1.01	12,005
90%	1,801	16.87	16,923	1.57	18,623	2.40	28,702	1.34	29,002	1.03	28,851
95%	2,125	29.28	29,504	1.80	31,396	2.97	5,844	1.46	53,647	1.04	53,635
99%	3,001	60.17	65,943	2.42	72,639	4.21	141,566	1.73	147,202	1.11	147,793
Maximum	8,679	989.66	1,415,464	3.31	290,613	5.00	763,301	2.32	511,846	1.36	590,895
n	15,072	-	15,072	-	15,072	-	9,284	-	9,284	-	9,284
Mean	1,013	8.48	8,447	1.06	8,342	1.58	13,543	1.00	13,543	1.00	13,543
Max/Mean	9	-	168	-	35	-	56	-	38	-	38

SDU = screener dwelling unit.

[1] For a key to modeling abbreviations, see Chapter 6, Exhibit 6.1.

Model Group 1 Overview

Selected Pair-Level Poststratification

In the selected pair-level poststratification step, 204 of 213 proposed factors were retained in the final model. All main and two-factor effects were retained at proposed levels. Of the 12 three-factor effects, 3 collapsed variables were kept in the model, and the rest were dropped because of convergence problems.

Respondent Pair-Level Nonresponse

In the respondent pair-level nonresponse step, all 213 proposed factors were retained in the final model.

Respondent Pair-Level Poststratification

In the respondent pair-level poststratification step, 214 of 223 proposed factors were retained in the final model. All main and two-factor effects were retained at proposed levels. Of the 12 three-factor effects, 3 collapsed variables were kept in the model, and the rest were dropped because of convergence problems.

Respondent Pair-Level Extreme Value Adjustment

This step used exactly the same variables as in the respondent pair-level poststratification step.

Exhibit H.1.1 Covariates for 2017 NSDUH Pair Weights (sel.pr.ps) Model Group 1: Northeast and South

Variables	Level	Proposed	Final	Comments
One-Factor Effects		**76**	**76**	
Intercept	1	1	1	All levels present.
State	26	25	25	All levels present.
Quarter	4	3	3	All levels present.
Population Density	4	3	3	All levels present.
Group Quarter	3	2	2	All levels present.
Household Size	3	2	2	All levels present.
Pair Age	15	14	14	All levels present.
Pair Gender	3	2	2	All levels present.
Pair Race/Ethnicity	10	9	9	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Rent/Housing	5	4	4	All levels present.
Segment % Black or African American	3	2	2	All levels present.
Segment % Hispanic or Latino	3	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
Two-Factor Effects		**125**	**125**	
Pair Race/Ethnicity (5 Levels) × Pair Age (6 Levels)	5 × 6	20	20	All levels present.
Pair Race/Ethnicity (5 Levels) × Pair Gender	5 × 3	8	8	All levels present.
Pair Gender × Pair Age (6 Levels)	3 × 6	10	10	All levels present.
State/Region × Pair Race/Ethnicity (5 Levels)	6 × 5	20	20	All levels present.
State/Region × Pair Age (6 Levels)	6 × 6	25	25	All levels present.
State/Region × Pair Gender	6 × 3	10	10	All levels present.
Rent/Housing × % Black or African American	5 × 3	8	8	All levels present.
Rent/Housing × % Hispanic or Latino	5 × 3	8	8	All levels present.
Rent/Housing × % Owner-Occupied	5 × 3	8	8	All levels present.
% Owner-Occupied × % Black or African American	3 × 3	4	4	All levels present.
% Owner-Occupied × % Hispanic or Latino	3 × 3	4	4	All levels present.
Three-Factor Effects		**12**	**3**	
Pair Race/Ethnicity (4 Levels) × Pair Gender × Pair Age (3 Levels)	4 × 3 × 3	12	3	Coll. (1,1,2), (2,1,2) & (3,1,2); (1,1,3), (2,1,3) & (3,1,3); (1,2,2), (2,2,2) & (3,2,2); conv.
Total		**213**	**204**	

Exhibit H.1.2 Covariates for 2017 NSDUH Pair Weights (res.pr.nr) Model Group 1: Northeast and South

Variables	Level	Proposed	Final	Comments
One-Factor Effects		**76**	**76**	
Intercept	1	1	1	All levels present.
State	26	25	25	All levels present.
Quarter	4	3	3	All levels present.
Population Density	4	3	3	All levels present.
Group Quarter	3	2	2	All levels present.
Household Size	3	2	2	All levels present.
Pair Age	15	14	14	All levels present.
Pair Gender	3	2	2	All levels present.
Pair Race/Ethnicity	10	9	9	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Rent/Housing	5	4	4	All levels present.
Segment % Black or African American	3	2	2	All levels present.
Segment % Hispanic or Latino	3	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
Two-Factor Effects		**125**	**125**	
Pair Race/Ethnicity (5 Levels) × Pair Age (6 Levels)	5 × 6	20	20	All levels present.
Pair Race/Ethnicity (5 Levels) × Pair Gender	5 × 3	8	8	All levels present.
Pair Gender × Pair Age (6 Levels)	3 × 6	10	10	All levels present.
State/Region × Pair Race/Ethnicity (5 Levels)	6 × 5	20	20	All levels present.
State/Region × Pair Age (6 Levels)	6 × 6	25	25	All levels present.
State/Region × Pair Gender	6 × 3	10	10	All levels present.
Rent/Housing × % Black or African American	5 × 3	8	8	All levels present.
Rent/Housing × % Hispanic or Latino	5 × 3	8	8	All levels present.
Rent/Housing × % Owner-Occupied	5 × 3	8	8	All levels present.
% Owner-Occupied × % Black or African American	3 × 3	4	4	All levels present.
% Owner-Occupied × % Hispanic or Latino	3 × 3	4	4	All levels present.
Three-Factor Effects		**12**	**12**	
Pair Race/Ethnicity (4 Levels) × Pair Gender × Pair Age (3 Levels)	4 × 3 × 3	12	12	All levels present.
Total		**213**	**213**	

Exhibit H.1.3 Covariates for 2017 NSDUH Pair Weights (res.pr.ps) Model Group 1: Northeast and South

Variables	Level	Proposed	Final	Comments
One-Factor Effects		**86**	**86**	
Intercept	1	1	1	All levels present.
State	26	25	25	All levels present.
Quarter	4	3	3	All levels present.
Population Density	4	3	3	All levels present.
Group Quarter	3	2	2	All levels present.
Household Size	3	2	2	All levels present.
Pair Age	15	14	14	All levels present.
Pair Gender	3	2	2	All levels present.
Pair Race/Ethnicity	10	9	9	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Rent/Housing	5	4	4	All levels present.
Segment % Black or African American	3	2	2	All levels present.
Segment % Hispanic or Latino	3	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
Pair Relationship	10	10	10	All levels present.
Two-Factor Effects		**125**	**125**	
Pair Race/Ethnicity (5 Levels) × Pair Age (6 Levels)	5 × 6	20	20	All levels present.
Pair Race/Ethnicity (5 Levels) × Pair Gender	5 × 3	8	8	All levels present.
Pair Gender × Pair Age (6 Levels)	3 × 6	10	10	All levels present.
State/Region × Pair Race/Ethnicity (5 Levels)	6 × 5	20	20	All levels present.
State/Region × Pair Age (6 Levels)	6 × 6	25	25	All levels present.
State/Region × Pair Gender	6 × 3	10	10	All levels present.
Rent/Housing × % Black or African American	5 × 3	8	8	All levels present.
Rent/Housing × % Hispanic or Latino	5 × 3	8	8	All levels present.
Rent/Housing × % Owner-Occupied	5 × 3	8	8	All levels present.
% Owner-Occupied × % Black or African American	3 × 3	4	4	All levels present.
% Owner-Occupied × % Hispanic or Latino	3 × 3	4	4	All levels present.
Three-Factor Effects		**12**	**3**	
Pair Race/Ethnicity (4 Levels) × Pair Gender × Pair Age (3 Levels)	4 × 3 × 3	12	3	Coll. (1,1,2), (2,1,2) & (3,1,2); (1,1,3), (2,1,3) & (3,1,3); (1,2,2), (2,2,2) & (3,2,2); conv.
Total		**223**	**214**	

Exhibit H.1.4 Covariates for 2017 NSDUH Pair Weights (res.pr.ev) Model Group 1: Northeast and South

This step used the same variables as the respondent pair-level poststratification step in Exhibit H.1.3.

This page intentionally left blank

Appendix H.2: Model Group 2: Midwest and West

(Alaska, Arizona, California, Colorado, Hawaii, Idaho, Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Montana, Nebraska, Nevada, New Mexico, North Dakota, Ohio, Oregon, South Dakota, Utah, Washington, Wisconsin, Wyoming)

This page intentionally left blank

Table H.2a 2017 Pair Weight GEM Modeling Summary (Model Group 2: Midwest and West)

Modeling Step[1]	Extreme Weight Proportions			UWE[2]	# Covariates[3]	Bounds[4]	
	% Unweighted	% Weighted	% Winsorized			Nominal	Realized
sel.pr.ps	4.28	19.55	9.92	17.8805	212	(0.29, 2.00)	(0.30, 2.00)
	2.06	5.27	0.60	4.5967	201	(0.30, 2.23)	(0.31, 2.23)
						(0.90, 1.16)	(0.90, 1.16)
res.pr.nr	2.20	7.39	0.92	4.7489	212	(1.03, 2.10)	(1.03, 2.10)
	2.07	5.36	0.87	6.1677	212	(1.00, 4.84)	(1.00, 4.82)
						N/A	N/A
res.pr.ps	2.14	6.21	1.01	6.1677	222	(0.51, 1.10)	(0.51, 1.10)
	0.57	1.38	0.09	6.2417	211	(0.33, 1.54)	(0.34, 1.53)
						N/A	N/A
res.pr.ev	0.57	1.38	0.09	6.2417	222	(0.97, 1.05)	(0.97, 1.05)
	0.00	0.03	0.00	6.2088	210	(0.95, 1.09)	(0.95, 1.09)
						N/A	N/A

GEM = generalized exponential model; N/A = not applicable.

[1] For a key to modeling abbreviations, see Chapter 6, Exhibit 6.1.

[2] Unequal weighting effect (UWE) defined as $1 + \left[(n-1)/n\right]*CV^2$, where CV = coefficient of variation of weights.

[3] Number of proposed covariates on top line and number finalized after modeling.

[4] Nominal bounds are used in defining maximum/minimum values for the GEM adjustment factors. The realized bound is the actual adjustment produced by the modeling. The first set of bounds listed is for high extreme values, the second is for nonextreme values, and the third is for low extreme values.

Table H.2b 2017 Distribution of Weight Adjustment Factors and Weight Products (Model Group 2: Midwest and West)

	SDU Weight	Pair Selection		sel.pr.ps[1]		res.pr.nr[1]		res.pr.ps[1]		res.pr.ev[1]	
Statistics	**1–11**	**pairwt12**	**1–12**	**pairwt13**	**1–13**	**pairwt14**	**1–14**	**pairwt15**	**1–15**	**pairwt16**	**1–16**
Minimum	21	1.02	34	0.04	56	0.52	60	0.15	57	0.67	56
1%	88	1.14	201	0.35	193	0.99	247	0.47	237	0.93	235
5%	129	1.40	399	0.55	375	1.05	489	0.62	479	0.97	480
10%	175	1.57	674	0.68	608	1.09	803	0.71	762	0.98	758
25%	470	2.52	1,480	0.85	1,02	1.19	1,785	0.90	1,761	0.99	1,755
Median	864	3.91	3,274	1.04	3,321	1.42	4,344	1.04	4,376	1.00	4,375
75%	1,284	8.04	7,559	1.24	7,878	1.82	11,089	1.17	11,021	1.01	10,967
90%	1,697	16.02	16,087	1.45	17,298	2.42	27,525	1.27	26,918	1.02	26,646
95%	1,965	26.37	27,342	1.57	30,111	2.89	51,616	1.33	52,109	1.03	52,161
99%	2,826	50.32	59,098	1.84	68,766	3.84	131,944	1.44	140,338	1.06	141,119
Maximum	8,022	2,463.78	2,350,151	2.23	248,443	4.82	627,261	1.53	502,020	1.09	475,394
n	13,706	-	13,706	-	13,706	-	8,420	-	8,420	-	8,420
Mean	935	8.01	7,701	1.05	7,825	1.61	12,575	1.02	12,575	1.00	12,575
Max/Mean	9	-	271	-	38	-	50	-	40	-	38

SDU = screener dwelling unit.

