2016 NATIONAL SURVEY ON DRUG USE AND HEALTH

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 2018

This page intentionally left blank

2016 NATIONAL SURVEY ON DRUG USE AND HEALTH: QUESTIONNAIRE DWELLING UNIT-LEVEL AND PERSON PAIR-LEVEL SAMPLING WEIGHT CALIBRATION

Prepared for the 2016 Methodological Resource Book (Section 12)

Contract No. HHSS283201300001C
RTI Project No. 0213985.006.106.006
Deliverable No. 49

RTI Authors:

Matthew Westlake
Jeff Laufenberg
Patrick Chen
Harper Gordek
Ralph Folsom*

* Retired from RTI

SAMHSA Authors:

Matthew Williams
Art Hughes**

** Retired from SAMHSA

RTI Project Director:

David Hunter

SAMHSA Project Officer:

Peter Tice

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

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

Prepared by RTI International, Research Triangle Park, North Carolina

March 2018

Recommended Citation: Center for Behavioral Health Statistics and Quality. (2018). *2016 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, Rockville, MD.

Acknowledgments

This report was prepared for the Substance Abuse and Mental Health Services Administration (SAMHSA), Center for Behavioral Health Statistics and Quality, by RTI International (a registered trademark and a trade name of Research Triangle Institute). Contributors to this report at RTI include Debbie Bond, Claudia Clark, and Margaret Johnson.

Table of Contents

Chapter **Page**

Table of Contents (continued)

List of Tables

Table **Page**

List of Tables (continued)

List of Exhibits

Exhibit **Page**

List of Exhibits (continued)

Exhibit **Page**

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, 2018a) 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 (2018a) 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 2016. 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, 2018b), 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 facilitates 50 percent overlap in third-stage units (area segments) within each successive 2-year period from 2014 through 2017. This designed sample overlap slightly increases 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 reduces segment listing costs because only one-half of the segments will need 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 2016 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, is 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 (2012–2016). 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 2012–2016 NSDUH Sample Sizes

Sample Unit		2012	2013	2014	2015	2016
SDU	Selected	178,586	190,067	154,533	165,328	173,149
	Completed	153,873	160,325	127,605	132,210	135,188
QDU	Selected	60,621	61,634	64,796	66,721	67,574
	Completed	48,850	48,896	49,672	50,119	50,095
Pair	Selected	27,035	27,108	26,844	27,778	28,033
	Completed	19,459	18,942	18,229	17,954	17,847
Person	Selected	87,656	88,742	91,640	94,499	95,607
	Completed	68,309	67,838	67,901	68,073	67,942

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], 2018a). 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

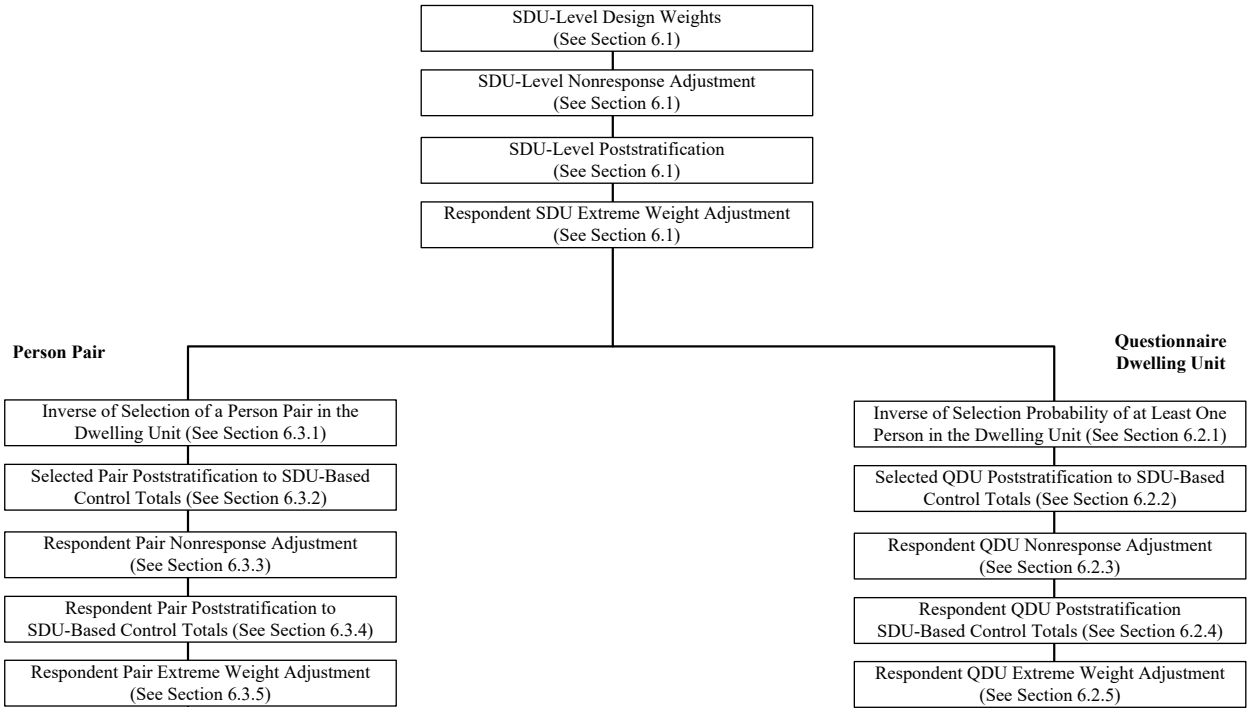
Some of the multiplicities, including counts of all possible pairs in a household for a given domain, were used for poststratification. Details are provided in Chapter 4. Additional information on the imputation of pair relationships, multiplicity factors, and household-level person counts for poststratification can be found in the NSDUH Methodological Resource Book editing and imputation report (CBHSQ, 2018a). 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, 2017b).

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, we performed a multivariate calibration with respect to a key set of pair domains. 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 2016 survey was the same as the one used in the 1999–2015 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 2002–2015 National Surveys on Drug Use and Health (NSDUHs),[6] the 2016 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, 2017a).

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 2016, 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 2016 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 2012–2016 NSDUHs

	2012			2013			2014			2015			2016		
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	178,586	153,873	86.16	190,067	160,325	84.35	154,533	127,605	82.57	165,328	132,210	79.97	173,149	135,188	78.08
QDUs															
Total QDUs	60,621	48,850	80.58	61,634	48,896	79.33	64,796	49,672	76.66	66,721	50,119	75.12	67,574	50,095	74.13
People															
Total People	87,656	68,309	77.93	88,742	67,838	76.44	91,640	67,901	74.10	94,499	68,073	72.04	95,607	67,942	71.06
12-17	27,147	22,473	82.78	27,630	22,494	81.41	21,392	17,007	79.50	21,859	16,911	77.36	22,323	17,081	76.52
18-25	28,639	22,529	78.67	28,921	22,214	76.81	21,726	16,449	75.71	23,211	17,097	73.66	22,836	16,435	71.97
26-34	8,304	6,484	78.08	8,210	6,310	76.86	14,004	10,252	73.21	14,720	10,446	70.96	15,022	10,528	70.08
35-49	12,364	9,076	73.41	12,566	9,058	72.08	19,065	13,590	71.28	19,341	13,304	68.79	19,988	13,572	67.90
50+	11,202	7,747	69.16	11,415	7,762	68.00	15,453	10,603	68.61	15,368	10,315	67.12	15,438	10,326	66.89
Non-Pairs[4]															
Total Non-Pairs	126,838	29,391	N/A	133,217	29,954	N/A	90,443	31,443	N/A	104,432	32,165	N/A	107,155	32,248	N/A
0,0	93,252	N/A	N/A	98,691	N/A	N/A	62,809	N/A	N/A	65,489	N/A	N/A	67,614	N/A	N/A
Total Singletons	33,586	29,391	87.51	34,526	29,954	86.76	37,952	31,443	82.85	38,943	32,165	82.60	39,541	32,248	81.56
0, 12-17	9,017	8,277	91.79	9,420	8,574	91.02	4,850	4,704	96.99	5,244	5,014	95.61	5,144	4,997	97.14
0, 18-25	10,325	9,461	91.63	10,535	9,475	89.94	7,250	6,647	91.68	7,583	7,102	93.66	7,647	6,895	90.17
0, 26-34	3,856	3,327	86.28	3,914	3,367	86.02	7,460	6,034	80.88	7,726	6,166	79.81	8,045	6,270	77.94
0, 35-49	4,368	3,645	83.45	4,506	3,736	82.91	8,074	6,450	79.89	8,093	6,320	78.09	8,442	6,596	78.13
0, 50+	6,020	4,681	77.76	6,151	4,802	78.07	10,318	7,608	73.74	10,297	7,563	73.45	10,263	7,490	72.98
Pairs															
Total Pairs[5]	27,035	19,459	71.98	27,108	18,942	69.88	26,844	18,229	67.91	27,778	17,954	64.63	28,033	17,847	63.66
12-17, 12-17	4,507	3,668	81.38	4,535	3,609	79.58	3,070	2,407	78.40	2,962	2,253	76.06	3,199	2,386	74.59
12-17, 18-25	3,627	2,759	76.07	3,662	2,754	75.20	2,443	1,832	74.99	2,571	1,795	69.82	2,548	1,774	69.62
12-17, 26-34	825	658	79.76	811	621	76.57	1,297	941	72.55	1,299	939	72.29	1,281	883	68.93
12-17, 35-49	3,813	2,812	73.75	3,834	2,756	71.88	5,530	3,940	71.25	5,654	3,888	68.77	5,829	3,930	67.42
12-17, 50+	851	631	74.15	833	571	68.55	1,132	776	68.55	1,167	769	65.90	1,123	725	64.56
18-25, 18-25	5,476	3,901	71.24	5,478	3,795	69.28	3,743	2,585	69.06	4,043	2,654	65.64	3,958	2,512	63.47
18-25, 26-34	1,079	794	73.59	1,034	690	66.73	1,378	870	63.13	1,577	975	61.83	1,429	886	62.00
18-25, 35-49	1,582	1,053	66.56	1,561	990	63.42	1,906	1,180	61.91	2,092	1,186	56.69	2,013	1,134	56.33
18-25, 50+	1,074	660	61.45	1,173	715	60.95	1,263	750	59.38	1,302	731	56.14	1,283	722	56.27
26-34, 26-34	880	604	68.64	822	581	70.68	1,356	865	63.79	1,492	870	58.31	1,518	905	59.62
26-34, 35-49	469	320	68.23	489	307	62.78	737	442	59.97	716	408	56.98	788	445	56.47
26-34, 50+	315	177	56.19	318	163	51.26	420	235	55.95	418	218	52.15	443	234	52.82
35-49, 35-49	833	487	58.46	857	512	59.74	1,160	658	56.72	1,158	635	54.84	1,213	627	51.69
35-49, 50+	466	272	58.37	462	245	53.03	498	262	52.61	470	232	49.36	490	213	43.47
50+, 50+	1,238	663	53.55	1,239	633	51.09	911	486	53.35	857	401	46.79	918	471	51.31

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, 2018b).

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, 2018b).

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 2016 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 2016 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 2016 survey. This DU gave rise to a selected pair of respondents, one aged 61 and the other aged 91. The selection probability in this DU was 0.15539 for a respondent aged 50 or older. Using the formula (2.6) in Chapter 2, the pair selection probability was computed to be 0.000664719. Therefore, the inverse of the selection probability, the pair-level design weight, was 1,504.40. 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 1999–2015 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 2016 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 modify 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 blocks 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 2016 NSDUH Methodological Resource Book (MRB) sample design report (Center for Behavioral Health Statistics and Quality [CBHSQ], 2017a).

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 2016 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 2016 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 2012–2016 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 2002–2016 survey pair data, the built-in ev control feature was used for each adjustment step.

Exhibit 6.1 Summary of 2016 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
#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, 2018b). 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 2016 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
#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, 2018b). 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	2012		2013		2014		2015		2016	
	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs
QDU										
Northeast	12,616	9,917	12,791	9,954	12,950	9,664	13,519	9,777	13,414	9,552
South	18,345	15,019	18,766	15,073	21,448	16,680	21,887	16,708	22,287	16,810
Midwest	16,984	13,687	17,207	13,519	15,276	11,618	15,808	11,698	16,025	11,768
West	12,676	10,227	12,870	10,350	15,122	11,710	15,507	11,936	15,848	11,965
Total	60,621	48,850	61,634	48,896	64,796	49,672	66,721	50,119	67,574	50,095
Model Group	**2012**		**2013**		**2014**		**2015**		**2016**	
	Selected Pairs	Completed Pairs	Selected Pairs	Completed Pairs	Selected Pairs	Completed Pairs	Selected Pairs	Completed Pairs	Selected Pairs	Completed Pairs
Pair										
Northeast + South	13,619	9,723	13,535	9,416	13,969	9,436	14,502	9,309	14,543	9,182
Midwest + West	13,416	9,736	13,573	9,526	12,875	8,793	13,276	8,645	13,490	8,665
Total	27,035	19,459	27,108	18,942	26,844	18,229	27,778	17,954	28,033	17,847

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 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 2016, 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 2016 NSDUH MRB sample design report (CBHSQ, 2017a), and for more detail on Weight Components #4 and #9 through #11, see the 2016 NSDUH MRB person-level sampling weight calibration report (CBHSQ, 2018b).

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 (2017a).

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 (2017a).

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 (2018b).

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, etc., 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, etc., 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, 2018a) 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 2016 National Survey on Drug Use and Health for various national domains. This table also shows the weighted response rates. Most domains reflect the overall 71.67 percent response rate, with most rates between 70 and 75 percent, although the highest response rate is 85.82 percent, from the Group category of the Group Quarters variable. The lowest response rate came from the 12-17 category of the Household Type variable, with 58.37 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 2016 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 57.33 percent. The response rates range from a low of 40.60 percent in the Pair Age Group category of 35-49, 50+ to a high of 76.26 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 2016 data, this step has been implemented for pair-level weighting to reduce final 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, 2018b). 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 2016 QDU domain household size of four or more, indicates a sample size of 10,805 with a total design-based weight of 15,111,787 and a census total of 14,807,696 with an initial slippage rate of 2.05 percent, which would imply a common weight adjustment approximately equal to 0.979877, 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 883, a design-based weight of 4,623,212, and a census total of 4,783,252, showing an initial slippage of -3.35 percent. The resultant required adjustment would be approximately equal to 1.034617, 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: 2016

Domain	*n*	Baseline (B)[1]	Final (F)[2]	(B-F)/F% (Estimate)	(B-F)/F% (SE)
Households with Family Income					
$0 - < $10,000	4,168	9,298,518 (283,518)	9,292,546 (283,116)	0.06	0.14
$10,000 - < $20,000	5,773	15,198,099 (354,978)	15,183,066 (354,358)	0.10	0.17
$20,000 - < $30,000	5,419	13,725,151 (309,103)	13,729,497 (309,820)	-0.03	-0.23
$30,000 - < $40,000	5,035	12,260,905 (278,738)	12,259,913 (279,733)	0.01	-0.36
$40,000 - < $50,000	5,037	12,463,008 (276,909)	12,466,385 (276,982)	-0.03	-0.03
$50,000 - < $75,000	7,807	19,155,322 (351,825)	19,163,945 (352,117)	-0.04	-0.08
$75,000 - < $100,000	5,899	14,443,230 (294,668)	14,446,430 (294,882)	-0.02	-0.07
$100,000+	10,957	26,944,722 (525,404)	26,947,173 (525,438)	-0.01	-0.01
Households with Number of Youths (< 18)					
0	23,370	81,644,655 (972,580)	81,648,205 (972,755)	-0.00	-0.02
1	10,689	17,600,029 (235,405)	17,603,629 (235,773)	-0.02	-0.16
2	9,360	14,924,452 (213,636)	14,921,141 (213,614)	0.02	0.01
3	4,427	6,418,785 (133,313)	6,411,240 (133,571)	0.12	-0.19
4+	2,249	2,901,034 (81,948)	2,904,740 (82,522)	-0.13	-0.70
Households with Insurance Coverage					
Yes	45,329	113,111,943 (1,014,939)	113,120,623 (1,013,996)	-0.01	0.09
No	4,766	10,377,012 (230,857)	10,368,332 (230,806)	0.08	0.02
Households with Number of Older Adults (65+)					
0	42,735	89,738,886 (754,369)	89,755,460 (754,315)	-0.02	0.01
1	4,785	21,819,510 (528,203)	21,800,165 (526,797)	0.09	0.27
2	2,517	11,729,767 (340,184)	11,730,840 (340,278)	-0.01	-0.03
3+	58	200,791 (34,882)	202,490 (35,208)	-0.84	-0.93

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: 2016

Drug	Mother User	*N*	Baseline[1]	Final[2]
Alcohol				
Lifetime	Yes	2,723	26.94 (1.40)	27.06 (1.39)
	No	434	14.83 (2.80)	14.64 (2.82)
	Overall	3,157	25.00 (1.22)	25.09 (1.22)
Past Year	Yes	2,221	21.49 (1.33)	21.52 (1.33)
	No	936	14.26 (2.06)	14.43 (2.06)
	Overall	3,157	19.20 (1.07)	19.29 (1.07)
Past Month	Yes	1,655	10.69 (1.21)	10.67 (1.21)
	No	1,502	7.07 (1.13)	7.17 (1.13)
	Overall	3,157	8.93 (0.85)	8.97 (0.85)
Cigarettes				
Lifetime	Yes	1,953	12.62 (1.14)	12.71 (1.16)
	No	1,204	6.03 (0.95)	6.28 (1.01)
	Overall	3,157	9.85 (0.79)	10.03 (0.81)
Past Year	Yes	794	13.05 (2.04)	13.23 (2.08)
	No	2,363	4.49 (0.63)	4.53 (0.63)
	Overall	3,157	6.25 (0.67)	6.35 (0.68)
Past Month	Yes	705	7.33 (1.60)	7.53 (1.67)
	No	2,452	1.90 (0.40)	1.89 (0.40)
	Overall	3,157	2.88 (0.46)	2.92 (0.47)

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: 2016

Drug	Father User	*N*	Baseline[1]	Final[2]
Alcohol				
Lifetime	Yes	1,723	24.20 (1.63)	24.26 (1.64)
	No	168	14.68 (4.12)	14.58 (4.06)
	Overall	1,891	23.14 (1.54)	23.20 (1.55)
Past Year	Yes	1,465	19.56 (1.54)	19.54 (1.54)
	No	426	14.20 (2.46)	14.27 (2.49)
	Overall	1,891	18.15 (1.32)	18.16 (1.32)
Past Month	Yes	1,223	8.59 (1.09)	8.60 (1.10)
	No	668	5.65 (1.20)	5.65 (1.21)
	Overall	1,891	7.48 (0.84)	7.49 (0.84)
Cigarettes				
Lifetime	Yes	1,288	11.77 (1.33)	11.75 (1.32)
	No	603	8.17 (1.75)	8.13 (1.77)
	Overall	1,891	10.45 (1.11)	10.43 (1.11)
Past Year	Yes	459	10.44 (2.27)	10.55 (2.27)
	No	1,432	5.14 (0.76)	5.19 (0.77)
	Overall	1,891	6.22 (0.78)	6.29 (0.80)
Past Month	Yes	387	4.68 (1.08)	4.97 (1.21)
	No	1,504	2.44 (0.51)	2.44 (0.51)
	Overall	1,891	2.84 (0.46)	2.90 (0.47)

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: 2016

Drug	Mother User	*N*	Baseline[1]	Final[2]
Any Illicit				
Lifetime	Yes	1,586	26.37 (1.71)	26.54 (1.72)
	No	1,571	16.92 (1.63)	16.96 (1.64)
	Overall	3,157	21.50 (1.21)	21.64 (1.22)
Past Year	Yes	416	20.41 (2.98)	20.48 (3.02)
	No	2,741	14.37 (1.03)	14.47 (1.05)
	Overall	3,157	15.07 (0.99)	15.18 (1.01)
Past Month	Yes	233	15.64 (3.64)	15.45 (3.61)
	No	2,924	7.02 (0.72)	7.06 (0.74)
	Overall	3,157	7.59 (0.72)	7.63 (0.73)
Marijuana				
Lifetime	Yes	1,455	19.64 (1.59)	19.92 (1.60)
	No	1,702	9.36 (1.27)	9.36 (1.28)
	Overall	3,157	13.83 (1.01)	14.00 (1.02)
Past Year	Yes	280	22.30 (3.89)	22.39 (3.91)
	No	2,877	10.80 (0.91)	10.91 (0.93)
	Overall	3,157	11.58 (0.89)	11.72 (0.91)
Past Month	Yes	176	20.30 (4.96)	20.10 (4.92)
	No	2,981	5.91 (0.68)	5.99 (0.70)
	Overall	3,157	6.53 (0.68)	6.63 (0.69)

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: 2016

Drug	Father User	*n*	Baseline[1]	Final[2]
Any Illicit				
Lifetime	Yes	1,121	25.66 (1.97)	25.72 (1.98)
	No	770	12.29 (1.86)	12.26 (1.89)
	Overall	1,891	19.74 (1.33)	19.80 (1.34)
Past Year	Yes	278	24.78 (3.97)	24.80 (4.01)
	No	1,613	13.19 (1.29)	13.15 (1.29)
	Overall	1,891	14.60 (1.23)	14.59 (1.23)
Past Month	Yes	168	13.74 (3.56)	13.48 (3.47)
	No	1,723	6.45 (0.88)	6.45 (0.89)
	Overall	1,891	6.97 (0.86)	6.96 (0.86)
Marijuana				
Lifetime	Yes	1,029	18.62 (1.88)	18.74 (1.89)
	No	862	6.56 (1.39)	6.47 (1.38)
	Overall	1,891	12.62 (1.15)	12.68 (1.16)
Past Year	Yes	215	27.09 (4.98)	27.14 (5.05)
	No	1,676	9.65 (1.08)	9.63 (1.08)
	Overall	1,891	11.07 (1.09)	11.10 (1.09)
Past Month	Yes	148	14.78 (4.07)	14.46 (3.96)
	No	1,743	5.45 (0.82)	5.45 (0.83)
	Overall	1,891	6.00 (0.81)	6.00 (0.81)

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: 2016

Drug	Parent Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Alcohol				
Lifetime	Yes	2,868	24.36 (1.31)	24.53 (1.33)
	No	2,080	27.09 (1.55)	27.09 (1.55)
	Overall	4,948	25.53 (1.04)	25.62 (1.04)
Past Year	Yes	2,868	19.30 (1.16)	19.46 (1.18)
	No	2,080	20.19 (1.40)	20.15 (1.40)
	Overall	4,948	19.68 (0.92)	19.76 (0.92)
Past Month	Yes	2,868	9.73 (0.95)	9.81 (0.96)
	No	2,080	6.97 (0.79)	6.96 (0.79)
	Overall	4,948	8.55 (0.66)	8.59 (0.67)
Cigarettes				
Lifetime	Yes	2,868	8.94 (0.86)	9.06 (0.88)
	No	2,080	12.25 (1.12)	12.42 (1.15)
	Overall	4,948	10.36 (0.70)	10.49 (0.72)
Past Year	Yes	2,868	5.71 (0.70)	5.81 (0.73)
	No	2,080	7.68 (0.97)	7.79 (1.00)
	Overall	4,948	6.55 (0.58)	6.66 (0.60)
Past Month	Yes	2,868	3.14 (0.56)	3.22 (0.59)
	No	2,080	3.28 (0.70)	3.32 (0.73)
	Overall	4,948	3.20 (0.44)	3.26 (0.46)

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: 2016

Drug	Parent Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Any Illicit				
Lifetime	Yes	2,868	19.76 (1.31)	19.95 (1.32)
	No	2,080	24.15 (1.46)	24.17 (1.48)
	Overall	4,948	21.64 (0.97)	21.75 (0.98)
Past Year	Yes	2,868	13.95 (1.01)	14.12 (1.03)
	No	2,080	17.38 (1.31)	17.36 (1.32)
	Overall	4,948	15.42 (0.82)	15.50 (0.83)
Past Month	Yes	2,868	6.71 (0.78)	6.86 (0.81)
	No	2,080	9.01 (0.93)	8.86 (0.91)
	Overall	4,948	7.70 (0.60)	7.71 (0.61)
Marijuana				
Lifetime	Yes	2,868	12.80 (1.13)	12.93 (1.14)
	No	2,080	15.66 (1.24)	15.77 (1.26)
	Overall	4,948	14.02 (0.83)	14.14 (0.84)
Past Year	Yes	2,868	10.73 (0.91)	10.90 (0.93)
	No	2,080	13.27 (1.17)	13.32 (1.19)
	Overall	4,948	11.82 (0.72)	11.93 (0.73)
Past Month	Yes	2,868	5.98 (0.75)	6.13 (0.78)
	No	2,080	7.55 (0.88)	7.47 (0.86)
	Overall	4,948	6.65 (0.58)	6.70 (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: 2016

Drug	Mother Talked about Dangers with Child	*N*	Baseline[1]	Final[2]
Alcohol				
Lifetime	0 times	241	10.10 (2.52)	10.06 (2.57)
	1-2 times	500	22.92 (3.30)	23.03 (3.32)
	A few times	793	21.80 (2.12)	22.00 (2.16)
	Many times	1,473	30.81 (1.94)	30.78 (1.92)
	Overall	3,007	25.12 (1.24)	25.19 (1.25)
Past Year	0 times	241	8.07 (2.31)	8.05 (2.36)
	1-2 times	500	15.40 (2.94)	15.57 (2.97)
	A few times	793	15.42 (1.73)	15.32 (1.72)
	Many times	1,473	25.12 (1.87)	25.26 (1.87)
	Overall	3,007	19.26 (1.09)	19.34 (1.10)
Past Month	0 times	241	3.12 (1.54)	3.09 (1.54)
	1-2 times	500	4.78 (1.32)	4.77 (1.32)
	A few times	793	5.97 (1.08)	5.97 (1.08)
	Many times	1,473	13.33 (1.59)	13.42 (1.58)
	Overall	3,007	8.93 (0.86)	8.98 (0.86)
Cigarettes				
Lifetime	0 times	241	3.61 (1.14)	3.66 (1.15)
	1-2 times	500	7.02 (1.79)	7.21 (1.84)
	A few times	793	7.01 (1.38)	7.10 (1.37)
	Many times	1,473	13.69 (1.34)	13.87 (1.37)
	Overall	3,007	9.82 (0.81)	9.97 (0.82)
Past Year	0 times	241	1.02 (0.53)	1.02 (0.52)
	1-2 times	500	4.77 (1.69)	4.79 (1.73)
	A few times	793	3.68 (0.94)	3.69 (0.93)
	Many times	1,473	9.11 (1.15)	9.22 (1.17)
	Overall	3,007	6.16 (0.68)	6.23 (0.69)
Past Month	0 times	241	0.36 (0.26)	0.37 (0.26)
	1-2 times	500	2.36 (1.18)	2.46 (1.30)
	A few times	793	1.54 (0.58)	1.51 (0.56)
	Many times	1,473	4.34 (0.84)	4.37 (0.86)
	Overall	3,007	2.88 (0.47)	2.91 (0.49)

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: 2016

Drug	Father Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Alcohol				
Lifetime	0 times	199	20.26 (4.62)	20.25 (4.58)
	1-2 times	390	17.11 (2.54)	17.25 (2.55)
	A few times	583	25.90 (2.62)	26.19 (2.67)
	Many times	572	24.68 (2.95)	24.62 (2.97)
	Overall	1,744	22.74 (1.55)	22.84 (1.56)
Past Year	0 times	199	14.40 (3.78)	14.22 (3.67)
	1-2 times	390	13.33 (2.23)	13.50 (2.27)
	A few times	583	21.19 (2.45)	21.45 (2.50)
	Many times	572	19.65 (2.75)	19.53 (2.76)
	Overall	1,744	17.98 (1.39)	18.04 (1.39)
Past Month	0 times	199	2.02 (0.89)	2.14 (0.93)
	1-2 times	390	6.21 (1.73)	6.25 (1.73)
	A few times	583	10.03 (1.74)	10.14 (1.80)
	Many times	572	8.15 (1.68)	8.05 (1.66)
	Overall	1,744	7.46 (0.89)	7.50 (0.90)
Cigarettes				
Lifetime	0 times	199	6.97 (3.72)	6.93 (3.71)
	1-2 times	390	9.40 (2.04)	9.42 (2.00)
	A few times	583	10.06 (1.54)	10.09 (1.56)
	Many times	572	11.29 (1.96)	11.39 (2.00)
	Overall	1,744	9.88 (1.08)	9.93 (1.08)
Past Year	0 times	199	3.10 (2.48)	3.01 (2.38)
	1-2 times	390	5.74 (1.66)	5.95 (1.67)
	A few times	583	6.98 (1.40)	7.08 (1.43)
	Many times	572	6.53 (1.25)	6.62 (1.27)
	Overall	1,744	6.02 (0.80)	6.13 (0.80)
Past Month	0 times	199	0.12 (0.09)	0.13 (0.10)
	1-2 times	390	1.35 (0.67)	1.51 (0.73)
	A few times	583	3.85 (0.96)	3.92 (0.98)
	Many times	572	3.61 (0.93)	3.71 (0.96)
	Overall	1,744	2.70 (0.47)	2.79 (0.49)