[1] For a key to modeling abbreviations, see Chapter 6, Exhibit 6.1.

Model Group 2 Overview

Selected Pair-Level Poststratification

In the selected pair-level poststratification step, 201 of 212 proposed factors were retained in the final model. All main and two-factor effects were retained at proposed levels. Of the 12 three-factor effects, 1 collapsed variable was kept in the model, and the rest were dropped because of convergence problems.

Respondent Pair-Level Nonresponse

In the respondent pair-level nonresponse step, all 212 proposed factors were retained in the final model.

Respondent Pair-Level Poststratification

In the respondent pair-level poststratification step, 211 of 222 proposed factors were retained in the final model. All main and two-factor effects were retained at proposed levels. Of the 12 three-factor effects, 1 collapsed variable was kept in the model, and the rest were dropped because of convergence problems.

Respondent Pair-Level Extreme Value Adjustment

The respondent pair-level extreme value adjustment step used 210 of 222 proposed factors in the final model. The main effect Race/Ethnicity of Householder categories Hispanic or Latino Black or African American and Hispanic or Latino Other were combined, but all other main and two-factor effects were retained at proposed levels. Of the 12 three-factor effects, 1 collapsed variable was kept in the model, and the rest were dropped because of convergence problems.

Exhibit H.2.1 Covariates for 2017 NSDUH Pair Weights (sel.pr.ps) Model Group 2: Midwest and West

Variables	Level	Proposed	Final	Comments
One-Factor Effects		**75**	**75**	
Intercept	1	1	1	All levels present.
State	25	24	24	All levels present.
Quarter	4	3	3	All levels present.
Population Density	4	3	3	All levels present.
Group Quarter	3	2	2	All levels present.
Household Size	3	2	2	All levels present.
Pair Age	15	14	14	All levels present.
Pair Gender	3	2	2	All levels present.
Pair Race/Ethnicity	10	9	9	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Rent/Housing	5	4	4	All levels present.
Segment % Black or African American	3	2	2	All levels present.
Segment % Hispanic or Latino	3	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
Two-Factor Effects		**125**	**125**	
Pair Race/Ethnicity (5 Levels) × Pair Age (6 Levels)	5 × 6	20	20	All levels present.
Pair Race/Ethnicity (5 Levels) × Pair Gender	5 × 3	8	8	All levels present.
Pair Gender × Pair Age (6 Levels)	3 × 6	10	10	All levels present.
State/Region × Pair Race/Ethnicity (5 Levels)	6 × 5	20	20	All levels present.
State/Region × Pair Age (6 Levels)	6 × 6	25	25	All levels present.
State/Region × Pair Gender	6 × 3	10	10	All levels present.
Rent/Housing × % Black or African American	5 × 3	8	8	All levels present.
Rent/Housing × % Hispanic or Latino	5 × 3	8	8	All levels present.
Rent/Housing × % Owner-Occupied	5 × 3	8	8	All levels present.
% Owner-Occupied × % Black or African American	3 × 3	4	4	All levels present.
% Owner-Occupied × % Hispanic or Latino	3 × 3	4	4	All levels present.
Three-Factor Effects		**12**	**1**	
Pair Race/Ethnicity (4 Levels) × Pair Gender × Pair Age (3 Levels)	4 × 3 × 3	12	1	Coll. (1,1,2), (2,1,2) & (3,1,2); conv.
Total		**212**	**201**	

Exhibit H.2.2 Covariates for 2017 NSDUH Pair Weights (res.pr.nr) Model Group 2: Midwest and West

Variables	Level	Proposed	Final	Comments
One-Factor Effects		**75**	**75**	
Intercept	1	1	1	All levels present.
State	25	24	24	All levels present.
Quarter	4	3	3	All levels present.
Population Density	4	3	3	All levels present.
Group Quarter	3	2	2	All levels present.
Household Size	3	2	2	All levels present.
Pair Age	15	14	14	All levels present.
Pair Gender	3	2	2	All levels present.
Pair Race/Ethnicity	10	9	9	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Rent/Housing	5	4	4	All levels present.
Segment % Black or African American	3	2	2	All levels present.
Segment % Hispanic or Latino	3	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
Two-Factor Effects		**125**	**125**	
Pair Race/Ethnicity (5 Levels) × Pair Age (6 Levels)	5 × 6	20	20	All levels present.
Pair Race/Ethnicity (5 Levels) × Pair Gender	5 × 3	8	8	All levels present.
Pair Gender × Pair Age (6 Levels)	3 × 6	10	10	All levels present.
State/Region × Pair Race/Ethnicity (5 Levels)	6 × 5	20	20	All levels present.
State/Region × Pair Age (6 Levels)	6 × 6	25	25	All levels present.
State/Region × Pair Gender	6 × 3	10	10	All levels present.
Rent/Housing × % Black or African American	5 × 3	8	8	All levels present.
Rent/Housing × % Hispanic or Latino	5 × 3	8	8	All levels present.
Rent/Housing × % Owner-Occupied	5 × 3	8	8	All levels present.
% Owner-Occupied × % Black or African American	3 × 3	4	4	All levels present.
% Owner-Occupied × % Hispanic or Latino	3 × 3	4	4	All levels present.
Three-Factor Effects		**12**	**12**	
Pair Race/Ethnicity (4 Levels) × Pair Gender × Pair Age (3 Levels)	4 × 3 × 3	12	12	All levels present.
Total		**212**	**212**	

Exhibit H.2.3 Covariates for 2017 NSDUH Pair Weights (res.pr.ps) Model Group 2: Midwest and West

Variables	Level	Proposed	Final	Comments
One-Factor Effects		**85**	**85**	
Intercept	1	1	1	All levels present.
State	26	24	24	All levels present.
Quarter	4	3	3	All levels present.
Population Density	4	3	3	All levels present.
Group Quarter	3	2	2	All levels present.
Household Size	3	2	2	All levels present.
Pair Age	15	14	14	All levels present.
Pair Gender	3	2	2	All levels present.
Pair Race/Ethnicity	10	9	9	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Rent/Housing	5	4	4	All levels present.
Segment % Black or African American	3	2	2	All levels present.
Segment % Hispanic or Latino	3	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
Pair Relationship	10	10	10	All levels present.
Two-Factor Effects		**125**	**125**	
Pair Race/Ethnicity (5 Levels) × Pair Age (6 Levels)	5 × 6	20	20	All levels present.
Pair Race/Ethnicity (5 Levels) × Pair Gender	5 × 3	8	8	All levels present.
Pair Gender × Pair Age (6 Levels)	3 × 6	10	10	All levels present.
State/Region × Pair Race/Ethnicity (5 Levels)	6 × 5	20	20	All levels present.
State/Region × Pair Age (6 Levels)	6 × 6	25	25	All levels present.
State/Region × Pair Gender	6 × 3	10	10	All levels present.
Rent/Housing × % Black or African American	5 × 3	8	8	All levels present.
Rent/Housing × % Hispanic or Latino	5 × 3	8	8	All levels present.
Rent/Housing × % Owner-Occupied	5 × 3	8	8	All levels present.
% Owner-Occupied × % Black or African American	3 × 3	4	4	All levels present.
% Owner-Occupied × % Hispanic or Latino	3 × 3	4	4	All levels present.
Three-Factor Effects		**12**	**1**	
Pair Race/Ethnicity (4 Levels) × Pair Gender × Pair Age (3 Levels)	4 × 3 × 3	12	1	Coll. (1,1,2), (2,1,2) & (3,1,2); conv.
Total		**222**	**211**	

Exhibit H.2.4 Covariates for 2017 NSDUH Pair Weights (res.pr.ev) Model Group 2: Midwest and West

Variables	Level	Proposed	Final	Comments
One-Factor Effects		**85**	**84**	
Intercept	1	1	1	All levels present.
State	26	24	24	All levels present.
Quarter	4	3	3	All levels present.
Population Density	4	3	3	All levels present.
Group Quarter	3	2	2	All levels present.
Household Size	3	2	2	All levels present.
Pair Age	15	14	14	All levels present.
Pair Gender	3	2	2	All levels present.
Pair Race/Ethnicity	10	9	9	All levels present.
Race/Ethnicity of Householder	6	5	4	Coll (2) &(3); conv.
Rent/Housing	5	4	4	All levels present.
Segment % Black or African American	3	2	2	All levels present.
Segment % Hispanic or Latino	3	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
Pair Relationship	10	10	10	All levels present.
Two-Factor Effects		**125**	**125**	
Pair Race/Ethnicity (5 Levels) × Pair Age (6 Levels)	5 × 6	20	20	All levels present.
Pair Race/Ethnicity (5 Levels) × Pair Gender	5 × 3	8	8	All levels present.
Pair Gender × Pair Age (6 Levels)	3 × 6	10	10	All levels present.
State/Region × Pair Race/Ethnicity (5 Levels)	6 × 5	20	20	All levels present.
State/Region × Pair Age (6 Levels)	6 × 6	25	25	All levels present.
State/Region × Pair Gender	6 × 3	10	10	All levels present.
Rent/Housing × % Black or African American	5 × 3	8	8	All levels present.
Rent/Housing × % Hispanic or Latino	5 × 3	8	8	All levels present.
Rent/Housing × % Owner-Occupied	5 × 3	8	8	All levels present.
% Owner-Occupied × % Black or African American	3 × 3	4	4	All levels present.
% Owner-Occupied × % Hispanic or Latino	3 × 3	4	4	All levels present.
Three-Factor Effects		**12**	**1**	
Pair Race/Ethnicity (4 Levels) × Pair Gender × Pair Age (3 Levels)	4 × 3 × 3	12	1	Coll. (1,1,2), (2,1,2) & (3,1,2); conv.
Total		**222**	**210**	

This page intentionally left blank

Appendix I: Evaluation of Calibration Weights: Pair-Level Response Rates

This page intentionally left blank

Table I.1 2017 NSDUH Person Pair-Level Response Rates

Domain	Selected Pairs	Respondent Pairs	% Interview Response Rate[1]
Total	28,778	17,704	54.90
Pair Age Group			
12–17, 12–17	3,261	2,371	74.61
12–17, 18–25	2,679	1,755	65.32
12–17, 26–34	1,338	910	66.74
12–17, 35–49	5,845	3,862	66.26
12–17, 50+	1,211	763	64.91
18–25, 18–25	4,167	2,524	59.67
18–25, 26–34	1,443	812	55.34
18–25, 35–49	2,084	1,206	60.29
18–25, 50+	1,309	670	51.32
26–34, 26–34	1,551	841	51.51
26–34, 35–49	810	447	56.22
26–34, 50+	460	235	48.38
35–49, 35–49	1,233	627	47.64
35–49, 50+	474	211	38.97
50+, 50+	913	470	48.01
Pair Race/Ethnicity			
Hispanic or Latino	4,991	3,083	54.68
Black or African American	2,771	1,918	62.26
White	15,850	9,585	54.71
Other	2,162	1,212	43.49
White & Black or African American	316	205	64.49
White & Hispanic or Latino	1,192	729	54.46
White & Other	965	630	62.23
Black or African American & Hispanic or Latino	146	88	53.32
Black or African American & Other	157	103	48.20
Hispanic or Latino & Other	228	151	58.34
Pair Gender			
Male, Male	6,140	3,609	54.48
Female, Female	6,164	4,056	59.74
Male, Female	16,474	10,039	53.64
Household Size			
Two	7,318	4,325	53.57
Three	9,133	5,674	54.59
Four or More	12,327	7,705	55.73

(continued)

Table I.1 2017 NSDUH Person Pair-Level Response Rates (continued)

Domain	Selected Pairs	Respondent Pairs	% Interview Response Rate[1]
Census Region			
Northeast	5,746	3,346	50.09
South	9,326	5,938	58.34
Midwest	6,765	4,162	56.52
West	6,941	4,258	52.16
Quarter			
Quarter 1	6,630	4,059	54.99
Quarter 2	7,446	4,566	56.07
Quarter 3	7,307	4,480	52.00
Quarter 4	7,395	4,599	56.55
% Hispanic or Latino in Segment			
50–100%	2,731	1,664	54.56
10–<50%	7,204	4,435	54.02
<10%	18,843	11,605	55.42
% Black or African American in Segment			
50–100%	1,840	1,255	61.86
10–<50%	5,535	3,482	56.90
<10%	21,403	12,967	53.67
% Owner-Occupied DUs in Segment			
50–100%	21,550	13,232	55.03
10–<50%	5,763	3,584	54.71
<10%	1,465	888	51.36
Combined Median Rent/Housing Value			
1st Quintile	4,645	3,052	59.34
2nd Quintile	6,462	4,204	58.97
3rd Quintile	6,791	4,156	56.40
4th Quintile	6,191	3,657	51.08
5th Quintile	4,689	2,635	50.13
Population Density			
Large MSA	12,726	7,503	52.62
Medium to Small MSA	13,766	8,751	57.61
Non-MSA, Urban	792	539	58.69
Non-MSA, Rural	1,494	911	61.20
Group Quarters			
Group	201	150	61.09
Non-Group	28,577	17,554	54.89

DU = dwelling unit, MSA = metropolitan statistical area.