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: 2016

Drug	Mother Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Any Illicit				
Lifetime	0 times	241	10.48 (2.50)	10.46 (2.56)
	1-2 times	500	18.35 (2.57)	19.07 (2.66)
	A few times	793	17.36 (2.01)	17.53 (2.02)
	Many times	1,473	26.56 (1.86)	26.61 (1.87)
	Overall	3,007	21.19 (1.20)	21.40 (1.22)
Past Year	0 times	241	6.47 (1.98)	6.46 (2.04)
	1-2 times	500	13.69 (2.34)	14.37 (2.44)
	A few times	793	11.73 (1.72)	11.73 (1.72)
	Many times	1,473	19.45 (1.67)	19.45 (1.68)
	Overall	3,007	15.19 (1.01)	15.32 (1.03)
Past Month	0 times	241	2.97 (1.02)	2.90 (1.00)
	1-2 times	500	3.40 (1.14)	3.69 (1.35)
	A few times	793	6.09 (1.10)	5.99 (1.07)
	Many times	1,473	10.82 (1.32)	10.89 (1.33)
	Overall	3,007	7.53 (0.73)	7.59 (0.75)
Marijuana				
Lifetime	0 times	241	4.45 (1.34)	4.34 (1.29)
	1-2 times	500	9.85 (2.06)	10.39 (2.16)
	A few times	793	12.30 (1.80)	12.37 (1.81)
	Many times	1,473	18.49 (1.71)	18.60 (1.72)
	Overall	3,007	14.03 (1.03)	14.19 (1.05)
Past Year	0 times	241	2.84 (1.11)	2.67 (1.03)
	1-2 times	500	8.83 (1.98)	9.42 (2.09)
	A few times	793	10.49 (1.69)	10.43 (1.68)
	Many times	1,473	15.33 (1.55)	15.43 (1.56)
	Overall	3,007	11.75 (0.91)	11.87 (0.93)
Past Month	0 times	241	2.10 (0.83)	2.02 (0.80)
	1-2 times	500	3.05 (1.12)	3.32 (1.33)
	A few times	793	5.73 (1.11)	5.62 (1.08)
	Many times	1,473	9.31 (1.26)	9.45 (1.28)
	Overall	3,007	6.59 (0.70)	6.67 (0.72)

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: 2016

Drug	Father Talked about Dangers with Child	*N*	Baseline[1]	Final[2]
Any Illicit				
Lifetime	0 times	199	13.91 (3.12)	13.49 (3.03)
	1-2 times	390	14.64 (2.69)	14.73 (2.67)
	A few times	583	20.23 (2.37)	20.48 (2.45)
	Many times	572	24.78 (3.08)	25.04 (3.13)
	Overall	1,744	19.55 (1.43)	19.69 (1.45)
Past Year	0 times	199	8.53 (3.15)	8.11 (2.94)
	1-2 times	390	9.98 (2.17)	10.12 (2.19)
	A few times	583	14.63 (1.94)	14.76 (1.97)
	Many times	572	20.45 (2.99)	20.57 (3.03)
	Overall	1,744	14.61 (1.31)	14.66 (1.32)
Past Month	0 times	199	3.42 (1.80)	3.04 (1.52)
	1-2 times	390	6.22 (1.92)	6.34 (1.94)
	A few times	583	6.53 (1.35)	6.52 (1.35)
	Many times	572	8.92 (1.86)	9.02 (1.90)
	Overall	1,744	6.81 (0.91)	6.81 (0.92)
Marijuana				
Lifetime	0 times	199	7.08 (3.01)	6.80 (2.82)
	1-2 times	390	7.77 (1.95)	7.91 (1.96)
	A few times	583	13.45 (1.88)	13.51 (1.91)
	Many times	572	17.11 (2.74)	17.47 (2.81)
	Overall	1,744	12.46 (1.23)	12.59 (1.24)
Past Year	0 times	199	6.45 (2.98)	6.17 (2.78)
	1-2 times	390	7.28 (1.93)	7.43 (1.94)
	A few times	583	11.38 (1.70)	11.47 (1.74)
	Many times	572	15.38 (2.65)	15.57 (2.70)
	Overall	1,744	11.06 (1.17)	11.15 (1.18)
Past Month	0 times	199	2.61 (1.61)	2.34 (1.34)
	1-2 times	390	5.17 (1.78)	5.25 (1.80)
	A few times	583	5.83 (1.27)	5.80 (1.27)
	Many times	572	8.07 (1.81)	8.18 (1.85)
	Overall	1,744	5.96 (0.87)	5.97 (0.88)

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

Center for Behavioral Health Statistics and Quality. (2017a). *2016 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. (2017b). *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. (2018a). *2016 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. (2018b). *2016 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 http://www.amstat.org/sections/srms/proceedings/

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 http://www.amstat.org/sections/srms/proceedings/

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 http://www.amstat.org/sections/srms/proceedings/

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 http://www.amstat.org/sections/srms/proceedings/

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

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, 2018a) 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 2016, 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 2016 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 2016 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	1.63	3.98	1.30	2.3356	243	(0.73, 1.90)	(0.74, 1.90)
	1.66	4.05	1.10	2.3374	243	(0.57, 2.80)	(0.58, 2.80)
						(0.90, 1.46)	(0.90, 1.46)
res.qdu.nr	1.63	3.81	1.00	2.3008	243	(1.00, 2.80)	(1.00, 2.80)
	1.28	3.74	0.91	2.4491	243	(1.00, 2.62)	(1.00, 2.60)
						(1.40, 1.69)	(1.40, 1.69)
res.qdu.ps	1.28	3.74	0.91	2.4491	243	(0.20, 1.70)	(0.91, 1.70)
	1.29	3.69	0.65	2.4459	243	(0.20, 5.00)	(0.71, 1.49)
						(0.90, 5.00)	(0.95, 1.04)

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 2016 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]	
	1-11	duwght12	1-12	duwght13	1-13	duwght14	1-14	duwght15	1-15
Minimum	22	1.00	31	0.23	20	0.54	23	0.67	24
1%	85	1.00	102	0.69	104	1.00	133	0.90	133
5%	126	1.00	171	0.80	170	1.06	215	0.96	216
10%	176	1.00	237	0.86	237	1.12	306	0.98	306
25%	276	1.00	529	0.92	521	1.23	648	0.99	641
Median	851	1.34	1,113	0.99	1,106	1.37	1,432	1.00	1,431
75%	1,077	2.04	1,882	1.07	1,887	1.54	2,663	1.01	2,676
90%	1,365	4.55	3,656	1.17	3,733	1.72	5,341	1.02	5,343
95%	1,626	6.49	5,151	1.25	5,290	1.85	7,836	1.03	7,808
99%	2,427	9.24	8,925	1.56	9,217	2.21	13,687	1.10	13,736
Maximum	8,510	12.65	39,464	2.80	34,191	2.80	29,207	1.49	25,371
n	13,414	-	13,414	-	13,414	-	9,552	-	9,552
Mean	805	2.10	1,626	1.01	1,634	1.40	2,294	1.00	2,226
Max/Mean	11	-	24	-	21	-	13	-	11

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

All 243 proposed effects were kept in the model.

Respondent Questionnaire Dwelling Unit-Level Nonresponse

All 243 proposed effects were kept in the model.

Respondent Questionnaire Dwelling Unit-Level Poststratification

All 243 proposed effects were kept in the model.

Exhibit C.1.1 Covariates for 2016 NSDUH Questionnaire Dwelling Unit Weights (sel.qdu.ps) Model Group 1: Northeast

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		**60**	**60**	
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)	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**	**243**	

Exhibit C.1.2 Covariates for 2016 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 1: Northeast

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		**60**	**60**	
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)	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**	**243**	

Exhibit C.1.3 Covariates for 2016 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 1: Northeast

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		**60**	**60**	
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)	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**	**243**	

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 2016 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	1.15	1.38	0.27	2.0027	300	(0.65, 1.10)	(0.65, 1.10)
	0.81	1.33	0.14	1.9870	300	(0.49, 2.06)	(0.49, 2.06)
						(0.90, 1.39)	(0.90, 1.39)
res.qdu.nr	0.84	1.68	0.14	1.9943	300	(1.00, 1.50)	(1.00, 1.50)
	0.66	0.83	0.12	2.1539	300	(1.00, 3.12)	(1.00, 3.12)
						(1.30, 1.62)	(1.30, 1.60)
res.qdu.ps	0.66	0.83	0.12	2.1539	300	(0.20, 1.10)	(0.90, 1.10)
	0.62	0.72	0.03	2.1549	300	(0.20, 5.00)	(0.73, 2.18)
						(0.90, 5.00)	(0.90, 1.07)

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 2016 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]	
	1-11	duwght12	1-12	duwght13	1-13	duwght14	1-14	duwght15	1-15
Minimum	27	1.00	27	0.40	52	0.64	58	0.52	58
1%	98	1.00	117	0.69	117	1.01	136	0.95	137
5%	132	1.00	201	0.82	205	1.06	256	0.99	256
10%	175	1.00	335	0.87	335	1.12	414	0.99	413
25%	586	1.00	753	0.94	742	1.22	907	1.00	909
Median	873	1.31	1,180	1.00	1,178	1.34	1,487	1.00	1,487
75%	1,112	2.07	1,925	1.06	1,938	1.48	2,652	1.00	2,646
90%	1,384	4.18	3,692	1.13	3,665	1.61	5,098	1.01	5,096
95%	1,590	6.06	5,255	1.20	5,213	1.72	7,723	1.01	7,724
99%	2,062	8.38	8,526	1.41	8,407	2.02	12,333	1.04	12,313
Maximum	4,717	11.43	27,378	2.74	18,220	3.12	24,955	2.18	24,961
n	16,025	-	16,025	-	16,025	-	11,768	-	11,768
Mean	860	2.01	1,686	1.00	1,682	1.36	2,290	1.00	2,290
Max/Mean	5	-	16	-	11	-	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

All 300 proposed effects were kept in the model.

Respondent Questionnaire Dwelling Unit-Level Nonresponse

All 300 proposed effects were kept in the model.

Respondent Questionnaire Dwelling Unit-Level Poststratification

All 300 proposed effects were kept in the model.

Exhibit C.2.1 Covariates for 2016 NSDUH Questionnaire Dwelling Unit Weights (sel.qdu.ps) Model Group 2: Midwest

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		**66**	**66**	
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)	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**	**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**	**300**	

Exhibit C.2.2 Covariates for 2016 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 2: Midwest

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		**66**	**66**	
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)	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**	**300**	

Exhibit C.2.3 Covariates for 2016 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 2: Midwest

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		**66**	**66**	
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)	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**	**300**	

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 2016 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	1.33	2.26	0.43	1.9061	339	(0.65, 1.90)	(0.66, 1.90)
	1.13	1.85	0.26	1.8915	339	(0.58, 2.27)	(0.59, 2.27)
						(0.90, 1.18)	(0.90, 1.18)
res.qdu.nr	1.08	2.33	0.37	1.8991	339	(1.00, 1.40)	(1.00, 1.40)
	0.63	0.77	0.10	2.0050	337	(1.00, 2.85)	(1.00, 2.85)
						(1.30, 2.48)	(1.30, 2.48)
res.qdu.ps	0.63	0.77	0.10	2.0050	339	(0.25, 1.10)	(0.25, 1.10)
	0.61	0.87	0.05	2.0059	339	(0.46, 3.07)	(0.47, 3.03)
						(0.90, 1.04)	(0.90, 1.03)

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 2016 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]	
	1-11	duwght12	1-12	duwght13	1-13	duwght14	1-14	duwght15	1-15
Minimum	12	1.00	12	0.40	14	0.53	14	0.17	12
1%	68	1.00	99	0.77	95	1.01	111	0.93	111
5%	147	1.00	258	0.85	259	1.07	332	0.98	334
10%	302	1.00	443	0.89	447	1.10	578	0.99	574
25%	737	1.00	933	0.94	933	1.19	1,164	1.00	1,164
Median	1,047	1.30	1,533	1.00	1,536	1.31	1,922	1.00	1,922
75%	1,454	2.20	2,533	1.06	2,540	1.44	3,281	1.00	3,283
90%	1,925	3.97	4,465	1.12	4,429	1.56	6,059	1.01	6,045
95%	2,231	5.34	6,193	1.17	6,155	1.66	8,744	1.02	8,766
99%	3,028	8.71	9,912	1.38	10,015	1.94	14,310	1.05	14,384
Maximum	9,132	10.97	24,754	3.09	24,905	2.85	32,140	3.03	32,010
n	22,287	-	22,287	-	22,287	-	16,810	-	16,810
Mean	1,119	1.98	2,109	1.01	2,106	1.33	2,792	1.00	2,792
Max/Mean	8	-	12	-	12	-	12	-	11

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

All 339 proposed effects were kept in the model.

Respondent Questionnaire Dwelling Unit-Level Nonresponse

Of the 339 proposed effects, 337 were kept in the model. The American Indian or Alaska Native and Asian Race categories were combined because small sample sizes led to convergence problems for Maryland and West Virginia.

Respondent Questionnaire Dwelling Unit-Level Poststratification

All 339 proposed effects were kept in the model.