[1] The weight used for calculating the response rate includes screener dwelling unit (SDU)- and pair-level design weights, SDU nonresponse and poststratification adjustments, and selected pair poststratification adjustment. This weight is the product of WT1*…*WT11*PRWT12*PRWT13.

Appendix J: Evaluation of Calibration Weights: Pair-Level Proportions of Extreme Values and Outwinsors

This page intentionally left blank

Table J.1 2017 NSDUH Selected Pair-Level Proportions of Extreme Values and Outwinsors

Domain	*n*	SDU-Level Weights[1] (SDUWT: WT1*...*WT11)			Before sel.pr.ps[1] (SDUWT*PRWT12)			After sel.pr.ps[1] (SDUWT*PRWT12*PRWT13)		
		% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]
Total	28,778	1.96	5.07	1.29	4.16	21.74	11.25	1.79	7.23	1.81
Pair Age Group										
12–17, 12–17	3,261	1.20	3.35	0.72	2.73	11.01	3.54	0.74	4.59	0.81
12–17, 18–25	2,679	1.57	4.11	1.06	6.38	22.84	8.73	1.23	4.43	0.55
12–17, 26–34	1,338	2.77	7.46	2.21	2.24	9.73	2.50	0.90	7.66	2.80
12–17, 35–49	5,845	1.57	4.52	1.29	1.93	8.90	2.67	0.51	1.68	0.30
12–17, 50+	1,211	1.16	3.05	0.74	1.32	5.76	1.58	0.08	0.57	0.14
18–25, 18–25	4,167	2.52	6.15	1.25	6.91	26.23	10.26	3.77	11.35	1.40
18–25, 26–34	1,443	4.44	9.64	2.49	5.61	16.93	5.42	3.12	7.76	1.11
18–25, 35–49	2,084	2.50	6.55	2.18	7.39	25.47	9.58	3.50	7.37	0.80
18–25, 50+	1,309	2.14	5.32	1.14	4.66	19.85	8.84	0.38	1.10	0.11
26–34, 26–34	1,551	1.93	4.90	1.48	1.68	6.31	1.95	1.55	3.76	0.50
26–34, 35–49	810	1.23	2.71	0.78	3.09	16.51	8.12	1.73	3.31	0.50
26–34, 50+	460	1.30	3.77	0.44	1.96	18.66	11.50	0.22	1.91	0.89
35–49, 35–49	1,233	1.46	3.28	0.53	3.57	27.49	16.42	2.84	3.69	0.55
35–49, 50+	474	1.48	4.00	1.41	5.91	34.08	16.73	4.85	21.37	5.41
50+, 50+	913	2.08	5.21	1.10	6.68	35.69	23.94	4.16	14.72	5.09
Pair Race/Ethnicity										
Hispanic or Latino	4,991	2.81	7.63	2.66	3.95	26.00	14.71	1.74	8.12	2.19
Black or African American	2,771	3.86	7.89	1.67	6.93	28.55	12.97	2.38	10.31	3.04
White	15,850	0.64	1.53	0.24	3.14	16.89	8.60	1.27	5.91	1.47
Other	2,162	4.53	10.50	2.07	6.29	32.29	20.11	2.82	8.72	1.67
White & Black or African American	316	5.38	11.56	2.62	10.44	42.10	24.58	1.90	17.73	6.18
White & Hispanic or Latino	1,192	2.10	5.27	1.21	5.03	14.57	4.27	2.94	5.51	0.68
White & Other	965	3.42	8.72	2.18	4.15	19.55	6.89	3.01	6.53	0.53
Black or African American & Hispanic or Latino	146	17.81	37.90	11.46	14.38	36.13	11.96	13.70	19.60	6.35
Black or African American & Other	157	1.91	4.12	1.05	2.55	5.43	2.36	2.55	3.17	0.49
Hispanic or Latino & Other	228	5.70	18.58	6.23	7.02	32.25	12.12	2.19	3.68	1.02
Pair Gender										
Male, Male	6,140	2.26	5.20	1.22	5.57	19.17	7.38	2.18	4.43	0.64
Female, Female	6,164	2.11	6.04	1.69	4.87	24.60	12.52	1.98	7.61	1.87
Male, Female	16,474	1.78	4.65	1.16	3.36	21.61	11.97	1.57	7.90	2.12
Household Size										
Two	7,318	1.57	3.74	0.94	0.93	2.59	0.71	0.45	1.25	0.23
Three	9,133	1.74	4.64	1.21	1.85	22.25	14.09	1.40	4.82	1.04
Four or More	12,327	2.34	6.11	1.54	7.78	30.68	14.93	2.87	11.47	3.01

Table J.1 2017 NSDUH Selected Pair-Level Proportions of Extreme Values and Outwinsors (continued)

Domain	*n*	SDU-Level Weights[1] (SDUWT: WT1*...*WT11)			Before sel.pr.ps[1] (SDUWT*PRWT12)			After sel.pr.ps[1] (SDUWT*PRWT12*PRWT13)		
		% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]
Census Region										
Northeast	5,746	2.09	6.73	1.86	4.26	20.10	8.80	1.41	6.16	1.59
South	9,326	1.43	3.59	0.84	3.80	23.30	12.81	1.50	7.77	2.32
Midwest	6,765	2.25	5.24	1.28	4.39	16.98	7.50	2.57	7.36	1.35
West	6,941	2.28	6.10	1.61	4.32	23.86	13.17	1.73	7.10	1.60
Quarter										
Quarter 1	6,630	2.61	6.60	1.65	5.25	24.51	12.41	2.52	7.12	1.36
Quarter 2	7,446	1.65	4.30	1.02	3.76	20.52	11.15	1.61	6.04	1.44
Quarter 3	7,307	2.11	5.25	1.38	4.01	25.89	14.37	1.68	10.18	2.92
Quarter 4	7,395	1.53	4.12	1.10	3.72	15.49	6.67	1.42	5.56	1.52
% Hispanic or Latino in Segment										
50–100%	2,731	1.61	5.08	1.84	3.73	31.55	19.95	1.21	8.72	2.28
10–<50%	7,204	2.53	6.13	1.65	4.37	18.61	7.76	2.48	8.18	1.85
<10%	18,843	1.79	4.52	0.99	4.13	20.75	10.70	1.61	6.43	1.69
% Black or African American in Segment										
50–100%	1,840	3.53	8.23	2.19	5.92	26.52	13.71	2.88	11.95	3.19
10–<50%	5,535	2.57	6.48	1.74	4.59	25.34	11.99	2.38	8.20	1.80
<10%	21,403	1.66	4.33	1.06	3.89	20.24	10.80	1.54	6.50	1.68
% Owner-Occupied DUs[1] in Segment										
50–100%	21,550	1.47	3.62	0.86	3.68	20.64	11.08	1.65	7.41	1.97
10–<50%	5,763	3.21	8.74	2.50	5.33	22.04	8.44	2.60	6.68	1.12
<10%	1,465	4.16	9.75	2.22	6.62	38.65	25.15	0.61	5.17	2.09
Combined Median Rent/Housing Value										
1st Quintile	4,645	1.46	3.99	1.08	3.51	20.43	12.12	1.49	4.05	0.97
2nd Quintile	6,462	1.66	3.63	0.88	4.16	25.28	14.14	1.36	5.98	1.37
3rd Quintile	6,791	2.31	6.53	1.83	4.24	21.77	10.12	2.19	9.24	2.00
4th Quintile	6,191	1.97	4.82	1.19	4.23	20.76	10.57	2.04	7.82	2.12
5th Quintile	4,689	2.32	5.91	1.34	4.56	19.64	9.35	1.77	7.65	2.28
Population Density										
Large MSA[1]	12,726	2.51	6.28	1.56	4.72	25.02	13.09	2.11	8.24	2.16
Medium to Small MSA[1]	13,766	1.62	3.90	1.06	3.66	17.68	9.32	1.54	6.32	1.48
Non-MSA,[1] Urban	792	1.01	1.65	0.39	3.66	15.59	5.35	1.01	0.87	0.17
Non-MSA,[1] Rural	1,494	0.87	0.75	0.11	4.15	14.36	4.47	1.81	3.83	0.50
Group Quarters										
Group	201	4.48	14.28	1.93	9.95	35.80	11.51	8.96	31.62	5.69
Non-Group	28,577	1.94	5.02	1.29	4.12	21.70	11.25	1.74	7.17	1.80

[1] This step used demographic variables from screener data for all selected person pairs; DU = dwelling unit, MSA = metropolitan statistical area, pr = pair, ps = poststratification adjustment, SDU = screener dwelling unit, sel = selected.

[2] Weighted extreme value proportion: $100*\sum_k w_{ek}/\sum_k w_k$, where w_{ek} denotes the weight for extreme values, and w_k denotes the weight for both extreme values and nonextreme values.

[3] Outwinsor weight proportion: $100*\sum_k(w_{ek} - b_k)/\sum_k w_k$, where b_k denotes the winsorized weight.