Exhibit C.3.1 Covariates for 2016 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**	**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.3.2 Covariates for 2016 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**	**211**	
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	62	Coll. (9,3) & (9,4), (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**	**337**	

Exhibit C.3.3 Covariates for 2016 NSDUH Questionnaire Dwelling Unit Weights (res.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**	**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**	

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 2016 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	1.36	2.08	0.45	2.1421	270	(0.77, 1.30)	(0.77, 1.30)
	1.07	1.97	0.26	2.1719	264	(0.51, 4.00)	(0.51, 4.00)
						(0.90, 1.21)	(0.90, 1.21)
res.qdu.nr	1.03	2.05	0.26	2.2291	270	(1.00, 1.50)	(1.00, 1.50)
	0.61	0.86	0.09	2.3767	264	(1.00, 3.51)	(1.00, 3.51)
						(1.30, 5.00)	(1.30, 1.30)
res.qdu.ps	0.61	0.86	0.09	2.3767	270	(0.20, 1.10)	(0.96, 1.10)
	0.69	1.24	0.05	2. 3793	264	(0.20, 5.00)	(0.93, 1.55)
						(0.90, 5.00)	(0.96, 0.99)

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 2016 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]	
	1-11	duwght12	1-12	duwght13	1-13	duwght14	1-14	duwght15	1-15
Minimum	29	1.00	29	0.46	19	0.69	21	0.59	21
1%	80	1.00	90	0.72	88	1.00	104	0.97	103
5%	113	1.00	141	0.84	142	1.07	173	0.99	173
10%	143	1.00	203	0.88	198	1.11	245	0.99	245
25%	262	1.00	450	0.94	454	1.19	563	1.00	562
Median	874	1.28	1,261	1.00	1,257	1.30	1,501	1.00	1,502
75%	1,495	2.03	2,215	1.07	2,254	1.42	2,900	1.00	2,897
90%	1,866	3.57	3,742	1.15	3,815	1.57	5,387	1.01	5,388
95%	2,072	5.02	5,458	1.21	5,495	1.68	7,860	1.01	7,841
99%	2,651	8.11	9,218	1.40	9,324	2.02	13,700	1.04	13,718
Maximum	7,413	11.28	17,470	4.00	24,048	3.51	31,020	1.55	31,262
n	15,848	-	15,848	-	15,848	-	11,965	-	11,965
Mean	954	1.89	1,728	1.01	1,747	1.33	2,314	1.00	2,314
Max/Mean	8	-	10	-	14	-	13	-	14

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, 264 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 all levels of Rent/Housing. Also combined were 50-100 and 10-<50 percent Black or African American × 0-<10 and 10-<50 percent Owner-Occupied.

Respondent Questionnaire Dwelling Unit-Level Nonresponse

This step used the same set of 264 effects as the selected questionnaire dwelling unit-level poststratification.

Respondent Questionnaire Dwelling Unit-Level Poststratification

This step used the same set of 264 effects as the selected questionnaire dwelling unit-level poststratification.

Exhibit C.4.1 Covariates for 2016 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**	**167**	
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	4	Coll. (1,1) & (2,1), (1,2) & (2,2), (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.4.2 Covariates for 2016 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 4: West

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

Exhibit C.4.3 Covariates for 2016 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 2016 NSDUH QDU-Level Response Rates

Domain	Selected QDU	Respondent QDU	% Interview Response Rate[1]
Total	67,574	50,095	71.67
Census Region			
Northeast	13,414	9,552	68.66
South	22,287	16,810	73.44
Midwest	16,025	11,768	70.39
West	15,848	11,965	72.27
Quarter			
Quarter 1	15,788	11,708	72.20
Quarter 2	17,471	12,939	71.17
Quarter 3	17,529	13,032	72.21
Quarter 4	16,786	12,416	71.09
Household Type			
12-17, 18-25, 26+	5,015	4,061	81.09
12-17, 18-25	61	48	85.68
12-17, 26+	14,976	11,913	79.65
18-25, 26+	11,540	8,622	74.72
12-17	20	13	58.37
18-25	5,278	4,113	77.82
26+	30,684	21,325	68.90
Race/Ethnicity of Householder			
Hispanic or Latino White	9,206	7,080	74.30
Hispanic or Latino Black or African American	222	178	73.31
Hispanic or Latino Other	582	454	76.84
Non-Hispanic or Latino White	44,087	32,005	70.32
Non-Hispanic or Latino Black or African American	7,881	6,317	77.95
Non-Hispanic or Latino Other	5,596	4,061	68.15
% Hispanic or Latino in Segment			
50-100%	5,161	3,985	74.56
10-<50%	17,079	12,775	72.27
<10%	45,334	33,335	71.02
% Black or African American in Segment			
50-100%	4,928	3,938	77.24
10-<50%	12,781	9,668	73.59
<10%	49,865	36,489	70.54
% Owner-Occupied DUs in Segment			
50-100%	49,452	36,355	71.20
10-<50%	14,247	10,771	72.83
<10%	3,875	2,969	73.41
Combined Median Rent/Housing Value			
1st Quintile	10,908	8,372	75.06
2nd Quintile	15,305	11,685	74.09
3rd Quintile	15,531	11,614	72.03
4th Quintile	14,721	10,642	70.22
5th Quintile	11,109	7,782	67.46
Population Density			
Large MSA	29,211	21,300	70.24
Medium to Small MSA	33,069	24,825	73.29
Non-MSA, Urban	1,672	1,305	77.13
Non-MSA, Rural	3,622	2,665	70.72
Group Quarters			
Group	529	475	85.82
Non-Group	67,045	49,620	71.60
Household Size			
One	9,217	6,799	72.01
Two	28,192	20,110	69.10
Three	16,495	12,381	73.40
Four or More	13,670	10,805	78.29

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 2016 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	67,574	1.87	3.72	0.92	1.35	2.33	0.55	1.15	2.15	0.38
Census Region										
Northeast	13,414	2.01	5.40	1.75	1.63	3.98	1.30	1.66	4.05	1.10
South	22,287	1.84	3.51	0.79	1.33	2.26	0.43	1.13	1.85	0.26
Midwest	16,025	1.91	3.23	0.54	1.15	1.38	0.27	0.81	1.33	0.14
West	15,848	1.75	3.30	0.88	1.36	2.08	0.45	1.07	1.97	0.26
Quarter										
Quarter 1	15,788	2.25	4.16	1.07	1.60	2.77	0.64	1.29	2.53	0.41
Quarter 2	17,471	1.66	3.56	0.84	1.15	2.17	0.53	1.06	2.01	0.43
Quarter 3	17,529	1.75	3.15	0.85	1.24	1.83	0.54	0.96	1.62	0.28
Quarter 4	16,786	1.85	3.99	0.92	1.45	2.56	0.50	1.30	2.46	0.40
Household Type										
12-17, 18-25, 26+	5,015	1.73	4.35	1.32	1.73	4.35	1.32	1.87	4.28	0.86
12-17, 18-25	61	1.64	5.01	0.27	0.00	0.00	0.00	0.00	0.00	0.00
12-17, 26+	14,976	1.83	3.79	0.96	1.84	3.79	0.96	1.49	3.39	0.66
18-25, 26+	11,540	1.77	3.73	0.95	1.66	3.56	0.85	1.33	3.31	0.64
12-17	20	5.00	3.06	0.23	0.00	0.00	0.00	0.00	0.00	0.00
18-25	5,278	3.39	5.69	1.20	3.05	5.08	1.12	2.08	4.55	0.76
26+	30,684	1.68	3.24	0.78	0.65	1.60	0.36	0.63	1.48	0.24
Race/Ethnicity of Householder										
Hispanic or Latino White	9,206	1.65	2.57	0.59	1.11	1.61	0.31	0.87	1.59	0.29
Hispanic or Latino Black or African American	222	54.50	77.39	33.40	41.44	65.66	30.03	52.25	70.12	21.21
Hispanic or Latino Other	582	22.68	45.58	13.86	14.43	25.32	6.56	13.75	24.44	4.61
Non-Hispanic or Latino White	44,087	0.94	1.59	0.28	0.79	1.19	0.17	0.54	0.98	0.09
Non-Hispanic or Latino Black or African American	7,881	2.35	4.08	0.89	1.48	2.41	0.48	1.56	2.40	0.38
Non-Hispanic or Latino Other	5,596	4.59	8.03	1.39	3.04	4.68	0.92	2.43	3.73	0.57

Table E.1 2016 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%	5,161	2.05	5.24	1.78	1.53	3.59	1.15	2.05	5.06	1.10
10-<50%	17,079	2.14	4.59	1.28	1.56	3.08	0.86	1.51	2.96	0.61
<10%	45,334	1.75	3.02	0.60	1.26	1.83	0.34	0.91	1.41	0.18
% Black or African American in Segment										
50-100%	4,928	1.93	4.02	1.10	1.14	2.36	0.75	1.56	3.23	0.78
10-<50%	12,781	2.89	5.89	1.60	1.74	3.20	0.89	1.64	3.22	0.75
<10%	49,865	1.60	3.01	0.69	1.28	2.08	0.44	0.98	1.74	0.24
% Owner-Occupied DUs in Segment										
50-100%	49,452	1.35	2.61	0.59	1.06	1.70	0.33	0.92	1.64	0.24
10-<50%	14,247	2.70	5.61	1.56	1.85	3.61	0.98	1.61	3.19	0.65
<10%	3,875	5.50	10.06	2.57	3.28	5.72	1.82	2.32	4.98	1.25
Combined Median Rent/Housing Value										
1st Quintile	10,908	1.56	3.38	0.77	1.05	1.77	0.34	0.90	1.81	0.26
2nd Quintile	15,305	1.82	3.37	0.92	1.25	2.15	0.61	1.05	2.26	0.50
3rd Quintile	15,531	1.96	3.90	1.13	1.48	2.54	0.75	1.20	2.10	0.41
4th Quintile	14,721	1.83	3.66	0.95	1.32	2.23	0.58	1.28	2.11	0.39
5th Quintile	11,109	2.16	4.19	0.73	1.66	2.86	0.38	1.27	2.43	0.30
Population Density										
Large MSA[1]	29,211	1.92	4.04	1.06	1.47	2.77	0.71	1.43	2.59	0.51
Medium to Small MSA[1]	33,069	1.92	3.53	0.81	1.33	2.01	0.41	0.98	1.81	0.26
Non-MSA,[1] Urban	1,672	1.97	2.48	0.33	1.20	0.76	0.10	0.84	1.12	0.14
Non-MSA,[1] Rural	3,622	0.91	1.57	0.36	0.69	0.71	0.13	0.50	0.49	0.06
Group Quarters										
Group	529	5.10	10.44	1.91	3.59	5.37	1.21	1.51	3.49	0.74
Non-Group	67,045	1.84	3.68	0.91	1.34	2.32	0.55	1.14	2.15	0.38
Household Size										
One	9,217	1.86	3.66	0.81	0.73	1.29	0.29	0.63	1.32	0.23
Two	28,192	1.87	3.48	0.83	1.31	2.21	0.49	0.99	1.82	0.29
Three	16,495	1.76	3.71	1.01	1.56	3.46	0.84	1.42	3.41	0.64
Four or More	13,670	2.01	4.19	1.04	1.63	3.65	0.98	1.49	3.59	0.72

[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 2016 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	50,095	1.12	2.38	0.40	0.76	1.33	0.24	0.76	1.42	0.15
Census Region										
Northeast	9,552	1.63	3.81	1.00	1.28	3.74	0.91	1.29	3.69	0.65
South	16,810	1.08	2.33	0.37	0.63	0.77	0.10	0.61	0.87	0.05
Midwest	11,768	0.84	1.68	0.14	0.66	0.83	0.12	0.62	0.72	0.03
West	11,965	1.03	2.05	0.26	0.61	0.86	0.09	0.69	1.24	0.05
Quarter										
Quarter 1	11,708	1.30	2.80	0.39	0.74	1.26	0.22	0.73	1.25	0.15
Quarter 2	12,939	1.05	2.04	0.40	0.76	1.37	0.26	0.71	1.28	0.15
Quarter 3	13,032	0.84	1.78	0.32	0.64	1.03	0.18	0.71	1.47	0.14
Quarter 4	12,416	1.30	2.90	0.50	0.89	1.65	0.32	0.89	1.69	0.16
Household Type										
12-17, 18-25, 26+	4,061	1.53	3.85	0.80	1.38	3.66	0.69	1.16	3.17	0.45
12-17, 18-25	48	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12-17, 26+	11,913	1.49	3.56	0.70	1.14	2.73	0.51	1.20	2.65	0.28
18-25, 26+	8,622	1.37	3.44	0.69	0.96	2.51	0.45	1.08	2.87	0.35
12-17	13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18-25	4,113	1.95	4.41	0.70	1.19	2.70	0.38	1.09	2.36	0.21
26+	21,325	0.57	1.72	0.24	0.26	0.68	0.13	0.25	0.82	0.07
Race/Ethnicity of Householder										
Hispanic or Latino White	7,080	0.73	1.33	0.23	0.55	1.13	0.20	0.59	1.34	0.13
Hispanic or Latino Black or African American	178	50.00	62.16	17.52	30.90	38.52	11.95	36.52	45.31	8.90
Hispanic or Latino Other	454	13.22	26.86	4.77	7.71	12.84	1.35	8.81	16.50	1.23
Non-Hispanic or Latino White	32,005	0.54	1.37	0.15	0.32	0.46	0.06	0.27	0.47	0.02
Non-Hispanic or Latino Black or African American	6,317	1.31	2.13	0.39	0.89	1.64	0.23	0.79	1.60	0.19
Non-Hispanic or Latino Other	4,061	2.54	4.26	0.67	2.27	3.91	0.70	2.39	3.71	0.31
% Hispanic or Latino in Segment										
50-100%	3,985	1.81	5.04	1.15	0.85	2.28	0.59	1.10	3.21	0.45
10-<50%	12,775	1.43	2.94	0.56	0.94	1.79	0.31	0.99	1.92	0.24
<10%	33,335	0.91	1.77	0.23	0.67	1.00	0.17	0.63	0.97	0.07

Table E.2 2016 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%	3,938	1.52	3.18	0.66	0.74	1.66	0.36	0.79	1.82	0.30
10-<50%	9,668	1.55	3.16	0.79	1.02	2.02	0.41	1.09	2.22	0.33
<10%	36,489	0.96	2.06	0.26	0.69	1.10	0.19	0.67	1.15	0.08
% Owner-Occupied DUs in Segment										
50-100%	36,355	0.89	1.95	0.27	0.61	0.92	0.15	0.61	1.01	0.09
10-<50%	10,771	1.53	3.22	0.69	1.02	2.26	0.45	1.04	2.32	0.28
<10%	2,969	2.32	4.65	1.00	1.62	3.13	0.69	1.62	3.37	0.51
Combined Median Rent/Housing Value										
1st Quintile	8,372	0.87	2.04	0.34	0.50	0.73	0.13	0.45	0.86	0.11
2nd Quintile	11,685	0.97	2.19	0.44	0.52	0.91	0.14	0.58	1.25	0.11
3rd Quintile	11,614	1.11	2.23	0.43	0.77	1.35	0.29	0.86	1.50	0.18
4th Quintile	10,642	1.15	1.93	0.37	0.80	1.45	0.33	0.81	1.44	0.20
5th Quintile	7,782	1.57	3.71	0.42	1.31	2.12	0.29	1.14	1.95	0.14
Population Density										
Large MSA[1]	21,300	1.42	2.64	0.48	1.02	1.86	0.37	0.98	1.85	0.22
Medium to Small MSA[1]	24,825	0.93	2.30	0.35	0.56	0.78	0.11	0.61	1.03	0.07
Non-MSA,[1] Urban	1,305	0.77	1.16	0.16	0.61	0.61	0.05	0.46	0.54	0.04
Non-MSA,[1] Rural	2,665	0.64	0.55	0.06	0.56	0.43	0.09	0.56	0.38	0.02
Group Quarters										
Group	475	1.26	2.10	0.25	0.21	0.14	0.02	0.21	0.16	0.01
Non-Group	49,620	1.11	2.38	0.40	0.76	1.33	0.25	0.77	1.43	0.15
Household Size										
One	6,799	0.56	1.13	0.10	0.25	0.31	0.04	0.28	0.54	0.02
Two	20,110	0.97	2.38	0.38	0.56	1.16	0.21	0.54	1.26	0.13
Three	12,381	1.36	3.50	0.71	0.95	2.24	0.43	0.98	2.16	0.23
Four or More	10,805	1.46	3.56	0.69	1.21	3.03	0.58	1.23	3.02	0.41

[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 2016 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	50,095	123,488,955	123,488,955	123,488,955	0.00	0.00
Census Region						
Northeast	9,552	21,913,889	21,913,889	21,913,889	0.00	-0.00
South	16,810	46,933,442	46,933,442	46,933,442	0.00	0.00
Midwest	11,768	26,953,036	26,953,036	26,953,036	0.00	-0.00
West	11,965	27,688,588	27,688,588	27,688,588	-0.00	-0.00
Quarter						
Quarter 1	11,708	30,697,116	30,697,116	30,697,116	0.00	0.00
Quarter 2	12,939	30,947,856	30,947,856	30,947,856	0.00	0.00
Quarter 3	13,032	30,919,079	30,919,079	30,919,079	0.00	0.00
Quarter 4	12,416	30,924,903	30,924,903	30,924,903	-0.00	0.00
Household Type						
12-17, 18-25, 26+	4,061	5,099,020	5,099,020	5,099,020	0.00	0.00
12-17, 18-25	48	61,350	61,350	61,350	0.00	0.00
12-17, 26+	11,913	13,548,942	13,548,942	13,548,942	0.00	-0.00
18-25, 26+	8,622	14,050,015	14,050,015	14,050,015	-0.00	0.00
12-17	13	9,818	9,818	9,818	0.00	0.00
18-25	4,113	5,705,413	5,705,413	5,705,413	0.00	0.00
26+	21,325	85,014,396	85,014,396	85,014,396	0.00	0.00
Race/Ethnicity of Householder						
Hispanic or Latino White	7,080	14,684,644	14,684,644	14,684,644	0.00	0.00
Hispanic or Latino Black or African American	178	900,407	900,407	900,407	-0.00	-0.00
Hispanic or Latino Other	454	1,183,038	1,183,038	1,183,038	-0.00	-0.00
Non-Hispanic or Latino White	32,005	82,691,913	82,691,913	82,691,913	0.00	-0.00
Non-Hispanic or Latino Black or African American	6,317	15,206,515	15,206,515	15,206,515	-0.00	0.00
Non-Hispanic or Latino Other	4,061	8,822,438	8,822,438	8,822,438	0.00	0.00
% Hispanic or Latino in Segment						
50-100%	3,985	9,838,395	9,838,395	9,838,395	-0.00	-0.00
10-<50%	12,775	35,814,248	35,814,248	35,814,248	0.00	-0.00
<10%	33,335	77,836,311	77,836,311	77,836,311	0.00	0.00
% Black or African American in Segment						
50-100%	3,938	9,268,402	9,268,402	9,268,402	0.00	0.00
10-<50%	9,668	25,109,886	25,109,886	25,109,886	0.00	-0.00
<10%	36,489	89,110,666	89,110,667	89,110,666	0.00	0.00
% Owner-Occupied DUs in Segment						
50-100%	36,355	90,331,647	90,331,647	90,331,647	0.00	0.00
10-<50%	10,771	26,404,004	26,404,004	26,404,004	0.00	0.00
<10%	2,969	6,753,304	6,753,304	6,753,304	-0.00	-0.00

Table F.1 2016 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	8,372	18,224,728	18,224,728	18,224,728	0.00	0.00
2nd Quintile	11,685	26,408,649	26,408,649	26,408,649	0.00	0.00
3rd Quintile	11,614	28,132,046	28,132,046	28,132,046	0.00	-0.00
4th Quintile	10,642	28,014,830	28,014,830	28,014,830	0.00	-0.00
5th Quintile	7,782	22,708,701	22,708,701	22,708,701	0.00	-0.00
Population Density						
Large MSA	21,300	64,605,903	64,605,903	64,605,903	0.00	-0.00
Medium to Small MSA	24,825	51,366,336	51,366,336	51,366,336	0.00	0.00
Non-MSA, Urban	1,305	2,522,824	2,522,824	2,522,824	0.00	0.00
Non-MSA, Rural	2,665	4,993,891	4,993,891	4,993,891	0.00	0.00
Group Quarters						
Group	475	552,960	552,960	552,960	0.00	0.00
Non-Group	49,620	122,935,994	122,935,994	122,935,994	0.00	0.00
Household Size						
One	6,799	33,501,523	33,494,459	32,864,568	1.94	1.92
Two	20,110	55,485,053	55,491,367	56,243,224	-1.35	-1.34
Three	12,381	19,390,591	19,390,859	19,573,466	-0.93	-0.93
Four or More	10,805	15,111,787	15,112,270	14,807,696	2.05	2.06

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 2016 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	67,574	12	518	929	1,283	9,132	1.41	12	710	1,295	2,177	39,464	2.07	14	705	1,296	2,196	34,191	2.06
Census Region																			
Northeast	13,414	22	276	851	1,077	8,510	1.46	31	529	1,113	1,882	39,464	2.34	20	521	1,106	1,887	34,191	2.34
South	22,287	12	737	1,047	1,454	9,132	1.33	12	933	1,533	2,533	24,754	1.91	14	933	1,536	2,540	24,905	1.89
Midwest	16,025	27	586	873	1,112	4,717	1.27	27	753	1,180	1,925	27,378	2.00	52	742	1,178	1,938	18,220	1.99
West	15,848	29	262	874	1,495	7,413	1.56	29	450	1,261	2,215	17,470	2.14	19	454	1,257	2,254	24,048	2.17
Quarter																			
Quarter 1	15,788	12	558	973	1,369	9,132	1.41	12	764	1,374	2,329	39,464	2.05	17	761	1,376	2,342	33,850	2.06
Quarter 2	17,471	27	498	905	1,257	8,339	1.40	27	691	1,273	2,113	35,642	2.06	20	678	1,268	2,136	34,191	2.07
Quarter 3	17,529	22	518	893	1,243	8,357	1.39	22	694	1,262	2,073	37,419	2.07	14	701	1,268	2,088	24,048	2.03
Quarter 4	16,786	15	501	944	1,284	6,774	1.43	15	703	1,295	2,227	24,274	2.08	16	692	1,289	2,224	24,905	2.08
Household Type																			
12-17, 18-25, 26+	5,015	29	565	959	1,349	8,339	1.43	29	565	959	1,349	8,339	1.43	19	556	952	1,368	7,338	1.42
12-17, 18-25	61	73	307	773	1,492	3,807	1.67	73	307	774	1,492	3,807	1.67	86	333	723	1,387	5,247	1.86
12-17, 26+	14,976	12	471	857	1,209	7,524	1.43	12	472	858	1,211	7,524	1.42	16	463	858	1,222	5,468	1.43
18-25, 26+	11,540	16	596	999	1,366	9,132	1.40	22	663	1,172	1,612	9,798	1.39	14	645	1,168	1,649	7,765	1.40
12-17	20	79	136	173	615	1,887	2.35	80	138	175	615	1,907	2.36	84	111	176	623	2,068	2.42
18-25	5,278	38	351	881	1,243	6,062	1.51	43	410	1,023	1,486	6,937	1.49	41	411	1,028	1,486	7,132	1.51
26+	30,684	13	540	940	1,278	8,357	1.39	30	1,229	2,117	3,611	39,464	1.73	39	1,213	2,105	3,628	34,191	1.73
Race/Ethnicity of Householder																			
Hispanic or Latino White	9,206	24	621	987	1,427	5,961	1.31	29	742	1,259	1,916	22,873	1.87	19	730	1,254	1,942	18,220	1.86
Hispanic or Latino Black or African American	222	57	880	1,852	3,501	8,510	1.58	57	1,101	2,463	4,472	39,464	3.09	77	1,453	2,686	4,936	34,191	2.46
Hispanic or Latino Other	582	12	209	827	1,934	9,132	2.17	12	312	1,137	2,587	16,259	2.51	17	307	1,105	2,875	15,460	2.49
Non-Hispanic or Latino White	44,087	27	436	903	1,205	5,665	1.39	27	686	1,275	2,226	24,754	2.09	21	691	1,281	2,251	24,905	2.10
Non-Hispanic or Latino Black or African American	7,881	38	777	1,040	1,371	7,413	1.28	38	940	1,451	2,317	22,879	1.87	37	909	1,451	2,306	19,993	1.86
Non-Hispanic or Latino Other	5,596	15	270	868	1,454	6,997	1.59	15	445	1,218	2,112	19,833	2.00	14	451	1,189	2,078	14,397	1.96
% Hispanic or Latino in Segment																			
50-100%	5,161	49	753	1,170	1,552	8,339	1.29	53	920	1,475	2,199	21,451	1.77	63	907	1,484	2,239	20,888	1.79
10-<50%	17,079	12	714	1,070	1,546	9,132	1.36	12	906	1,551	2,539	39,464	1.95	17	892	1,541	2,530	34,191	1.93
<10%	45,334	15	351	864	1,158	8,510	1.41	15	597	1,183	2,022	27,378	2.14	14	593	1,185	2,033	30,677	2.15

Table G.1 2016 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%	4,928	12	738	1,004	1,300	8,510	1.31	12	877	1,385	2,225	35,642	1.94	14	844	1,398	2,261	34,191	1.95
10-<50%	12,781	18	751	1,025	1,387	7,524	1.30	18	924	1,456	2,309	39,464	1.90	17	914	1,447	2,318	33,850	1.88
<10%	49,865	22	392	891	1,250	9,132	1.45	27	621	1,238	2,135	37,419	2.13	19	620	1,238	2,149	30,677	2.13
% Owner-Occupied DUs[1] in Segment																			
50-100%	49,452	12	509	919	1,260	9,132	1.40	12	711	1,286	2,170	24,754	2.03	19	705	1,291	2,193	30,677	2.04
10-<50%	14,247	15	574	965	1,336	8,510	1.42	15	748	1,334	2,211	37,419	2.10	14	743	1,327	2,222	24,048	2.08
<10%	3,875	16	422	922	1,392	6,062	1.53	24	601	1,250	2,122	39,464	2.43	33	586	1,234	2,138	34,191	2.33
Combined Median Rent/Housing Value																			
1st Quintile	10,908	22	365	794	1,088	6,126	1.44	29	574	1,113	1,952	23,122	2.24	19	569	1,104	1,970	24,905	2.25
2nd Quintile	15,305	37	453	887	1,194	9,132	1.41	37	649	1,223	2,011	39,464	2.12	24	647	1,225	2,054	34,191	2.10
3rd Quintile	15,531	16	515	946	1,299	8,510	1.42	27	704	1,301	2,154	37,419	2.07	21	698	1,297	2,151	22,270	2.06
4th Quintile	14,721	15	613	962	1,355	7,524	1.38	15	784	1,364	2,289	27,378	1.99	16	786	1,374	2,291	20,560	1.98
5th Quintile	11,109	12	679	1,082	1,453	6,469	1.36	12	876	1,514	2,500	24,274	1.94	14	865	1,491	2,503	30,677	1.96
Population Density																			
Large MSA[1]	29,211	12	867	1,128	1,526	9,132	1.26	12	1,061	1,643	2,623	39,464	1.80	14	1,047	1,637	2,615	34,191	1.80
Medium to Small MSA[1]	33,069	22	294	757	1,074	6,997	1.49	27	472	1,029	1,817	27,378	2.29	20	477	1,033	1,846	24,905	2.29
Non-MSA,[1] Urban	1,672	29	266	686	1,019	3,026	1.46	29	435	984	1,776	13,158	2.22	19	439	986	1,802	14,807	2.25
Non-MSA,[1] Rural	3,622	43	186	547	931	4,477	1.61	44	323	846	1,671	14,784	2.43	32	321	850	1,672	15,452	2.43
Group Quarters																			
Group	529	41	138	365	1,117	4,791	2.15	44	202	442	1,382	14,844	2.93	58	181	441	1,348	12,571	3.33
Non-Group	67,045	12	526	931	1,283	9,132	1.41	12	718	1,300	2,183	39,464	2.06	14	714	1,301	2,202	34,191	2.06
Household Size																			
One	9,217	20	456	891	1,193	5,216	1.40	54	1,257	2,569	5,254	39,464	1.80	63	1,269	2,538	5,096	34,191	1.80
Two	28,192	22	504	924	1,254	8,357	1.40	24	829	1,558	2,597	37,419	1.72	32	814	1,548	2,601	20,560	1.73
Three	16,495	12	527	932	1,297	9,132	1.42	12	576	1,023	1,575	13,258	1.56	16	569	1,025	1,586	10,280	1.57
Four or More	13,670	13	557	976	1,391	8,339	1.42	22	569	1,000	1,457	19,833	1.47	14	556	989	1,464	8,006	1.46