Table J.2 2017 NSDUH Respondent Pair-Level Proportions of Extreme Values and Outwinsors

Domain	*n*	Before res.pr.nr[1] (SDUWT*PRWT12*PRWT13)			After res.pr.nr[1] (SDUWT*PRWT12*...*PRWT14)		
		% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]
Total	17,704	2.03	9.62	2.30	2.36	11.04	2.89
Pair Age Group							
12–17, 12–17	2,371	0.93	5.44	0.91	0.34	1.60	0.28
12–17, 18–25	1,755	1.94	6.69	0.80	1.42	6.73	1.08
12–17, 26–34	910	1.21	6.60	1.69	0.66	4.05	0.35
12–17, 35–49	3,862	0.60	1.78	0.29	0.60	2.85	0.54
12–17, 50+	763	0.26	1.21	0.14	0.26	1.20	0.72
18–25, 18–25	2,524	3.45	11.10	1.65	4.71	14.61	2.56
18–25, 26–34	812	3.57	8.85	1.21	6.03	16.24	3.14
18–25, 35–49	1,206	3.81	7.03	0.66	4.48	10.32	1.97
18–25, 50+	670	1.49	2.54	0.19	1.64	4.44	1.62
26–34, 26–34	841	1.07	3.22	0.46	2.26	8.89	2.22
26–34, 35–49	447	2.24	4.33	0.56	3.80	8.52	2.50
26–34, 50+	235	2.55	12.12	3.30	1.70	8.24	0.41
35–49, 35–49	627	3.03	5.92	0.74	6.38	14.28	2.97
35–49, 50+	211	8.06	34.53	9.24	6.64	24.67	7.97
50+, 50+	470	7.23	22.94	7.26	5.74	19.33	6.22
Pair Race/Ethnicity							
Hispanic or Latino	3,083	1.88	10.21	2.65	2.21	12.36	3.48
Black or African American	1,918	2.97	13.52	3.86	1.88	12.18	3.07
White	9,585	1.49	7.90	1.69	1.89	8.55	2.33
Other	1,212	2.56	10.96	2.76	5.61	23.60	6.16
White & Black or African American	205	1.95	26.09	10.72	3.90	31.31	5.70
White & Hispanic or Latino	729	2.74	6.91	1.14	2.19	5.28	1.06
White & Other	630	4.44	14.25	1.97	2.22	3.81	0.60
Black or African American & Hispanic or Latino	88	14.77	16.72	2.22	18.18	22.41	5.47
Black or African American & Other	103	2.91	4.79	0.90	6.80	18.82	3.78
Hispanic or Latino & Other	151	1.32	1.70	0.73	2.65	13.02	1.98
Pair Gender							
Male, Male	3,609	2.30	5.45	0.81	3.08	7.80	1.87
Female, Female	4,056	2.17	8.38	2.00	2.02	10.57	2.76
Male, Female	10,039	1.87	11.20	2.82	2.24	12.07	3.22
Household Size							
Two	4,325	0.65	2.94	0.42	0.69	2.59	0.56
Three	5,674	1.53	9.38	2.27	2.40	11.94	3.36
Four or More	7,705	3.17	12.97	3.23	3.27	14.81	3.83

Table J.2 2017 NSDUH Respondent Pair-Level Proportions of Extreme Values and Outwinsors (continued)

Domain	*n*	Before res.pr.nr[1] (SDUWT*PRWT12*PRWT13)			After res.pr.nr[1] (SDUWT*PRWT12*...*PRWT14)		
		% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]
Census Region							
Northeast	3,346	1.73	8.71	1.73	3.17	13.23	3.73
South	5,938	1.82	10.06	2.63	2.17	12.97	3.74
Midwest	4,162	2.76	10.36	2.52	2.26	8.30	1.69
West	4,258	1.83	8.93	1.98	2.09	8.88	2.05
Quarter							
Quarter 1	4,059	2.96	11.23	2.11	3.08	11.52	2.42
Quarter 2	4,566	2.04	9.94	2.48	2.01	9.59	2.13
Quarter 3	4,480	1.63	9.93	2.70	2.66	12.91	3.66
Quarter 4	4,599	1.59	7.47	1.94	1.78	10.12	3.36
% Hispanic or Latino in Segment							
50–100%	1,664	1.14	8.41	2.16	1.14	10.14	3.57
10–<50%	4,435	2.82	10.89	1.97	3.02	11.12	2.47
<10%	11,605	1.85	9.24	2.49	2.28	11.18	2.97
% Black or African American in Segment							
50–100%	1,255	3.51	17.01	5.80	3.35	16.00	4.07
10–<50%	3,482	2.44	10.25	2.27	2.53	13.12	3.45
<10%	12,967	1.77	8.61	1.92	2.22	9.97	2.62
% Owner-Occupied DUs[1] in Segment							
50–100%	13,232	1.87	9.44	2.36	2.27	10.92	2.86
10–<50%	3,584	2.96	10.93	2.19	2.85	11.74	3.21
<10%	888	0.56	3.04	0.52	1.80	8.51	1.23
Combined Median Rent/Housing Value							
1st Quintile	3,052	1.54	6.73	1.49	1.57	9.66	2.16
2nd Quintile	4,204	1.52	8.27	2.19	1.31	9.09	2.53
3rd Quintile	4,156	2.57	11.10	2.45	2.41	11.42	3.46
4th Quintile	3,657	2.13	9.38	2.58	3.23	11.53	2.56
5th Quintile	2,635	2.39	11.98	2.55	3.68	13.02	3.50
Population Density							
Large MSA[1]	7,503	2.49	11.49	2.82	2.97	11.98	3.13
Medium to Small MSA[1]	8,751	1.71	8.21	1.91	1.95	10.09	2.79
Non-MSA,[1] Urban	539	1.86	3.52	0.23	2.23	10.39	1.32
Non-MSA,[1] Rural	911	1.32	0.73	0.12	1.32	5.91	1.09
Group Quarters							
Group	150	6.00	20.60	4.46	4.67	18.83	4.80
Non-Group	17,554	1.99	9.59	2.29	2.34	11.02	2.89

[1] This step used demographic variables from screener data for all responding person pairs; DU = dwelling unit, MSA = metropolitan statistical area, nr = nonresponse adjustment, pr = pair, res = respondent, SDU = screener dwelling unit.

[2] Weighted extreme value proportion: $100*\sum_k w_{ek}/\sum_k w_k$, where w_{ek} denotes the weight for extreme values, and w_k denotes the weight for both extreme values and nonextreme values.

[3] Outwinsor weight proportion: $100*\sum_k(w_{ek} - b_k)/\sum_k w_k$, where b_k denotes the winsorized weight.

Table J.3 2017 NSDUH Respondent Pair-Level Proportions of Extreme Values and Outwinsors

Domain	*n*	Before res.pr.ps[1] (SDUWT*PRWT12*...*PRWT14)			After res.pr.ps[1] (SDUWT*PRWT12*...*PRWT15)			Final Weight: After res.pr.ev[1] (SDUWT*PRWT12*...*PRWT16)		
		% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]
Total	17,704	2.35	8.35	1.67	0.94	2.87	0.39	0.38	1.41	0.11
Pair Age Group										
12–17, 12–17	2,368	0.38	2.13	0.46	0.42	2.62	0.30	0.34	2.14	0.14
12–17, 18–25	1,758	1.42	6.62	1.10	0.40	1.93	0.43	0.23	0.99	0.07
12–17, 26–34	894	0.78	4.17	0.40	0.67	4.90	0.53	0.22	1.51	0.06
12–17, 35–49	3,870	0.78	3.42	0.56	0.31	1.70	0.14	0.28	1.44	0.05
12–17, 50+	771	0.26	1.18	0.25	0.52	1.64	0.16	0.39	1.30	0.13
18–25, 18–25	2,467	4.82	14.69	3.03	2.15	5.94	0.42	0.45	1.77	0.12
18–25, 26–34	839	5.60	15.10	2.94	1.79	7.31	1.45	0.72	3.31	0.23
18–25, 35–49	1,176	5.53	13.42	3.61	1.79	4.39	0.37	0.60	1.44	0.08
18–25, 50+	683	1.32	4.36	1.50	1.02	3.27	0.47	0.59	2.16	0.21
26–34, 26–34	871	2.18	8.66	2.13	0.34	1.72	0.49	0.23	1.59	0.15
26–34, 35–49	445	3.60	10.25	2.67	1.57	4.21	0.38	0.45	2.61	0.28
26–34, 50+	249	0.00	0.00	0.00	0.40	1.51	0.28	0.40	1.42	0.19
35–49, 35–49	628	6.05	9.22	1.74	2.71	5.97	1.15	0.80	2.41	0.21
35–49, 50+	213	4.69	12.10	0.68	0.94	2.62	0.47	0.47	0.97	0.02
50+, 50+	472	4.24	13.22	2.69	0.42	0.83	0.06	0.00	0.00	0.00
Pair Race/Ethnicity										
Hispanic or Latino	3,116	2.18	10.40	2.32	0.93	2.46	0.34	0.45	1.05	0.05
Black or African American	1,881	1.70	4.07	0.77	0.64	1.80	0.22	0.05	0.01	0.00
White	9,247	1.92	6.67	1.07	0.68	1.83	0.17	0.02	0.18	0.00
Other	1,183	6.34	22.29	5.02	3.55	12.21	2.21	3.72	12.12	0.99
White & Black or African American	189	2.12	3.28	0.38	1.59	7.11	0.67	0.00	0.00	0.00
White & Hispanic or Latino	765	2.22	3.92	0.82	0.52	0.46	0.10	0.00	0.00	0.00
White & Other	815	2.45	5.98	1.51	0.61	2.66	0.42	0.12	1.84	0.39
Black or African American & Hispanic or Latino	120	10.00	21.26	4.63	5.00	11.45	0.89	4.17	11.06	0.48
Black or African American & Other	204	2.94	2.75	0.59	0.00	0.00	0.00	0.00	0.00	0.00
Hispanic or Latino & Other	184	2.17	20.22	8.40	1.63	0.75	0.06	0.00	0.00	0.00
Pair Gender										
Male, Male	3,609	3.33	8.24	2.02	1.52	4.64	0.54	0.61	2.50	0.18
Female, Female	4,054	1.97	7.22	1.46	0.89	2.46	0.32	0.49	1.31	0.09
Male, Female	10,041	2.15	8.71	1.64	0.76	2.49	0.37	0.25	1.13	0.09
Household Size										
Two	4,325	0.72	2.86	0.65	0.30	1.52	0.41	0.16	0.65	0.06
Three	5,674	2.38	11.25	2.44	0.78	3.61	0.39	0.35	1.68	0.11
Four or More	7,705	3.24	9.62	1.80	1.43	3.17	0.38	0.52	1.65	0.13

Table J.3 2017 NSDUH Respondent Pair-Level Proportions of Extreme Values and Outwinsors (continued)

Domain	*n*	Before res.pr.ps[1] (SDUWT*PRWT12*...*PRWT14)			After res.pr.ps[1] (SDUWT*PRWT12*...*PRWT15)			Final Weight: After res.pr.ev[1] (SDUWT*PRWT12*...*PRWT16)		
		% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]
Census Region										
Northeast	3,346	3.32	12.73	2.88	2.09	6.37	1.11	1.32	4.27	0.37
South	5,938	2.11	8.92	1.92	0.83	3.03	0.41	0.39	1.78	0.11
Midwest	4,162	2.16	6.17	0.92	0.77	1.96	0.14	0.00	0.00	0.00
West	4,258	2.11	6.23	1.07	0.38	0.95	0.05	0.00	0.00	0.00
Quarter										
Quarter 1	4,059	3.10	8.51	2.01	1.23	3.54	0.47	0.44	1.45	0.09
Quarter 2	4,566	1.91	6.42	1.20	0.92	3.15	0.43	0.39	1.97	0.18
Quarter 3	4,480	2.68	9.51	1.74	0.92	2.72	0.40	0.42	1.24	0.08
Quarter 4	4,599	1.80	8.97	1.75	0.74	2.08	0.27	0.26	0.97	0.07
% Hispanic or Latino in Segment										
50–100%	1,664	1.14	9.69	2.40	0.66	1.38	0.13	0.30	0.57	0.03
10–<50%	4,435	3.02	7.98	1.63	1.53	5.57	0.80	0.86	2.85	0.21
<10%	11,605	2.27	8.26	1.55	0.76	1.81	0.24	0.21	0.85	0.07
% Black or African American in Segment										
50–100%	1,255	2.95	7.86	2.11	1.35	3.93	0.58	0.72	1.82	0.16
10–<50%	3,482	2.41	8.40	1.50	1.21	4.58	0.71	0.75	3.62	0.28
<10%	12,967	2.28	8.39	1.68	0.83	2.29	0.28	0.25	0.75	0.05
% Owner-Occupied DUs[1] in Segment										
50–100%	13,232	2.21	7.64	1.42	0.76	2.19	0.25	0.25	0.85	0.07
10–<50%	3,584	3.04	11.35	2.75	1.67	5.63	0.97	0.84	3.67	0.27
<10%	888	1.58	7.45	1.24	0.68	3.27	0.40	0.45	1.78	0.11
Combined Median Rent/Housing Value										
1st Quintile	3,052	1.44	5.62	1.08	0.62	2.42	0.26	0.10	0.72	0.09
2nd Quintile	4,204	1.43	5.96	1.10	0.55	1.55	0.24	0.21	0.82	0.08
3rd Quintile	4,156	2.45	10.42	1.68	0.79	2.82	0.53	0.31	1.34	0.12
4th Quintile	3,657	3.23	8.58	2.12	1.59	4.33	0.62	0.71	2.16	0.15
5th Quintile	2,635	3.49	10.05	2.17	1.29	2.94	0.21	0.61	1.71	0.08
Population Density										
Large MSA[1]	7,503	2.85	8.29	1.80	1.51	3.83	0.53	0.77	2.04	0.13
Medium to Small MSA[1]	8,751	2.03	8.58	1.56	0.54	1.77	0.23	0.10	0.65	0.09
Non-MSA,[1] Urban	539	2.23	10.39	1.27	0.37	0.38	0.04	0.00	0.00	0.00
Non-MSA,[1] Rural	911	1.32	5.91	1.16	0.55	0.23	0.01	0.00	0.00	0.00
Group Quarters										
Group	150	4.00	16.65	5.05	0.00	0.00	0.00	0.00	0.00	0.00
Non-Group	17,554	2.34	8.33	1.67	0.95	2.88	0.39	0.38	1.41	0.11

Table J.3 2017 NSDUH Respondent Pair-Level Proportions of Extreme Values and Outwinsors (continued)

Domain	*n*	Before res.pr.ps[1] (SDUWT*PRWT12*...*PRWT14)			After res.pr.ps[1] (SDUWT*PRWT12*...*PRWT15)			Final Weight: After res.pr.ev[1] (SDUWT*PRWT12*...*PRWT16)		
		% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]
Pair Relationship Domain[4]										
Parent-Child (12-14)	2,810	0.85	3.23	0.38	0.39	1.87	0.16	0.25	0.94	0.07
Parent-Child (12–17)	5,049	0.69	3.10	0.50	0.38	1.94	0.19	0.26	1.39	0.07
Parent-Child (12-20)	5,901	1.32	4.53	0.70	0.51	2.41	0.20	0.32	1.57	0.08
Sibling (12-14)-Sibling (15-17)	1,397	0.14	0.39	0.06	0.43	2.63	0.39	0.21	1.32	0.10
Sibling (12–17)-Sibling (18–25)	1,583	1.52	6.80	1.10	0.44	2.09	0.46	0.25	1.07	0.08
Spouse-Spouse/Partner-Partner	3,467	2.02	9.20	1.58	0.55	1.78	0.33	0.23	1.00	0.10
Spouse-Spouse/Partner-Partner with Children (Younger Than 18)	1,711	2.16	10.58	1.72	1.05	4.08	0.77	0.41	2.15	0.21

[1] This step used demographic variables from questionnaire data for all responding person pairs; DU = dwelling unit, ev = extreme value adjustment, MSA = metropolitan statistical area, pr = pair, ps = poststratification adjustment, res = respondent, SDU = screener dwelling unit.

[2] Weighted extreme value proportion: $100*\sum_k w_{ek}/\sum_k w_k$, where w_{ek} denotes the weight for extreme values, and w_k denotes the weight for both extreme values and nonextreme values.

[3] Outwinsor weight proportion: $100*\sum_k (w_{ek} - b_k)/\sum_k w_k$, where b_k denotes the winsorized weight.

[4] Parent-child (15-17) was not included here since extreme values were not controlled with this domain.

This page intentionally left blank

Appendix K: Evaluation of Calibration Weights: Pair-Level Slippage Rates

This page intentionally left blank

Table K.1 2017 NSDUH Respondent Pair-Level Slippage Rates

Domain	*n*	Initial Total (*I*)[1]	Final Total (*F*)[2]	Control Total from SDU (*C*)	(*I* - *C*)/*C*%	(*F* - *C*)/*C*%
Total	17,704	231,612,324	231,612,324	231,612,324	-0.00	0.00
Pair Age Group						
12–17, 12–17	2,368	7,324,911	7,284,264	7,284,264	0.56	0.00
12–17, 18–25	1,758	8,013,602	7,992,602	7,992,602	0.26	0.00
12–17, 26–34	894	4,854,885	4,878,622	4,878,622	-0.49	0.00
12–17, 35–49	3,870	28,956,026	28,987,236	28,987,236	-0.11	0.00
12–17, 50+	771	13,642,863	13,433,303	13,433,303	1.56	0.00
18–25, 18–25	2,467	12,390,009	12,591,242	12,591,242	-1.60	0.00
18–25, 26–34	839	6,996,757	7,154,932	7,154,932	-2.21	-0.00
18–25, 35–49	1,176	16,429,426	16,363,237	16,363,237	0.40	-0.00
18–25, 50+	683	20,375,398	20,228,372	20,228,372	0.73	-0.00
26–34, 26–34	871	12,175,597	11,604,298	11,604,298	4.92	-0.00
26–34, 35–49	445	8,182,814	8,580,123	8,580,123	-4.63	-0.00
26–34, 50+	249	13,707,437	13,662,522	13,662,522	0.33	-0.00
35–49, 35–49	628	18,452,992	18,703,711	18,703,711	-1.34	-0.00
35–49, 50+	213	17,689,609	17,673,729	17,673,729	0.09	0.00
50+, 50+	472	42,419,998	42,474,129	42,474,129	-0.13	0.00
Pair Race/Ethnicity						
Hispanic or Latino	3,116	42,563,144	42,642,479	42,642,479	-0.19	-0.00
Black or African American	1,881	23,644,438	24,610,120	24,610,120	-3.92	0.00
White	9,247	117,830,822	121,979,499	121,979,499	-3.40	0.00
Other	1,183	18,703,665	18,807,852	18,807,852	-0.55	0.00
White & Black or African American	189	2,689,265	2,753,841	2,753,841	-2.34	0.00
White & Hispanic or Latino	765	10,326,191	9,841,953	9,841,953	4.92	0.00
White & Other	815	9,137,016	6,662,557	6,662,557	37.14	0.00
Black or African American & Hispanic or Latino	120	1,817,027	1,722,974	1,722,974	5.46	0.00
Black or African American & Other	204	2,780,680	1,016,681	1,016,681	173.51	0.00
Hispanic or Latino & Other	184	2,120,076	1,574,368	1,574,368	34.66	0.00
Pair Gender						
Male, Male	3,609	41,204,571	41,284,758	41,284,758	-0.19	-0.00
Female, Female	4,054	42,216,197	42,208,910	42,208,910	0.02	0.00
Male, Female	10,041	148,191,556	148,118,656	148,118,656	0.05	-0.00
Pair Relationship Domain[3,4,5]						
Parent-Child (12–14)*	2,810	11,674,474	12,839,100	12,839,100	-9.07	0.00
Parent-Child (12–17)*	5,049	23,799,976	25,533,394	25,533,394	-6.79	0.00
Parent-Child (15–17)*	2,239	12,125,502	12,694,293	12,694,293	-4.48	0.00
Parent-Child (12–20)*	5,901	32,908,524	34,657,985	34,657,985	-5.05	0.00
Parent*-Child (12–14)	2,810	17,634,419	19,672,411	19,672,411	-10.36	0.00
Parent*-Child (12–17)	5,049	30,864,749	32,780,882	32,780,882	-5.85	0.00
Parent*-Child (15–17)	2,239	18,910,589	19,096,877	19,462,074	-2.83	-1.88
Parent*-Child (12–20)	5,901	39,065,918	40,558,323	40,558,323	-3.68	0.00
Sibling (12–14)-Sibling (15–17)*	1,397	3,667,270	4,036,271	4,036,271	-9.14	0.00
Sibling (12–17)-Sibling (18–25)*	1,583	6,072,503	6,307,282	6,307,282	-3.72	0.00
Spouse-Spouse/Partner-Partner	3,467	77,826,506	76,125,486	76,125,486	2.23	-0.00
Spouse-Spouse/Partner-Partner with Children (Younger Than 18)	1,711	27,509,375	30,338,039	30,338,039	-9.32	-0.00

(continued)

Table K.1 2017 NSDUH Respondent Pair-Level Slippage Rates (continued)

Domain	*n*	Initial Total (*I*)[1]	Final Total (*F*)[2]	Control Total from SDU (*C*)	(*I* - *C*)/*C*%	(*F* - *C*)/*C*%
Household Size						
Two	4,325	57,711,968	57,711,968	57,711,968	0.00	-0.00
Three	5,674	58,983,740	58,983,740	58,983,740	-0.00	0.00
Four or More	7,705	114,916,616	114,916,616	114,916,616	0.00	0.00
Census Region						
Northeast	3,346	41,047,147	41,047,147	41,047,147	-0.00	0.00
South	5,938	84,683,572	84,683,572	84,683,572	-0.00	0.00
Midwest	4,162	45,290,792	45,290,792	45,290,792	0.00	-0.00
West	4,258	60,590,812	60,590,812	60,590,812	-0.00	-0.00
Quarter						
Quarter 1	4,059	57,766,086	57,766,086	57,766,086	0.00	0.00
Quarter 2	4,566	57,843,136	57,843,137	57,843,137	-0.00	-0.00
Quarter 3	4,480	57,989,093	57,989,093	57,989,093	0.00	-0.00
Quarter 4	4,599	58,014,008	58,014,008	58,014,008	0.00	0.00
% Hispanic or Latino in Segment						
50–100%	1,664	28,169,477	28,169,477	28,169,477	-0.00	-0.00
10–<50%	4,435	68,404,923	68,404,923	68,404,923	-0.00	0.00
<10%	11,605	135,037,923	135,037,923	135,037,923	0.00	0.00
% Black or African American in Segment						
50–100%	1,255	16,323,467	16,323,467	16,323,467	-0.00	0.00
10–<50%	3,482	47,104,333	47,104,333	47,104,333	0.00	-0.00
<10%	12,967	168,184,524	168,184,524	168,184,524	-0.00	0.00
% Owner-Occupied DUs in Segment						
50–100%	13,232	183,070,086	183,070,086	183,070,086	-0.00	-0.00
10–<50%	3,584	44,541,285	44,541,285	44,541,285	0.00	0.00
<10%	888	4,000,952	4,000,952	4,000,952	-0.00	-0.00
Combined Median Rent/Housing Value						
1st Quintile	3,052	31,380,370	31,380,370	31,380,370	0.00	-0.00
2nd Quintile	4,204	47,983,686	47,983,686	47,983,686	0.00	0.00
3rd Quintile	4,156	54,509,285	54,509,285	54,509,285	0.00	-0.00
4th Quintile	3,657	52,764,087	52,764,087	52,764,087	0.00	0.00
5th Quintile	2,635	44,974,896	44,974,896	44,974,896	-0.00	0.00
Population Density						
Large MSA	7,503	132,414,913	132,414,913	132,414,913	-0.00	0.00
Medium to Small MSA	8,751	86,718,978	86,718,978	86,718,978	0.00	-0.00
Non-MSA, Urban	539	4,505,927	4,505,927	4,505,927	0.00	0.00
Non-MSA, Rural	911	7,972,506	7,972,506	7,972,506	0.00	0.00
Group Quarters						
Group	150	561,103	561,103	561,103	0.00	0.00
Non-Group	17,554	231,051,221	231,051,221	231,051,221	-0.00	0.00

DU = dwelling unit, MSA = metropolitan statistical area, SDU = screener dwelling unit.

[1] WT1*...*WT11*PRWT12*...*PRWT14 (before respondent person pair poststratification and respondent person pair extreme value adjustment).

[2] WT1*...*WT11*PRWT12*...*PRWT16 (after respondent person pair poststratification and respondent person pair extreme value adjustment).

[3] The member of the pair that is the focus is designated with an asterisk (*).

[4] The parent-child (15–17) pair domains were not controlled for within the modeling and thus have higher slippage rates than the other domains listed. However, since these domains are a subset of other controlled domains, the rates are not large.

[5] Slippage rates were not calculated for the sibling-sibling domains with the younger child as the focus since no household counts for this domain were calculated and are required to construct the appropriate controls totals.