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

Table G.2 2016 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	50,095	14	684	1,256	2,109	24,905	2.07	14	867	1,631	2,926	32,140	2.20	12	864	1,631	2,929	32,010	2.21
Census Region																			
Northeast	9,552	20	498	1,058	1,821	20,560	2.30	23	648	1,432	2,663	29,207	2.45	24	641	1,431	2,676	25,371	2.45
South	16,810	14	921	1,510	2,439	24,905	1.90	14	1,164	1,922	3,281	32,140	2.01	12	1,164	1,922	3,283	32,010	2.01
Midwest	11,768	52	719	1,135	1,857	17,607	1.99	58	907	1,487	2,652	24,955	2.15	58	909	1,487	2,646	24,961	2.15
West	11,965	19	432	1,198	2,135	24,048	2.23	21	563	1,501	2,900	31,020	2.38	21	562	1,502	2,897	31,262	2.38
Quarter																			
Quarter 1	11,708	17	734	1,335	2,266	22,270	2.08	17	928	1,726	3,099	31,020	2.21	14	928	1,729	3,093	31,111	2.22
Quarter 2	12,939	20	660	1,225	2,052	19,993	2.06	23	837	1,602	2,861	24,197	2.17	24	835	1,599	2,863	24,178	2.17
Quarter 3	13,032	14	681	1,237	2,013	24,048	2.04	14	859	1,604	2,799	29,571	2.17	12	855	1,604	2,801	31,262	2.17
Quarter 4	12,416	16	676	1,241	2,109	24,905	2.10	17	850	1,618	2,956	32,140	2.25	19	848	1,619	2,959	32,010	2.24
Household Type																			
12-17, 18-25, 26+	4,061	19	556	952	1,371	7,338	1.42	21	646	1,185	1,690	6,920	1.44	21	640	1,183	1,691	6,736	1.43
12-17, 18-25	48	86	336	792	1,532	5,247	1.85	87	386	937	1,727	8,075	2.12	86	386	931	1,683	8,316	2.16
12-17, 26+	11,913	16	468	855	1,220	5,468	1.43	17	558	1,065	1,543	9,045	1.45	14	557	1,064	1,546	5,627	1.44
18-25, 26+	8,622	14	645	1,166	1,647	7,765	1.40	14	826	1,531	2,208	10,476	1.43	12	822	1,531	2,210	9,179	1.43
12-17	13	108	111	177	471	1,629	2.33	185	324	471	627	2,594	1.88	189	326	471	635	2,594	1.87
18-25	4,113	41	400	1,033	1,506	7,132	1.50	48	507	1,311	1,934	9,284	1.51	43	508	1,315	1,939	8,270	1.51
26+	21,325	46	1,189	2,083	3,591	24,905	1.73	58	1,619	2,979	5,300	32,140	1.75	56	1,612	2,976	5,299	32,010	1.76
Race/Ethnicity of Householder																			
Hispanic or Latino White	7,080	19	712	1,225	1,889	16,038	1.83	25	877	1,540	2,454	22,135	2.01	25	876	1,541	2,454	21,535	2.01
Hispanic or Latino Black or African American	178	77	1,469	2,673	4,864	24,048	2.01	77	1,731	3,302	6,663	29,571	2.13	78	1,750	3,429	6,150	31,262	2.15
Hispanic or Latino Other	454	17	314	1,125	2,739	15,460	2.42	17	441	1,356	3,410	22,081	2.59	14	441	1,349	3,412	23,233	2.62
Non-Hispanic or Latino White	32,005	21	663	1,233	2,158	24,905	2.14	21	869	1,663	3,048	32,140	2.23	21	868	1,662	3,049	32,010	2.23
Non-Hispanic or Latino Black or African American	6,317	41	900	1,426	2,234	19,993	1.84	51	1,041	1,704	2,851	24,981	2.00	50	1,041	1,701	2,850	25,215	2.00
Non-Hispanic or Latino Other	4,061	14	420	1,089	1,960	14,397	1.98	14	552	1,454	2,834	25,533	2.22	12	549	1,451	2,842	24,368	2.23
% Hispanic or Latino in Segment																			
50-100%	3,985	63	889	1,460	2,156	20,888	1.79	64	1,123	1,791	2,894	29,571	1.95	61	1,123	1,787	2,897	31,070	1.94
10-<50%	12,775	17	855	1,485	2,439	24,905	1.95	17	1,073	1,919	3,353	32,140	2.07	14	1,071	1,919	3,346	32,010	2.07
<10%	33,335	14	575	1,151	1,964	22,270	2.16	14	742	1,507	2,761	31,020	2.29	12	739	1,508	2,758	31,111	2.29

Table G.2 2016 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%	3,938	14	832	1,370	2,187	19,993	1.90	14	969	1,638	2,783	24,981	2.05	12	969	1,639	2,775	25,215	2.05
10-<50%	9,668	17	892	1,421	2,239	24,048	1.88	17	1,116	1,814	3,034	29,571	2.01	14	1,109	1,812	3,037	31,262	2.01
<10%	36,489	19	598	1,194	2,063	24,905	2.15	21	769	1,578	2,910	32,140	2.28	21	768	1,579	2,912	32,010	2.28
% Owner-Occupied DUs[1] in Segment																			
50-100%	36,355	19	685	1,247	2,104	24,905	2.06	21	877	1,642	2,946	32,140	2.18	21	876	1,643	2,945	32,010	2.18
10-<50%	10,771	14	717	1,290	2,135	24,048	2.10	14	883	1,629	2,915	31,020	2.25	12	878	1,626	2,915	31,262	2.25
<10%	2,969	33	560	1,197	2,079	20,888	2.18	34	713	1,530	2,694	29,571	2.39	33	700	1,528	2,693	31,070	2.40
Combined Median Rent/Housing Value																			
1st Quintile	8,372	19	562	1,078	1,910	24,905	2.27	29	695	1,380	2,510	32,140	2.39	31	698	1,382	2,504	32,010	2.39
2nd Quintile	11,685	24	637	1,193	1,994	20,888	2.09	25	799	1,516	2,652	29,571	2.21	24	795	1,516	2,652	31,070	2.21
3rd Quintile	11,614	21	679	1,270	2,072	22,270	2.07	21	846	1,635	2,868	31,020	2.21	21	846	1,637	2,869	31,111	2.21
4th Quintile	10,642	16	768	1,335	2,213	20,560	2.01	17	984	1,759	3,106	27,068	2.14	19	981	1,758	3,110	24,302	2.14
5th Quintile	7,782	14	828	1,424	2,403	21,233	1.95	14	1,098	2,011	3,482	29,207	2.05	12	1,102	2,011	3,482	28,953	2.05
Population Density																			
Large MSA[1]	21,300	14	1,017	1,583	2,501	22,270	1.80	14	1,321	2,112	3,531	31,020	1.93	12	1,319	2,117	3,539	31,111	1.93
Medium to Small MSA[1]	24,825	20	469	1,010	1,801	24,905	2.31	21	610	1,305	2,402	32,140	2.43	21	609	1,305	2,403	32,010	2.43
Non-MSA,[1] Urban	1,305	19	432	970	1,783	14,807	2.23	29	539	1,211	2,315	19,385	2.32	31	543	1,210	2,297	19,288	2.31
Non-MSA,[1] Rural	2,665	32	304	830	1,571	13,081	2.46	32	410	1,098	2,151	21,236	2.59	33	412	1,099	2,146	21,235	2.59
Group Quarters																			
Group	475	58	174	402	1,292	12,571	3.48	66	222	471	1,473	20,037	3.90	66	221	471	1,436	22,948	4.03
Non-Group	49,620	14	694	1,262	2,114	24,905	2.07	14	881	1,640	2,938	32,140	2.20	12	879	1,641	2,940	32,010	2.20
Household Size																			
One	6,799	63	1,266	2,475	4,884	24,905	1.80	105	1,683	3,394	6,962	32,140	1.85	110	1,674	3,388	6,959	32,010	1.85
Two	20,110	33	787	1,494	2,485	20,560	1.75	34	1,022	2,010	3,552	27,068	1.85	33	1,020	2,004	3,552	22,955	1.85
Three	12,381	16	556	1,002	1,553	10,280	1.58	17	694	1,297	2,054	15,923	1.68	14	694	1,297	2,055	16,556	1.67
Four or More	10,805	14	549	982	1,451	8,006	1.46	14	659	1,246	1,856	9,045	1.53	12	655	1,246	1,860	8,275	1.52

[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 $1 + \left[(n-1)/n\right] * CV^2$, 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 2016, 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 2016 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 2016 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	3.91	16.88	7.72	8.4557	213	(0.20, 1.25)	(0.20, 1.25)
	1.21	3.63	0.48	4.1561	205	(0.20, 2.64)	(0.20, 2.64)
						(0.90, 1.10)	(0.90, 0.90)
res.pr.nr	1.26	3.29	0.44	4.1931	213	(1.02, 1.80)	(1.02, 1.80)
	1.15	3.74	0.67	5.1666	213	(1.00, 4.38)	(1.00, 4.36)
						(1.70, 1.80)	(1.70, 1.70)
res.pr.ps	1.22	3.81	0.71	5.1666	223	(0.62, 1.10)	(0.62, 1.10)
	0.37	1.10	0.05	5.3211	215	(0.32, 1.63)	(0.33, 1.61)
						(0.90, 1.41)	(1.41, 1.41)
res.pr.ev	0.37	1.10	0.05	5.3211	223	(0.99, 1.06)	(0.99, 1.06)
	0.00	0.00	0.00	5.3116	215	(0.96, 1.06)	(0.96, 1.06)
						(0.95, 1.05)	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 2016 Distribution of Weight Adjustment Factors and Weight Products (Model Group 1: Northeast and South)

	SDU Weight	Pair Selection Prob		sel.pr.ps[1]		res.pr.nr[1]		res.pr.ps[1]		res.pr.ev[1]	
	1-11	pairwt12	1-12	pairwt13	1-13	pairwt14	1-14	pairwt15	1-15	pairwt16	1-16
Minimum	15	1.02	17	0.02	3	0.64	4	0.32	2	0.75	2
1%	88	1.18	217	0.20	164	1.01	191	0.38	146	0.96	145
5%	160	1.41	529	0.30	409	1.07	502	0.54	438	0.98	437
10%	229	1.61	902	0.53	677	1.11	883	0.70	823	0.99	820
25%	638	2.61	1,927	0.83	1,813	1.22	2,350	0.91	2,251	0.99	2,249
Median	986	4.12	3,893	1.05	3,985	1.44	5,506	1.03	5,512	1.00	5,519
75%	1,352	8.41	8,591	1.27	8,810	1.76	12,970	1.13	13,114	1.01	13,149
90%	1,855	17.40	17,775	1.60	19,213	2.18	29,504	1.23	29,050	1.01	29,097
95%	2,177	28.84	32,125	1.90	32,456	2.50	52,717	1.32	52,770	1.02	52,954
99%	3,098	60.49	71,760	2.20	74,705	3.34	141,565	1.48	142,881	1.05	143,866
Maximum	8,510	1,206.57	1,298,482	2.64	247,868	4.36	457,845	1.61	509,876	1.06	493,922
n	14,543	-	14,543	-	14,543	-	9,182	-	9,182	-	9,182
Mean	1,046	8.52	8,746	1.06	8,616	1.56	13,647	1.00	13,674	1.00	13,647
Max/Mean	8	-	148	-	29	-	34	-	37	-	36

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, 205 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, 4 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, 215 of 223 proposed factors were retained in the final model. All main and two-factor effects were retained at the proposed levels. Of the 12 three-factor effects, 4 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 2016 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**	**4**	
Pair Race/Ethnicity (4 Levels) × Pair Gender × Pair Age (3 Levels)	4 × 3 × 3	12	4	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); (1,2,3), (2,2,3) & (3,2,3); conv.
Total		**213**	**205**	

Exhibit H.1.2 Covariates for 2016 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 2016 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**	**4**	
Pair Race/Ethnicity (4 Levels) × Pair Gender × Pair Age (3 Levels)	4 × 3 × 3	12	4	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); (1,2,3), (2,2,3) & (3,2,3); conv.
Total		**223**	**215**	

Exhibit H.1.4 Covariates for 2016 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 2016 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	4.00	19.91	9.84	16.071	212	(0.24, 2.50)	(0.25, 2.49)
	2.24	6.49	0.98	4.4258	200	(0.22, 2.79)	(0.23, 2.77)
						(0.95, 1.12)	(0.95, 1.12)
res.pr.nr	2.15	6.66	1.14	4.7236	212	(1.02, 1.79)	(1.02, 1.79)
	1.93	4.50	0.65	5.8513	211	(1.00, 4.32)	(1.00, 4.29)
						N/A	N/A
res.pr.ps	1.98	6.00	0.93	5.8513	222	(0.39, 1.13)	(0.39, 1.13)
	0.80	2.17	0.13	5.8138	209	(0.39, 1.54)	(0.40, 1.52)
						N/A	N/A
res.pr.ev	0.80	2.17	0.13	5.8138	222	(0.98, 1.12)	(0.99, 1.12)
	0.03	0.03	0.00	5.7723	209	(0.89, 1.12)	(0.93, 1.07)
						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 2016 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]	
	1-11	pairwt12	1-12	pairwt13	1-13	pairwt14	1-14	pairwt15	1-15	pairwt16	1-16
Minimum	27	1.02	53	0.07	24	0.52	24	0.15	16	0.67	16
1%	87	1.22	193	0.31	164	0.93	178	0.49	180	0.93	177
5%	127	1.43	388	0.54	364	1.03	430	0.66	437	0.98	435
10%	168	1.66	649	0.66	598	1.06	702	0.78	722	0.98	718
25%	447	2.62	1,526	0.85	1,454	1.16	1,789	0.91	1,825	0.99	1,816
Median	884	4.15	3,382	1.02	3,315	1.36	4,304	1.05	4,399	1.00	4,411
75%	1,297	8.26	7,745	1.22	7,998	1.70	11,099	1.16	10,865	1.01	10,852
90%	1,716	16.19	16,228	1.43	17,936	2.18	27,060	1.26	27,062	1.02	26,752
95%	1,972	26.97	27,433	1.61	29,884	2.60	49,744	1.33	46,830	1.03	46,987
99%	2,578	55.28	61,865	1.99	69,550	3.47	128,528	1.43	133,893	1.05	134,950
Maximum	6,896	1,775.06	2,116,210	2.77	301,034	4.29	471,208	1.52	422,499	1.09	374,658
n	13,490	-	13,490	-	13,490	-	8,665	-	8,665	-	8,665
Mean	930	8.22	7,816	1.04	7,825	1.52	12,182	1.03	12,182	1.00	12,182
Max/Mean	7	-	271	-	38	-	39	-	35	-	31

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, 200 of 212 proposed factors were retained in the final model. All main and two-factor effects were retained at proposed levels. None of the 12 three-factor effects were kept in the model because of convergence problems.

Respondent Pair-Level Nonresponse

In the respondent pair-level nonresponse step, 211 of 212 proposed factors were retained in the final model. The Group Quarters categories for College Dorms and Other Group Quarters were collapsed because of convergence problems.

Respondent Pair-Level Poststratification

In the respondent pair-level poststratification step, 209 of 222 proposed factors were retained in the final model. Similar to the respondent pair-level nonresponse step, the Group Quarters categories for College Dorms and Other Group Quarters were collapsed because of convergence problems. All other main and two-factor effects were retained to proposed levels, whereas none of the 12 three-factor effects were kept in the model because of convergence problems.

Respondent Pair-Level Extreme Value Adjustment

The respondent pair-level extreme value adjustment step used the same variables as the respondent pair-level poststratification step.

Exhibit H.2.1 Covariates for 2016 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**	**0**	
Pair Race/Ethnicity (4 Levels) × Pair Gender × Pair Age (3 Levels)	4 × 3 × 3	12	0	Drop all; conv.
Total		**212**	**200**	

Exhibit H.2.2 Covariates for 2016 NSDUH Pair Weights (res.pr.nr) Model Group 2: Midwest and West

Variables	Level	Proposed	Final	Comments
One-Factor Effects		**75**	**74**	
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	1	Coll. (1) & (2); conv.
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**	**211**	

Exhibit H.2.3 Covariates for 2016 NSDUH Pair Weights (res.pr.ps) 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	1	Coll. (1) & (2); conv.
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**	**0**	
Pair Race/Ethnicity (4 Levels) × Pair Gender × Pair Age (3 Levels)	4 × 3 × 3	12	0	Drop all; conv.
Total		**222**	**209**	

Exhibit H.2.4 Covariates for 2016 NSDUH Pair Weights (res.pr.ev) Model Group 2: Midwest and West

This step used the same variables as the respondent pair-level poststratification step in Exhibit H.2.3.

This page intentionally left blank

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

This page intentionally left blank

Table I.1 2016 NSDUH Person Pair-Level Response Rates

Domain	Selected Pairs	Respondent Pairs	% Interview Response Rate[1]
Total	28,033	17,847	57.33
Pair Age Group			
12-17, 12-17	3,199	2,393	76.26
12-17, 18-25	2,548	1,770	69.94
12-17, 26-34	1,281	902	72.29
12-17, 35-49	5,829	3,929	67.59
12-17, 50+	1,123	711	61.77
18-25, 18-25	3,958	2,560	65.58
18-25, 26-34	1,429	856	60.32
18-25, 35-49	2,013	1,153	59.11
18-25, 50+	1,283	710	55.09
26-34, 26-34	1,518	906	57.31
26-34, 35-49	788	426	52.22
26-34, 50+	443	224	52.26
35-49, 35-49	1,213	625	48.23
35-49, 50+	490	218	40.60
50+, 50+	918	464	52.41
Pair Race/Ethnicity			
Hispanic or Latino	4,909	3,253	61.83
Black or African American	2,845	2,070	66.03
White	15,501	9,534	55.35
Other	2,054	1,208	44.22
White & Black or African American	273	176	62.14
White & Hispanic or Latino	1,133	742	59.90
White & Other	864	559	61.33
Black or African American & Hispanic or Latino	139	99	72.29
Black or African American & Other	141	97	69.90
Hispanic or Latino & Other	174	109	45.98
Pair Gender			
Male, Male	5,972	3,721	56.62
Female, Female	6,088	4,157	62.80
Male, Female	15,973	9,969	55.97
Household Size			
Two	7,064	4,221	54.21
Three	8,816	5,609	57.81
Four or More	12,153	8,017	58.61

Table I.1 2016 NSDUH Person Pair-Level Response Rates (continued)

Domain	Selected Pairs	Respondent Pairs	% Interview Response Rate[1]
Census Region			
Northeast	5,368	3,159	51.35
South	9,175	6,023	60.74
Midwest	6,624	4,255	59.69
West	6,866	4,410	54.90
Quarter			
Quarter 1	6,606	4,317	59.11
Quarter 2	7,211	4,500	55.63
Quarter 3	7,183	4,575	57.58
Quarter 4	7,033	4,455	56.99
% Hispanic or Latino in Segment			
50-100%	2,667	1,751	58.68
10-<50%	7,316	4,646	55.85
<10%	18,050	11,450	57.84
% Black or African American in Segment			
50-100%	1,902	1,366	64.39
10-<50%	5,436	3,522	57.19
<10%	20,695	12,959	56.70
% Owner-Occupied DUs in Segment			
50-100%	20,952	13,239	56.94
10-<50%	5,589	3,609	58.41
<10%	1,492	999	63.85
Combined Median Rent/Housing Value			
1st Quintile	4,349	2,940	64.06
2nd Quintile	6,323	4,202	61.41
3rd Quintile	6,515	4,164	57.83
4th Quintile	6,310	3,855	52.56
5th Quintile	4,536	2,686	53.92
Population Density			
Large MSA	12,392	7,632	54.79
Medium to Small MSA	13,515	8,839	60.79
Non-MSA, Urban	652	437	62.78
Non-MSA, Rural	1,474	939	58.79
Group Quarters			
Group	220	168	69.98
Non-Group	27,813	17,679	57.30

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 2016 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	28,033	1.65	3.83	0.92	4.00	19.10	9.31	1.78	6.70	1.35
Pair Age Group										
12-17, 12-17	3,199	0.91	1.95	0.47	2.72	10.52	2.91	0.78	2.98	0.42
12-17, 18-25	2,548	1.45	4.13	1.27	6.00	21.72	7.91	1.06	2.89	0.41
12-17, 26-34	1,281	2.11	5.06	1.40	1.56	3.96	1.12	1.01	3.40	0.34
12-17, 35-49	5,829	1.30	3.47	1.00	1.73	8.43	1.92	0.74	2.20	0.30
12-17, 50+	1,123	1.42	3.72	0.92	1.42	7.87	2.25	0.53	2.79	0.74
18-25, 18-25	3,958	2.73	5.74	1.11	8.21	30.71	13.08	2.60	8.24	1.02
18-25, 26-34	1,429	2.52	5.66	1.46	4.20	14.63	5.51	3.85	8.74	1.52
18-25, 35-49	2,013	1.94	3.57	0.76	6.76	19.99	6.53	4.22	6.56	0.85
18-25, 50+	1,283	1.40	2.90	0.57	4.13	18.01	6.17	1.17	3.81	0.68
26-34, 26-34	1,518	1.65	3.45	0.68	1.58	9.92	5.06	1.12	2.94	0.34
26-34, 35-49	788	1.78	4.33	1.03	3.05	8.50	2.20	4.06	8.99	1.55
26-34, 50+	443	0.68	1.99	0.59	2.48	19.99	12.20	1.58	8.50	1.65
35-49, 35-49	1,213	1.15	2.42	0.60	2.89	23.02	14.99	1.81	3.20	0.48
35-49, 50+	490	1.43	3.81	0.89	4.69	31.26	18.20	2.45	11.03	3.06
50+, 50+	918	1.53	3.07	0.39	5.77	27.14	16.01	4.14	13.04	3.13
Pair Race/Ethnicity										
Hispanic or Latino	4,909	2.55	7.11	2.49	3.52	19.19	8.78	1.75	7.64	1.93
Black or African American	2,845	2.28	3.93	0.59	4.92	17.34	6.44	2.07	6.91	1.06
White	15,501	0.67	1.15	0.14	3.48	18.89	10.04	1.10	4.92	0.99
Other	2,054	4.04	9.16	1.96	6.52	20.12	8.06	4.24	11.43	2.17
White & Black or African American	273	2.20	2.35	0.19	6.23	28.54	17.18	4.40	4.98	0.80
White & Hispanic or Latino	1,133	3.00	6.50	1.21	3.97	18.19	9.10	2.47	10.39	1.84
White & Other	864	1.97	4.39	0.66	4.28	17.91	9.04	3.13	9.13	2.34
Black or African American & Hispanic or Latino	139	11.51	26.19	7.89	12.23	31.91	8.46	13.67	21.57	3.99
Black or African American & Other	141	1.42	4.10	1.64	4.26	21.13	8.36	4.96	10.13	1.50
Hispanic or Latino & Other	174	6.32	17.39	3.25	7.47	38.84	7.35	2.87	19.88	1.41
Pair Gender										
Male, Male	5,972	1.74	4.27	1.02	5.19	16.02	5.19	2.43	5.12	0.61
Female, Female	6,088	1.72	4.02	0.92	4.32	19.03	8.82	1.56	5.34	0.71
Male, Female	15,973	1.59	3.58	0.88	3.43	19.94	10.57	1.63	7.53	1.75
Household Size										
Two	7,064	1.56	3.29	0.62	0.74	1.24	0.21	0.52	1.14	0.18
Three	8,816	1.55	3.84	1.03	1.61	24.25	16.68	1.55	5.96	1.10
Four or More	12,153	1.78	4.10	1.00	7.63	25.24	9.82	2.68	9.77	2.05

Table J.1 2016 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	5,368	1.90	5.69	1.81	4.36	18.79	8.82	1.49	7.42	1.32
South	9,175	1.65	3.72	0.79	3.74	17.39	7.89	1.19	4.71	0.95
Midwest	6,624	1.39	2.58	0.47	4.06	23.61	14.39	2.55	6.92	1.24
West	6,866	1.72	3.84	0.92	4.01	18.03	7.46	2.07	8.81	2.02
Quarter										
Quarter 1	6,606	1.82	4.36	1.10	4.86	22.61	12.62	2.03	7.67	1.57
Quarter 2	7,211	1.30	2.98	0.65	3.29	17.54	9.31	1.33	6.03	1.44
Quarter 3	7,183	1.64	3.66	0.97	3.66	18.23	7.98	1.70	6.56	1.04
Quarter 4	7,033	1.86	4.29	0.96	4.27	17.91	7.21	2.10	6.53	1.37
% Hispanic or Latino in Segment										
50-100%	2,667	1.54	4.19	1.41	3.37	19.22	10.33	1.61	6.12	1.29
10-<50%	7,316	2.04	5.41	1.42	4.32	17.91	6.76	2.01	7.98	1.81
<10%	18,050	1.51	2.88	0.54	3.96	19.67	10.40	1.72	6.13	1.12
% Black or African American in Segment										
50-100%	1,902	2.21	4.58	1.15	4.63	18.66	7.58	2.84	8.16	1.19
10-<50%	5,436	2.32	5.59	1.56	4.49	19.00	7.76	1.77	5.26	0.96
<10%	20,695	1.43	3.19	0.69	3.81	19.17	9.91	1.69	6.97	1.48
% Owner-Occupied DUs[1] in Segment										
50-100%	20,952	1.29	2.92	0.62	3.60	19.89	10.22	1.87	6.86	1.47
10-<50%	5,589	2.24	5.31	1.54	4.99	16.24	6.11	1.88	6.43	0.94
<10%	1,492	4.56	10.24	2.56	5.90	16.91	6.33	0.27	2.18	0.42
Combined Median Rent/Housing Value										
1st Quintile	4,349	1.29	3.50	0.77	3.24	19.02	10.17	1.77	5.79	0.95
2nd Quintile	6,323	1.61	3.35	0.86	3.95	19.38	9.57	1.57	5.88	1.33
3rd Quintile	6,515	1.75	3.83	0.98	4.02	14.80	5.71	1.64	4.54	0.65
4th Quintile	6,310	1.68	4.36	1.18	4.45	20.28	8.86	1.87	9.29	2.10
5th Quintile	4,536	1.87	3.91	0.68	4.12	22.43	13.26	2.18	7.54	1.56
Population Density										
Large MSA[1]	12,392	1.74	4.20	1.11	4.20	21.07	10.75	1.90	7.61	1.56
Medium to Small MSA[1]	13,515	1.66	3.50	0.73	3.94	16.49	7.36	1.70	5.52	1.07
Non-MSA,[1] Urban	652	1.53	1.88	0.28	3.53	12.52	5.15	1.99	2.84	0.32
Non-MSA,[1] Rural	1,474	0.81	2.40	0.28	3.12	16.47	7.72	1.49	6.32	1.48
Group Quarters										
Group	220	3.64	7.97	1.58	11.36	40.65	12.77	6.36	22.92	3.07
Non-Group	27,813	1.64	3.80	0.92	3.94	19.04	9.30	1.75	6.65	1.35

[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 2016 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	17,847	1.77	6.71	1.41	1.64	6.99	1.06
Pair Age Group							
12-17, 12-17	2,393	0.71	2.19	0.28	0.29	1.81	0.14
12-17, 18-25	1,770	1.24	3.14	0.36	1.13	3.91	0.51
12-17, 26-34	902	1.33	4.06	0.44	0.33	0.38	0.08
12-17, 35-49	3,929	0.61	1.86	0.22	0.59	2.56	0.41
12-17, 50+	711	0.56	4.21	1.21	0.14	0.56	0.21
18-25, 18-25	2,560	2.77	8.49	1.09	2.89	11.63	1.71
18-25, 26-34	856	4.56	12.60	2.15	7.01	13.35	2.46
18-25, 35-49	1,153	4.60	8.54	1.07	4.25	10.44	2.06
18-25, 50+	710	1.27	3.36	0.37	1.55	4.64	1.12
26-34, 26-34	906	0.88	2.58	0.30	0.55	1.83	0.33
26-34, 35-49	426	5.16	11.81	1.66	0.70	0.80	0.06
26-34, 50+	224	1.79	11.95	4.00	2.23	11.75	1.14
35-49, 35-49	625	1.60	2.27	0.40	1.28	1.77	0.34
35-49, 50+	218	2.29	10.78	1.96	2.75	11.94	2.20
50+, 50+	464	3.45	12.36	3.60	3.88	12.91	1.48
Pair Race/Ethnicity							
Hispanic or Latino	3,253	1.78	10.43	2.52	1.35	8.12	1.34
Black or African American	2,070	1.88	5.17	0.98	0.63	2.26	0.21
White	9,534	1.15	4.74	0.94	1.26	5.90	0.82
Other	1,208	3.56	9.24	1.63	5.96	17.71	2.99
White & Black or African American	176	3.98	5.28	0.68	1.14	0.83	0.09
White & Hispanic or Latino	742	2.02	7.04	1.19	1.75	5.90	0.93
White & Other	559	4.11	11.79	3.42	1.61	7.58	0.93
Black or African American & Hispanic or Latino	99	12.12	24.33	4.68	5.05	9.36	2.14
Black or African American & Other	97	6.19	12.32	2.10	2.06	6.31	1.13
Hispanic or Latino & Other	109	2.75	9.33	0.67	11.93	26.45	5.15
Pair Gender							
Male, Male	3,721	2.55	5.44	0.75	2.15	6.53	1.10
Female, Female	4,157	1.59	5.06	0.76	1.37	6.29	1.46
Male, Female	9,969	1.55	7.60	1.81	1.56	7.31	0.93
Household Size							
Two	4,221	0.59	0.98	0.15	0.43	0.66	0.13
Three	5,609	1.55	7.59	1.56	1.32	8.30	1.47
Four or More	8,017	2.54	8.84	1.90	2.51	9.39	1.30