Appendix L: Evaluation of Calibration Weights: Pair-Level Weight Summary Statistics

This page intentionally left blank

Table L.1 2017 NSDUH Selected Pair-Level Weight Summary Statistics

Domain	*n*	SDU-Level Weights[1] (SDUWT: WT1*...*WT11)						Before sel.pr.ps[1] (SDUWT*PRWT12)						After sel.pr.ps[1] (SDUWT*PRWT12*PRWT13)					
		Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
Total	28,778	7	543	912	1,305	8,679	1.43	11	1,627	3,452	7,853	2,350,151	13.88	10	1,519	3,519	8,096	290,613	4.74
Pair Age Group																			
12–17, 12–17	3,261	9	420	829	1,208	5,613	1.45	11	834	1,591	2,689	42,626	2.19	10	655	1,589	2,863	25,669	2.20
12–17, 18–25	2,679	10	580	944	1,328	7,475	1.39	33	1,136	1,985	3,652	56,361	2.39	33	1,119	2,123	3,860	23,111	1.91
12–17, 26–34	1,338	10	530	880	1,313	5,800	1.47	101	1,590	2,687	4,498	88,542	2.61	38	1,164	2,250	4,195	107,081	3.30
12–17, 35–49	5,845	12	468	850	1,210	7,049	1.47	66	1,799	3,293	5,715	112,703	2.45	63	1,603	3,251	6,052	76,715	2.27
12–17, 50+	1,211	7	579	939	1,398	4,806	1.38	74	4,836	8,478	12,593	191,602	2.10	84	4,177	8,373	13,353	94,414	2.00
18–25, 18–25	4,167	12	541	959	1,354	6,209	1.42	97	1,090	1,938	3,331	96,871	2.74	56	948	1,948	4,028	25,621	2.09
18–25, 26–34	1,443	19	593	955	1,370	8,022	1.46	226	1,884	3,393	5,494	120,269	2.74	102	1,545	3,055	6,049	60,860	2.50
18–25, 35–49	2,084	25	602	928	1,311	8,679	1.44	90	2,805	4,947	8,481	167,591	3.07	112	2,689	5,286	9,572	58,442	2.08
18–25, 50+	1,309	49	764	1,088	1,523	4,884	1.30	872	6,941	10,816	16,747	415,871	3.10	553	6,543	11,725	19,356	82,740	1.77
26–34, 26–34	1,551	46	533	900	1,291	6,332	1.45	306	3,422	5,922	9,239	234,308	2.35	166	2,821	5,351	9,051	134,560	2.25
26–34, 35–49	810	32	592	936	1,298	5,354	1.39	453	4,351	7,588	11,635	482,443	5.92	111	3,796	6,996	11,601	172,536	2.96
26–34, 50+	460	60	706	1,074	1,495	4,035	1.35	1,236	13,719	23,044	33,475	1,415,464	6.09	1,045	11,828	22,828	37,240	261,310	1.99
35–49, 35–49	1,233	27	566	923	1,284	5,469	1.39	334	4,543	8,612	13,190	570,940	7.82	551	4,635	9,027	15,184	162,725	3.14
35–49, 50+	474	20	597	981	1,382	6,492	1.42	1,009	11,700	21,432	32,833	872,615	4.43	1,354	12,589	24,298	45,304	274,340	2.25
50+, 50+	913	40	606	937	1,330	4,530	1.35	1,900	20,331	35,894	47,938	2,350,151	7.03	3,420	22,134	39,343	55,408	290,613	1.74
Pair Race/Ethnicity																			
Hispanic or Latino	4,991	7	651	1,055	1,475	8,679	1.40	11	2,021	3,984	8,564	2,053,194	19.48	19	1,853	3,990	8,871	290,613	4.64
Black or African American	2,771	27	745	1,079	1,465	6,643	1.31	64	2,058	3,927	8,592	474,410	7.25	27	1,856	4,035	8,496	261,922	5.18
White	15,850	16	485	874	1,184	4,338	1.38	25	1,493	3,190	7,534	1,415,464	9.66	18	1,412	3,258	7,715	285,564	4.78
Other	2,162	21	333	815	1,487	5,469	1.64	29	1,270	3,053	7,764	2,350,151	39.04	10	1,154	3,197	8,671	239,957	4.69
White & Black or African American	316	40	761	1,096	1,382	3,830	1.40	201	1,941	4,120	8,996	773,508	18.00	108	1,714	4,046	7,556	250,895	6.68
White & Hispanic or Latino	1,192	32	576	940	1,425	4,873	1.42	84	1,794	3,960	8,699	136,934	3.15	57	1,656	4,303	9,188	149,094	3.44
White & Other	965	12	380	733	1,194	6,327	1.58	73	1,392	3,086	6,704	247,222	5.84	104	1,366	3,300	7,647	142,364	4.08
Black or African American & Hispanic or Latino	146	57	617	1,168	1,970	6,467	1.67	169	1,779	4,445	10,570	102,404	3.14	198	2,926	6,903	12,511	107,081	2.80
Black or African American & Other	157	47	570	900	1,201	4,063	1.40	323	1,775	3,084	5,384	69,906	3.48	134	1,418	3,910	6,689	104,819	4.15
Hispanic or Latino & Other	228	24	326	808	1,347	6,600	1.89	100	1,181	3,056	7,855	202,231	6.53	56	922	2,556	6,930	115,741	5.20

(continued)

Table L.1 2017 NSDUH Selected Pair-Level Weight Summary Statistics (continued)

Domain	*n*	SDU-Level Weights[1] (SDUWT: WT1*...*WT11)						Before sel.pr.ps[1] (SDUWT*PRWT12)						After sel.pr.ps[1] (SDUWT*PRWT12*PRWT13)					
		Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
Pair Gender																			
Male, Male	6,140	7	554	909	1,325	6,467	1.43	11	1,572	3,209	7,265	872,615	6.91	18	1,428	3,408	7,414	193,571	3.66
Female, Female	6,164	10	514	910	1,318	8,679	1.47	28	1,591	3,272	6,940	1,415,464	14.91	19	1,494	3,320	7,220	273,810	4.62
Male, Female	16,474	10	548	914	1,292	8,022	1.41	25	1,666	3,637	8,411	2,350,151	14.87	10	1,553	3,643	8,706	290,613	4.88
Household Size																			
Two	7,318	24	546	903	1,245	6,307	1.39	84	1,633	3,788	8,612	183,922	3.13	38	1,254	3,222	8,112	143,788	3.67
Three	9,133	20	531	898	1,300	8,022	1.43	31	1,447	2,857	5,407	2,053,194	24.40	10	1,444	3,049	6,239	290,613	5.09
Four or More	12,327	7	551	930	1,354	8,679	1.45	11	1,765	4,049	9,257	2,350,151	13.99	19	1,780	4,187	9,605	285,564	4.88
Census Region																			
Northeast	5,746	15	310	805	1,061	5,949	1.47	29	1,299	2,760	6,695	593,743	8.54	10	1,166	2,820	7,089	273,810	5.47
South	9,326	7	725	1,072	1,498	8,679	1.34	11	2,139	4,236	9,065	1,415,464	11.08	18	1,964	4,253	9,088	290,613	4.52
Midwest	6,765	51	613	855	1,095	5,178	1.30	97	1,578	3,059	6,598	872,615	7.64	105	1,521	3,185	6,810	220,458	4.31
West	6,941	21	356	897	1,487	8,022	1.56	34	1,348	3,590	9,064	2,350,151	22.53	56	1,254	3,530	9,226	248,443	4.64
Quarter																			
Quarter1	6,630	22	594	981	1,413	8,679	1.43	34	1,802	3,779	8,495	2,053,194	14.49	24	1,760	4,007	8,933	261,922	4.23
Quarter2	7,446	10	544	869	1,256	7,498	1.42	28	1,544	3,336	7,481	1,415,464	13.64	18	1,450	3,466	7,945	290,613	4.67
Quarter3	7,307	7	488	899	1,313	6,643	1.45	11	1,558	3,347	7,726	2,350,151	19.45	22	1,381	3,216	7,652	278,811	5.62
Quarter4	7,395	16	562	901	1,266	7,049	1.41	25	1,625	3,391	7,828	564,595	6.17	10	1,513	3,431	8,147	285,564	4.45
% Hispanic or Latino in Segment																			
50–100%	2,731	46	782	1,251	1,597	8,679	1.29	73	2,609	4,993	10,928	2,053,194	19.68	27	2,068	4,558	10,509	290,613	4.42
10–<50%	7,204	7	668	1,062	1,546	8,022	1.38	11	2,077	4,114	9,088	398,064	5.10	19	2,014	4,558	9,821	261,310	4.04
<10%	18,843	12	413	845	1,158	6,332	1.44	25	1,397	3,005	7,003	2,350,151	15.38	10	1,326	3,034	7,113	285,564	5.10
% Black or African American in Segment																			
50–100%	1,840	7	729	1,035	1,345	6,298	1.33	11	2,042	3,844	8,518	773,508	9.66	22	1,923	4,142	8,947	261,922	5.03
10–<50%	5,535	12	683	1,006	1,409	8,022	1.36	28	1,930	3,855	8,350	1,067,384	9.82	10	1,726	3,987	8,602	290,613	4.47
<10%	21,403	12	462	879	1,271	8,679	1.46	29	1,512	3,316	7,690	2,350,151	15.58	18	1,425	3,364	7,882	285,564	4.78
% Owner-Occupied DUs[1] in Segment																			
50–100%	21,550	7	533	894	1,271	7,498	1.41	11	1,631	3,490	8,025	2,350,151	13.07	18	1,625	3,681	8,468	290,613	4.71
10–<50%	5,763	12	595	987	1,390	8,679	1.46	30	1,686	3,503	7,598	474,410	5.94	24	1,652	3,723	8,128	218,170	4.20
<10%	1,465	32	521	951	1,427	6,188	1.51	43	1,379	2,942	6,116	2,053,194	59.59	10	534	1,168	2,786	115,741	6.84
Combined Median Rent/Housing Value																			
1st Quintile	4,645	27	378	796	1,159	6,643	1.45	34	1,302	2,781	6,393	2,350,151	28.97	27	1,177	2,771	6,383	239,957	4.83
2nd Quintile	6,462	15	489	877	1,249	7,498	1.40	36	1,501	3,181	7,345	2,053,194	18.58	27	1,353	3,138	7,468	261,922	4.87
3rd Quintile	6,791	16	534	901	1,290	8,679	1.48	25	1,617	3,444	7,731	1,067,384	10.36	10	1,522	3,648	8,165	290,613	4.78
4th Quintile	6,191	7	592	940	1,367	6,327	1.41	11	1,775	3,801	8,252	1,415,464	11.23	22	1,641	3,806	8,572	277,251	4.70
5th Quintile	4,689	10	733	1,071	1,482	7,049	1.35	28	2,005	4,310	9,733	564,595	6.47	18	1,971	4,519	10,019	285,564	4.36

(continued)

Table L.1 2017 NSDUH Selected Pair-Level Weight Summary Statistics (continued)

Domain	*n*	SDU-Level Weights[1] (SDUWT: WT1*...*WT11)						Before sel.pr.ps[1] (SDUWT*PRWT12)						After sel.pr.ps[1] (SDUWT*PRWT12*PRWT13)					
		Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
Population Density																			
Large MSA[1]	12,726	7	831	1,127	1,549	8,679	1.29	11	2,427	4,788	10,084	2,053,194	11.67	18	2,366	5,027	10,618	290,613	4.13
Medium to Small MSA[1]	13,766	15	337	747	1,101	6,643	1.49	29	1,202	2,647	6,168	2,350,151	17.07	10	1,129	2,596	6,138	277,251	5.17
Non-MSA,[1] Urban	792	24	328	732	1,033	3,815	1.45	84	1,096	2,502	5,290	139,921	4.81	88	1,038	2,491	5,300	114,998	4.57
Non-MSA,[1] Rural	1,494	27	195	570	942	2,829	1.56	36	887	2,004	4,811	124,091	5.01	31	916	2,111	5,036	153,026	4.85
Group Quarters																			
Group	201	90	297	707	991	4,209	1.75	145	822	1,465	3,071	25,254	2.85	103	573	1,326	2,893	25,621	3.02
Non-Group	28,577	7	546	913	1,307	8,679	1.43	11	1,640	3,475	7,892	2,350,151	13.84	10	1,534	3,536	8,139	290,613	4.73

[1] This step used demographic variables from screener data for all selected person pairs; DU = dwelling unit, MSA = metropolitan statistical area, pr = pair, ps = poststratification, SDU = screener dwelling unit, sel = selected.

[2] Q1 and Q3 refer to the first and third quartile of the weight distribution.

[3] Unequal weighting effect (UWE) is defined as $1 + [(n - 1)/n]*CV^2$, where CV = coefficient of variation of weights.

Table L.2 2017 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.nr)

Domain	*n*	Before res.pr.nr[1] (SDUWT*PRWT12*PRWT13)						After res.pr.nr[1] (SDUWT*PRWT12*...*PRWT14)					
		Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
Total	17,704	10	1,411	3,235	7,391	290,613	4.73	10	1,939	4,755	11,595	763,301	6.35
Pair Age Group													
12–17, 12–17	2,371	10	652	1,578	2,973	25,669	2.23	10	835	2,048	3,963	28,951	2.19
12–17, 18–25	1,755	63	1,117	2,084	3,806	23,111	1.94	66	1,569	3,109	5,964	38,112	2.03
12–17, 26–34	910	38	1,116	2,246	4,078	81,803	2.98	51	1,430	3,151	6,012	99,704	2.98
12–17, 35–49	3,862	64	1,570	3,214	6,084	76,715	2.29	74	2,134	4,501	9,266	103,664	2.49
12–17, 50+	763	292	3,831	8,134	13,555	94,414	2.13	487	4,962	11,136	20,344	160,277	2.36
18–25, 18–25	2,524	56	933	1,927	4,002	22,096	2.11	60	1,280	2,868	6,524	40,938	2.27
18–25, 26–34	812	102	1,530	3,000	5,871	57,565	2.55	143	2,179	4,550	10,085	105,595	3.02
18–25, 35–49	1,206	245	2,739	5,357	9,701	58,442	2.11	277	4,040	8,421	16,624	135,071	2.22
18–25, 50+	670	797	6,429	11,475	18,881	77,417	1.81	861	10,005	20,453	38,171	330,689	2.12
26–34, 26–34	841	166	2,693	5,238	8,813	134,560	2.33	533	4,046	7,919	15,192	273,308	3.17
26–34, 35–49	447	111	3,395	6,491	11,178	172,536	3.44	133	5,106	9,608	18,570	236,189	3.68
26–34, 50+	235	1,045	11,702	21,307	33,474	217,403	2.01	1,965	20,658	40,007	72,618	315,206	1.97
35–49, 35–49	627	678	3,904	8,696	14,749	160,583	3.26	820	6,115	13,475	28,302	427,673	4.02
35–49, 50+	211	1,809	11,556	21,836	34,587	255,788	2.49	3,174	24,642	52,596	99,417	558,249	2.36
50+, 50+	470	3,420	19,709	36,679	51,816	290,613	1.82	4,931	36,782	76,613	112,246	763,301	1.93
Pair Race/Ethnicity													
Hispanic or Latino	3,083	22	1,706	3,648	8,015	290,613	4.67	51	2,329	5,438	12,749	627,261	6.65
Black or African American	1,918	27	1,775	3,775	7,733	261,922	5.10	27	2,220	4,959	10,854	504,622	6.88
White	9,585	18	1,337	3,017	7,109	285,564	4.60	26	1,865	4,489	11,161	763,301	6.38
Other	1,212	10	994	2,525	6,771	239,957	5.39	10	1,394	4,060	13,873	395,516	6.03
White & Black or African American	205	108	1,569	3,734	7,389	250,895	7.94	115	1,838	5,110	10,407	419,215	9.15
White & Hispanic or Latino	729	72	1,503	3,815	8,285	149,094	3.41	97	2,244	5,903	13,025	234,743	4.24
White & Other	630	104	1,306	3,138	7,488	142,364	4.32	108	1,733	4,308	10,492	209,719	4.67
Black or African American & Hispanic or Latino	88	302	3,332	6,682	11,674	72,395	2.48	441	4,027	10,765	22,482	212,056	3.31
Black or African American & Other	103	134	1,418	3,188	6,010	27,444	2.17	345	2,527	5,817	11,524	82,604	2.76
Hispanic or Latino & Other	151	56	921	2,303	6,159	115,741	5.72	60	1,450	3,356	8,348	179,409	6.28
Pair Gender													
Male, Male	3,609	18	1,337	3,243	6,959	136,394	3.46	26	2,007	5,079	11,613	353,737	4.48
Female, Female	4,056	28	1,355	3,071	6,870	250,895	4.24	51	1,797	4,321	10,286	627,261	6.47
Male, Female	10,039	10	1,449	3,289	7,832	290,613	5.04	10	1,978	4,839	12,212	763,301	6.53
Household Size													
Two	4,325	38	1,184	2,878	7,424	143,788	3.82	51	1,525	3,883	10,918	241,885	5.01
Three	5,674	10	1,321	2,808	5,706	290,613	5.41	10	1,838	4,048	9,059	705,026	7.44
Four or More	7,705	19	1,692	3,866	8,831	285,564	4.67	27	2,420	5,953	14,185	763,301	6.30

(continued)

Table L.2 2017 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.nr) (continued)

Domain	*n*	Before res.pr.nr[1] (SDUWT*PRWT12*PRWT13)						After res.pr.nr[1] (SDUWT*PRWT12*...*PRWT14)					
		Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
Census Region													
Northeast	3,346	10	1,032	2,557	6,384	255,788	5.03	10	1,343	3,688	10,334	558,249	6.71
South	5,938	18	1,916	4,052	8,589	290,613	4.47	26	2,622	5,817	13,050	763,301	6.32
Midwest	4,162	108	1,408	2,922	6,314	220,458	4.51	115	1,968	4,305	9,941	550,629	5.42
West	4,258	56	1,124	3,023	7,693	239,957	4.84	60	1,572	4,391	12,249	627,261	6.43
Quarter													
Quarter1	4,059	24	1,651	3,712	8,120	261,922	4.36	43	2,268	5,676	12,871	468,755	5.53
Quarter2	4,566	18	1,339	3,156	7,200	290,613	5.06	26	1,791	4,473	11,061	592,733	6.17
Quarter3	4,480	22	1,272	2,945	6,715	277,251	5.09	51	1,865	4,606	11,279	705,026	6.87
Quarter4	4,599	10	1,441	3,184	7,478	285,564	4.42	10	1,876	4,494	11,300	763,301	6.84
% Hispanic or Latino in Segment													
50–100%	1,664	27	1,888	4,208	9,489	290,613	4.56	27	2,506	6,114	14,935	627,261	6.49
10–<50%	4,435	22	1,859	4,176	8,837	220,458	3.78	51	2,633	6,396	15,150	525,656	4.86
<10%	11,605	10	1,250	2,827	6,523	285,564	5.19	10	1,707	4,049	10,032	763,301	7.00
% Black or African American in Segment													
50–100%	1,255	22	1,837	3,788	8,129	261,922	5.59	51	2,281	5,188	11,682	468,755	6.20
10–<50%	3,482	10	1,656	3,713	8,055	290,613	4.46	10	2,315	5,394	12,660	592,733	6.09
<10%	12,967	18	1,322	3,059	7,099	285,564	4.68	26	1,821	4,552	11,288	763,301	6.43
% Owner-Occupied DUs[1] in Segment													
50–100%	13,232	18	1,525	3,408	7,816	290,613	4.65	26	2,123	5,047	12,153	763,301	6.21
10–<50%	3,584	24	1,520	3,399	7,289	218,170	4.35	27	1,974	4,815	11,360	627,261	6.15
<10%	888	10	483	1,065	2,432	115,741	5.88	10	686	1,553	4,305	123,781	5.19
Combined Median Rent/Housing Value													
1st Quintile	3,052	27	1,146	2,697	5,900	239,957	4.75	27	1,506	3,687	8,722	471,021	6.90
2nd Quintile	4,204	27	1,283	2,911	6,811	261,922	5.10	27	1,651	3,813	9,883	627,261	7.38
3rd Quintile	4,156	10	1,465	3,444	7,610	290,613	4.75	10	2,009	4,987	11,592	592,733	6.71
4th Quintile	3,657	22	1,522	3,422	7,755	277,251	4.61	51	2,302	5,559	13,349	705,026	5.52
5th Quintile	2,635	18	1,772	4,010	8,948	285,564	4.19	26	2,632	6,641	16,578	763,301	5.16
Population Density													
Large MSA[1]	7,503	18	2,223	4,673	9,575	290,613	4.12	26	3,220	7,277	16,442	763,301	5.16
Medium to Small MSA[1]	8,751	10	1,085	2,471	5,678	277,251	5.22	10	1,463	3,410	8,526	705,026	7.61
Non-MSA,[1] Urban	539	88	975	2,198	4,812	62,177	3.81	95	1,236	3,011	7,486	201,597	5.55
Non-MSA,[1] Rural	911	31	919	2,207	5,194	90,937	4.33	39	1,182	2,843	7,172	214,986	5.85
Group Quarters													
Group	150	103	518	1,250	2,693	13,501	2.59	115	759	1,701	4,283	40,938	3.10
Non-Group	17,554	10	1,425	3,255	7,437	290,613	4.71	10	1,962	4,791	11,668	763,301	6.32

[1] This step used demographic variables from screener data for all selected person pairs; DU = dwelling unit, MSA = metropolitan statistical area, nr = nonresponse adjustment, pr = pair, res = respondent, SDU = screener dwelling unit.

[2] Q1 and Q3 refer to the first and third quartile of the weight distribution.

[3] Unequal weighting effect (UWE) is defined as $1 + [(n - 1)/n]*CV^2$, where CV = coefficient of variation of weights.

Table L.3 2017 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.ps and res.pr.ev)

Domain	*n*	Before res.pr.ps[1] (SDUWT*PRWT12*...*PRWT14)						After res.pr.ps[1] (SDUWT*PRWT12*...*PRWT15)						Final Weight: After res.pr.ev[1] (SDUWT*PRWT12*...*PRWT16)					
		Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
Total	17,704	10	1,939	4,755	11,595	763,301	6.35	17	1,836	4,660	11,540	511,846	6.25	17	1,830	4,659	11,589	509,895	6.22
Pair Age Group																			
12–17, 12–17	2,368	10	837	2,049	3,964	39,256	2.24	17	762	1,924	3,998	31,140	2.33	17	760	1,917	4,003	28,680	2.32
12–17, 18–25	1,758	66	1,569	3,127	5,986	38,112	2.03	53	1,399	3,142	6,156	34,440	2.01	51	1,385	3,142	6,170	28,178	1.98
12–17, 26–34	894	51	1,420	3,136	6,117	99,704	3.03	27	1,294	3,070	6,359	91,200	3.07	26	1,285	3,055	6,320	79,902	3.04
12–17, 35–49	3,870	74	2,132	4,520	9,273	103,664	2.47	57	2,038	4,441	9,540	88,168	2.41	56	2,031	4,436	9,536	89,956	2.41
12–17, 50+	771	358	4,913	11,125	20,543	160,277	2.37	246	4,310	10,076	20,720	149,411	2.43	243	4,284	10,038	20,698	149,980	2.44
18–25, 18–25	2,467	60	1,285	2,861	6,530	62,562	2.32	51	1,223	2,912	6,797	32,750	2.22	49	1,219	2,933	6,858	32,749	2.20
18–25, 26–34	839	143	2,098	4,382	9,808	105,595	3.05	142	1,778	4,138	10,169	99,022	3.13	143	1,794	4,119	10,170	88,180	3.01
18–25, 35–49	1,176	277	3,978	8,373	16,787	235,783	2.44	132	3,752	8,413	17,944	91,998	2.19	131	3,729	8,494	18,002	79,683	2.18
18–25, 50+	683	861	9,844	20,304	37,641	330,689	2.15	584	9,136	19,044	37,907	215,749	2.07	554	9,140	19,031	38,206	206,759	2.07
26–34, 26–34	871	533	4,046	7,947	15,726	273,308	3.13	258	3,357	6,623	13,528	289,980	3.62	248	3,299	6,572	13,633	300,425	3.67
26–34, 35–49	445	133	5,079	9,379	17,871	236,189	3.64	72	4,510	9,744	17,480	250,926	3.84	70	4,432	9,719	17,610	250,560	3.86
26–34, 50+	249	1,254	18,074	37,602	69,911	315,206	2.07	921	16,421	34,365	69,137	332,169	2.21	906	16,236	34,542	69,711	330,744	2.22
35–49, 35–49	628	820	6,048	13,163	26,938	504,622	4.35	527	5,891	11,911	26,055	494,639	4.70	518	5,957	12,015	26,102	488,199	4.72
35–49, 50+	213	3,147	23,527	52,642	105,364	558,249	2.29	2,214	21,567	50,988	105,505	508,111	2.34	2,170	21,070	50,535	105,551	509,895	2.32
50+, 50+	472	4,931	36,143	76,518	111,324	763,301	1.93	3,232	37,295	77,919	114,117	511,846	1.74	3,210	37,435	77,714	114,500	498,139	1.72
Pair Race/Ethnicity																			
Hispanic or Latino	3,116	51	2,333	5,459	12,701	627,261	6.64	49	2,235	5,400	12,631	508,503	6.59	51	2,237	5,375	12,733	488,199	6.59
Black or African American	1,881	27	2,220	4,963	10,999	504,622	6.57	18	2,115	4,941	11,081	486,945	6.93	18	2,114	4,952	11,037	479,562	6.92
White	9,247	10	1,847	4,466	11,157	763,301	6.31	17	1,872	4,617	11,528	511,846	5.97	17	1,871	4,611	11,560	509,895	5.92
Other	1,183	51	1,477	4,286	14,140	395,516	5.74	27	1,318	4,310	14,123	359,517	5.83	26	1,303	4,288	14,133	361,531	5.76
White & Black or African American	189	183	1,874	4,579	9,688	419,215	9.28	111	1,928	4,420	10,426	452,574	9.62	111	1,964	4,508	10,291	453,463	9.54
White & Hispanic or Latino	765	77	2,136	5,270	12,822	238,330	4.85	50	1,737	4,708	11,989	247,069	5.08	51	1,734	4,642	12,012	251,101	5.13
White & Other	815	118	1,752	4,555	10,925	471,021	6.21	46	1,265	3,222	8,166	234,346	5.43	45	1,250	3,210	8,259	232,815	5.46
Black or African American & Hispanic or Latino	120	60	2,761	7,517	16,393	212,056	3.81	61	2,642	7,122	16,403	166,153	3.36	61	2,642	7,049	16,210	166,489	3.36
Black or African American & Other	204	101	2,241	5,450	11,421	391,046	7.31	23	722	1,791	4,238	110,558	6.46	22	712	1,792	4,253	109,521	6.43
Hispanic or Latino & Other	184	134	1,647	3,812	9,544	330,689	7.98	69	1,260	3,005	7,939	153,696	5.38	68	1,261	3,027	7,861	156,678	5.39

(continued)

Table L.3 2017 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.ps and res.pr.ev) (continued)

Domain	*n*	Before res.pr.ps[1] (SDUWT*PRWT12*...*PRWT14)						After res.pr.ps[1] (SDUWT*PRWT12*...*PRWT15)						Final Weight: After res.pr.ev[1] (SDUWT*PRWT12*...*PRWT16)					
		Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
Pair Gender																			
Male, Male	3,609	26	2,009	5,065	11,596	353,737	4.47	30	1,905	4,936	11,338	291,850	4.66	30	1,898	4,950	11,460	273,947	4.61
Female, Female	4,054	51	1,798	4,321	10,323	627,261	6.47	23	1,764	4,216	10,401	491,970	6.16	22	1,760	4,221	10,375	475,394	6.15
Male, Female	10,041	10	1,975	4,839	12,212	763,301	6.53	17	1,851	4,749	12,055	511,846	6.42	17	1,846	4,749	12,097	509,895	6.39
Household Size																			
Two	4,325	51	1,525	3,883	10,918	241,885	5.01	27	1,402	3,727	10,538	248,605	5.29	26	1,395	3,722	10,547	246,813	5.30
Three	5,674	10	1,838	4,048	9,059	705,026	7.44	17	1,742	4,043	9,220	508,503	6.78	17	1,746	4,030	9,271	487,227	6.65
Four or More	7,705	27	2,420	5,953	14,185	763,301	6.30	23	2,262	5,914	14,058	511,846	6.24	22	2,257	5,869	14,113	509,895	6.23
Census Region																			
Northeast	3,346	10	1,343	3,688	10,334	558,249	6.71	17	1,201	3,695	10,351	505,150	6.42	17	1,194	3,647	10,418	488,637	6.40
South	5,938	26	2,622	5,817	13,050	763,301	6.32	23	2,410	5,558	13,007	511,846	6.12	22	2,393	5,537	13,076	509,895	6.09
Midwest	4,162	115	1,968	4,305	9,941	550,629	5.42	97	1,966	4,394	10,169	466,939	5.42	95	1,962	4,412	10,096	469,229	5.43
West	4,258	60	1,572	4,391	12,249	627,261	6.43	57	1,567	4,352	12,107	502,020	6.54	56	1,568	4,331	12,127	475,394	6.49
Quarter																			
Quarter1	4,059	43	2,268	5,676	12,871	468,755	5.53	44	2,150	5,469	13,087	502,020	5.87	43	2,137	5,423	13,198	488,199	5.84
Quarter2	4,566	26	1,791	4,473	11,061	592,733	6.17	18	1,664	4,383	10,930	508,503	6.39	18	1,650	4,387	10,948	484,385	6.36
Quarter3	4,480	51	1,865	4,606	11,279	705,026	6.87	23	1,782	4,604	11,337	508,111	6.51	22	1,778	4,599	11,376	509,895	6.49
Quarter4	4,599	10	1,876	4,494	11,300	763,301	6.84	17	1,811	4,339	11,054	511,846	6.19	17	1,802	4,326	11,023	498,139	6.16
% Hispanic or Latino in Segment																			
50–100%	1,664	27	2,506	6,114	14,935	627,261	6.49	18	2,396	5,984	14,807	508,503	6.34	18	2,409	5,943	14,850	484,385	6.29
10–<50%	4,435	51	2,633	6,396	15,150	525,656	4.86	27	2,484	6,205	15,126	508,111	5.04	26	2,476	6,231	15,163	509,895	5.02
<10%	11,605	10	1,707	4,049	10,032	763,301	7.00	17	1,603	4,000	10,102	511,846	6.74	17	1,597	3,987	10,038	498,139	6.72
% Black or African American in Segment																			
50–100%	1,255	51	2,281	5,188	11,682	468,755	6.20	43	2,073	4,814	11,483	486,945	6.60	42	2,056	4,862	11,655	479,562	6.61
10–<50%	3,482	10	2,315	5,394	12,660	592,733	6.09	17	2,045	5,103	12,478	508,503	6.31	17	2,030	5,146	12,482	509,895	6.26
<10%	12,967	26	1,821	4,552	11,288	763,301	6.43	30	1,748	4,502	11,296	511,846	6.19	30	1,747	4,495	11,289	498,139	6.17
% Owner-Occupied DUs[1] in Segment																			
50–100%	13,232	26	2,123	5,047	12,153	763,301	6.21	30	2,013	4,941	12,032	511,846	6.06	30	2,005	4,957	12,033	509,895	6.04
10–<50%	3,584	27	1,974	4,815	11,360	627,261	6.15	18	1,843	4,687	11,532	491,970	6.26	18	1,841	4,712	11,621	475,394	6.19
<10%	888	10	686	1,553	4,305	123,781	5.19	17	593	1,454	4,003	136,147	5.89	17	588	1,447	3,952	139,270	5.97
Combined Median Rent/Housing Value																			
1st Quintile	3,052	27	1,506	3,687	8,722	471,021	6.90	23	1,413	3,685	8,623	359,517	6.41	22	1,398	3,699	8,652	361,531	6.41
2nd Quintile	4,204	27	1,651	3,813	9,883	627,261	7.38	18	1,553	3,774	9,704	491,970	7.44	18	1,538	3,748	9,699	479,562	7.38
3rd Quintile	4,156	10	2,009	4,987	11,592	592,733	6.71	17	1,921	4,824	11,586	508,503	6.65	17	1,919	4,846	11,597	509,895	6.61
4th Quintile	3,657	51	2,302	5,559	13,349	705,026	5.52	44	2,165	5,405	13,754	507,991	5.43	43	2,143	5,365	13,771	487,227	5.43
5th Quintile	2,635	26	2,632	6,641	16,578	763,301	5.16	27	2,597	6,562	15,808	511,846	5.08	26	2,586	6,523	15,890	498,139	5.04

(continued)

Table L.3 2017 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.ps and res.pr.ev) (continued)

Domain	*n*	Before res.pr.ps[1] (SDUWT*PRWT12*...*PRWT14)						After res.pr.ps[1] (SDUWT*PRWT12*...*PRWT15)						Final Weight: After res.pr.ev[1] (SDUWT*PRWT12*...*PRWT16)					
		Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
Population Density																			
Large MSA[1]	7,503	26	3,220	7,277	16,442	763,301	5.16	18	3,078	7,195	16,247	511,846	5.16	18	3,070	7,181	16,193	498,139	5.15
Medium to Small MSA[1]	8,751	10	1,463	3,410	8,526	705,026	7.61	17	1,398	3,380	8,576	508,111	7.30	17	1,391	3,372	8,587	509,895	7.24
Non-MSA,[1] Urban	539	95	1,236	3,011	7,486	201,597	5.55	85	1,238	2,981	7,342	166,023	5.68	84	1,226	3,000	7,335	172,395	5.74
Non-MSA,[1] Rural	911	39	1,182	2,843	7,172	214,986	5.85	37	1,092	2,818	7,550	171,880	5.45	34	1,087	2,795	7,553	169,626	5.47
Group Quarters																			
Group	150	115	759	1,701	4,283	40,938	3.10	51	778	1,814	4,092	24,664	2.74	49	768	1,846	4,055	23,770	2.75
Non-Group	17,554	10	1,962	4,791	11,668	763,301	6.32	17	1,859	4,694	11,602	511,846	6.22	17	1,854	4,701	11,623	509,895	6.19
Pair Relationship Domain[4]																			
Parent-Child (12–14)	2,810	51	1,994	4,132	8,797	103,889	2.71	27	2,142	4,544	9,824	104,977	2.66	26	2,130	4,520	9,824	105,448	2.65
Parent-Child (12–17)	5,049	51	2,115	4,613	9,955	144,303	2.87	27	2,151	4,752	10,388	149,411	2.84	26	2,132	4,732	10,399	149,980	2.85
Parent-Child (12–20)	5,901	51	2,291	5,131	11,192	186,849	3.02	27	2,338	5,238	11,697	215,749	3.02	26	2,319	5,245	11,702	206,759	3.01
Sibling (12–14)-Sibling (15–17)	1,397	10	810	2,056	3,864	24,806	2.10	17	804	2,115	4,165	31,140	2.23	17	809	2,116	4,178	28,680	2.19
Sibling (12–17)-Sibling (18–25)	1,583	66	1,569	3,170	6,068	38,112	2.02	53	1,508	3,281	6,358	34,440	1.98	51	1,489	3,263	6,386	28,178	1.95
Spouse-Spouse/ Partner-Partner	3,467	60	2,055	5,948	16,970	763,301	6.08	58	1,756	5,383	15,731	511,846	6.08	56	1,747	5,349	15,677	509,895	6.04
Spouse-Spouse/ Partner-Partner with Children (Younger Than 18)	1,711	60	2,098	5,554	12,602	763,301	8.86	61	2,181	5,894	13,742	511,846	8.31	61	2,216	5,896	13,718	509,895	8.24

[1] This step used demographic variables from questionnaire data for all selected person pairs; DU = dwelling unit, ev = extreme value adjustment, MSA = metropolitan statistical area, pr = pair, ps = poststratification adjustment, res = respondent, SDU = screener dwelling unit.
[2] Q1 and Q3 refer to the first and third quartile of the weight distribution.
[3] Unequal weighting effect (UWE) is defined as $1 + [(n - 1)/n]*CV^2$, where CV = coefficient of variation of weights.
[4] Parent-child (15–17) was not included here since extreme values were not controlled with this domain.