Table J.2 2016 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	3,159	1.49	6.58	1.49	1.49	7.19	1.20
South	6,023	1.26	4.39	0.61	1.15	6.04	1.06
Midwest	4,255	2.30	6.16	1.17	1.90	5.47	0.86
West	4,410	2.15	10.82	2.79	2.18	9.29	1.10
Quarter							
Quarter 1	4,317	2.29	8.02	1.61	1.90	7.64	1.32
Quarter 2	4,500	1.11	5.54	1.34	1.53	6.77	1.02
Quarter 3	4,575	1.66	7.18	1.48	1.42	6.27	0.80
Quarter 4	4,455	2.04	6.00	1.20	1.73	7.27	1.09
% Hispanic or Latino in Segment							
50-100%	1,751	1.94	9.73	1.93	2.23	10.50	1.82
10-<50%	4,646	2.02	7.94	1.91	2.09	7.84	0.97
<10%	11,450	1.64	5.44	1.04	1.37	5.80	0.95
% Black or African American in Segment							
50-100%	1,366	2.12	4.41	0.67	1.39	4.39	0.55
10-<50%	3,522	1.76	5.63	0.95	1.82	5.08	0.77
<10%	12,959	1.74	7.27	1.62	1.62	7.78	1.19
% Owner-Occupied DUs[1] in Segment							
50-100%	13,239	1.82	7.11	1.55	1.64	7.21	1.01
10-<50%	3,609	1.97	5.45	0.90	2.02	6.39	1.26
<10%	999	0.40	2.78	0.78	0.30	2.79	1.01
Combined Median Rent/Housing Value							
1st Quintile	2,940	1.84	6.15	0.82	1.02	4.91	0.83
2nd Quintile	4,202	1.48	6.95	1.68	1.31	5.67	0.58
3rd Quintile	4,164	1.99	5.03	0.64	1.63	5.34	0.87
4th Quintile	3,855	1.61	7.20	1.94	2.33	9.32	1.22
5th Quintile	2,686	2.05	8.45	1.91	1.86	8.83	1.73
Population Density							
Large MSA[1]	7,632	1.87	7.49	1.71	2.08	8.64	1.37
Medium to Small MSA[1]	8,839	1.65	5.98	1.11	1.29	4.98	0.66
Non-MSA,[1] Urban	437	2.06	1.57	0.37	1.83	2.63	0.28
Non-MSA,[1] Rural	939	1.92	5.42	0.72	1.28	3.62	0.56
Group Quarters							
Group	168	5.36	20.56	2.11	2.38	8.03	1.10
Non-Group	17,679	1.74	6.66	1.41	1.63	6.98	1.06

[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 2016 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	17,847	1.61	4.91	0.85	0.59	1.71	0.15	0.03	0.14	0.06
Pair Age Group										
12-17, 12-17	2,386	0.29	1.82	0.14	0.08	0.39	0.00	0.00	0.00	0.00
12-17, 18-25	1,774	1.13	4.22	0.54	0.23	1.10	0.07	0.00	0.00	0.00
12-17, 26-34	883	0.34	0.40	0.09	0.11	0.13	0.02	0.00	0.00	0.00
12-17, 35-49	3,930	0.64	2.59	0.49	0.33	0.80	0.02	0.03	0.01	0.00
12-17, 50+	725	0.14	0.56	0.20	0.00	0.00	0.00	0.00	0.00	0.00
18-25, 18-25	2,512	2.99	12.04	2.04	1.00	3.50	0.19	0.00	0.00	0.00
18-25, 26-34	886	6.77	14.18	3.42	1.81	5.42	0.85	0.23	1.99	0.63
18-25, 35-49	1,134	4.85	11.23	2.25	1.59	3.05	0.21	0.09	0.01	0.00
18-25, 50+	722	1.52	5.12	1.55	0.69	2.33	0.64	0.28	0.85	0.46
26-34, 26-34	905	0.55	4.20	1.15	0.77	3.67	0.10	0.00	0.00	0.00
26-34, 35-49	445	0.90	2.36	0.44	1.80	2.53	0.13	0.00	0.00	0.00
26-34, 50+	234	0.43	2.18	0.22	0.00	0.00	0.00	0.00	0.00	0.00
35-49, 35-49	627	1.44	2.93	0.40	0.48	0.96	0.05	0.00	0.00	0.00
35-49, 50+	213	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
50+, 50+	471	2.34	7.18	0.67	0.85	2.28	0.10	0.00	0.00	0.00
Pair Race/Ethnicity										
Hispanic or Latino	3,272	1.25	3.93	0.84	0.37	0.75	0.04	0.00	0.00	0.00
Black or African American	2,005	0.90	2.96	0.41	0.70	2.52	0.13	0.05	0.24	0.05
White	9,262	1.18	3.51	0.57	0.51	1.56	0.16	0.02	0.18	0.10
Other	1,157	5.79	17.11	3.49	2.68	5.77	0.49	0.26	0.18	0.00
White & Black or African American	180	1.67	7.33	1.68	0.00	0.00	0.00	0.00	0.00	0.00
White & Hispanic or Latino	752	2.39	7.32	0.62	0.13	0.12	0.02	0.00	0.00	0.00
White & Other	749	1.47	1.79	0.10	0.00	0.00	0.00	0.00	0.00	0.00
Black or African American & Hispanic or Latino	144	3.47	10.08	1.50	0.00	0.00	0.00	0.00	0.00	0.00
Black or African American & Other	172	2.91	5.59	1.10	0.00	0.00	0.00	0.00	0.00	0.00
Hispanic or Latino & Other	154	6.49	17.18	3.02	0.65	1.21	0.00	0.00	0.00	0.00
Pair Gender										
Male, Male	3,714	2.29	8.01	1.43	0.75	2.38	0.20	0.05	0.27	0.08
Female, Female	4,148	1.37	5.44	1.16	0.68	2.47	0.15	0.02	0.00	0.00
Male, Female	9,985	1.45	3.88	0.60	0.50	1.31	0.13	0.03	0.14	0.07
Household Size										
Two	4,221	0.43	0.65	0.13	0.17	0.29	0.02	0.02	0.00	0.00
Three	5,609	1.23	5.48	0.86	0.53	1.65	0.09	0.02	0.10	0.02
Four or More	8,017	2.49	6.69	1.20	0.86	2.43	0.24	0.05	0.22	0.11

Table J.3 2016 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	3,159	1.36	3.27	0.56	0.54	1.39	0.17	0.03	0.20	0.08
South	6,023	1.16	4.15	0.80	0.30	1.05	0.03	0.00	0.00	0.00
Midwest	4,255	1.76	4.04	0.63	1.13	2.13	0.12	0.07	0.07	0.00
West	4,410	2.24	7.74	1.28	0.52	2.54	0.30	0.05	0.34	0.17
Quarter										
Quarter 1	4,317	1.92	4.57	0.90	0.74	1.50	0.08	0.00	0.00	0.00
Quarter 2	4,500	1.53	5.27	0.79	0.51	1.74	0.29	0.09	0.50	0.24
Quarter 3	4,575	1.36	3.79	0.55	0.48	1.00	0.05	0.00	0.00	0.00
Quarter 4	4,455	1.64	6.01	1.17	0.65	2.61	0.16	0.04	0.06	0.00
% Hispanic or Latino in Segment										
50-100%	1,751	2.17	8.42	1.34	0.63	1.14	0.06	0.00	0.00	0.00
10-<50%	4,646	2.17	6.86	1.26	0.56	2.47	0.31	0.06	0.40	0.20
<10%	11,450	1.29	3.14	0.53	0.60	1.43	0.08	0.03	0.03	0.00
% Black or African American in Segment										
50-100%	1,366	1.61	4.66	0.74	0.59	0.98	0.04	0.00	0.00	0.00
10-<50%	3,522	1.96	6.03	0.77	0.77	2.22	0.12	0.03	0.13	0.02
<10%	12,959	1.51	4.61	0.89	0.55	1.64	0.16	0.04	0.16	0.08
% Owner-Occupied DUs[1] in Segment										
50-100%	13,239	1.54	4.58	0.71	0.53	1.45	0.13	0.02	0.10	0.05
10-<50%	3,609	2.19	6.43	1.43	0.91	2.85	0.23	0.06	0.33	0.10
<10%	999	0.40	3.11	0.98	0.30	1.18	0.11	0.10	0.04	0.00
Combined Median Rent/Housing Value										
1st Quintile	2,940	1.12	2.81	0.35	0.78	1.26	0.11	0.03	0.21	0.04
2nd Quintile	4,202	1.14	2.59	0.55	0.31	1.16	0.24	0.02	0.30	0.20
3rd Quintile	4,164	1.59	4.13	0.69	0.65	1.85	0.09	0.02	0.00	0.00
4th Quintile	3,855	2.36	8.40	1.31	0.80	2.89	0.20	0.03	0.15	0.06
5th Quintile	2,686	1.82	5.33	1.14	0.45	0.95	0.06	0.07	0.07	0.00
Population Density										
Large MSA[1]	7,632	2.04	6.40	1.18	0.68	2.12	0.21	0.05	0.24	0.11
Medium to Small MSA[1]	8,839	1.23	2.87	0.40	0.52	1.03	0.07	0.02	0.00	0.00
Non-MSA,[1] Urban	437	1.83	1.85	0.21	0.46	0.17	0.01	0.00	0.00	0.00
Non-MSA,[1] Rural	939	1.49	4.05	0.76	0.64	3.22	0.07	0.00	0.00	0.00
Group Quarters										
Group	168	2.98	8.69	1.27	0.00	0.00	0.00	0.00	0.00	0.00
Non-Group	17,679	1.60	4.90	0.85	0.60	1.72	0.15	0.03	0.14	0.06

Table J.3 2016 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)	2,749	0.33	0.79	0.19	0.36	0.85	0.03	0.00	0.00	0.00
Parent-Child (12-17)	5,048	0.48	1.70	0.30	0.28	0.55	0.02	0.02	0.00	0.00
Parent-Child (12-20)	5,846	0.94	3.25	0.62	0.46	0.97	0.06	0.03	0.01	0.00
Sibling (12-14)-Sibling (15-17)	1,421	0.28	2.06	0.17	0.00	0.00	0.00	0.00	0.00	0.00
Sibling (12-17)-Sibling (18-25)	1,626	0.92	3.45	0.44	0.25	1.19	0.08	0.00	0.00	0.00
Spouse-Spouse/Partner-Partner	3,363	1.16	3.63	0.52	0.54	1.33	0.08	0.00	0.00	0.00
Spouse-Spouse/Partner-Partner with Children (Younger Than 18)	1,727	0.98	2.31	0.42	0.75	1.04	0.05	0.00	0.00	0.00

[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 2016 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	17,847	230,864,859	230,864,859	230,864,859	-0.00	-0.00
Pair Age Group						
12-17, 12-17	2,386	7,336,452	7,341,504	7,341,504	-0.07	-0.00
12-17, 18-25	1,774	8,065,065	7,998,777	7,998,777	0.83	-0.00
12-17, 26-34	883	4,623,212	4,783,252	4,783,252	-3.35	-0.00
12-17, 35-49	3,930	29,397,023	29,253,475	29,253,475	0.49	-0.00
12-17, 50+	725	13,040,321	12,965,562	12,965,562	0.58	-0.00
18-25, 18-25	2,512	12,795,161	12,908,047	12,908,047	-0.87	-0.00
18-25, 26-34	886	7,093,411	7,246,957	7,246,957	-2.12	-0.00
18-25, 35-49	1,134	16,495,107	16,646,249	16,646,249	-0.91	-0.00
18-25, 50+	722	20,328,920	20,305,643	20,305,643	0.11	-0.00
26-34, 26-34	905	10,985,093	11,142,380	11,142,380	-1.41	-0.00
26-34, 35-49	445	9,026,624	9,063,688	9,063,688	-0.41	0.00
26-34, 50+	234	13,529,828	13,329,300	13,329,300	1.50	0.00
35-49, 35-49	627	18,453,002	18,245,751	18,245,751	1.14	0.00
35-49, 50+	213	17,144,310	17,597,338	17,597,338	-2.57	0.00
50+, 50+	471	42,551,329	42,036,936	42,036,936	1.22	-0.00
Pair Race/Ethnicity						
Hispanic or Latino	3,272	41,952,526	41,873,987	41,873,987	0.19	0.00
Black or African American	2,005	23,546,680	24,745,828	24,745,828	-4.85	-0.00
White	9,262	119,112,601	123,921,085	123,921,085	-3.88	0.00
Other	1,157	17,807,097	18,181,994	18,181,994	-2.06	-0.00
White & Black or African American	180	2,188,193	2,420,405	2,420,405	-9.59	0.00
White & Hispanic or Latino	752	10,356,957	9,378,332	9,378,332	10.43	0.00
White & Other	749	9,848,723	6,484,680	6,484,680	51.88	0.00
Black or African American & Hispanic or Latino	144	1,634,895	1,435,435	1,435,435	13.90	0.00
Black or African American & Other	172	2,200,077	1,027,530	1,027,530	114.11	0.00
Hispanic or Latino & Other	154	2,217,110	1,395,583	1,395,583	58.87	-0.00
Pair Gender						
Male, Male	3,714	41,616,672	41,631,043	41,631,043	-0.03	-0.00
Female, Female	4,148	41,986,396	42,104,835	42,104,835	-0.28	-0.00
Male, Female	9,985	147,261,791	147,128,981	147,128,981	0.09	0.00
Pair Relationship Domain[3,4,5]						
Parent-Child (12-14)*	2,749	11,633,859	12,615,724	12,615,724	-7.78	-0.00
Parent-Child (12-17)*	5,048	23,763,084	25,365,956	25,365,956	-6.32	-0.00
Parent-Child (15-17)*	2,299	12,129,225	12,750,231	12,750,231	-4.87	-0.00
Parent-Child (12-20)*	5,846	32,424,952	34,511,441	34,511,441	-6.05	-0.00
Parent*-Child (12-14)	2,749	17,717,120	19,339,786	19,339,786	-8.39	-0.00
Parent*-Child (12-17)	5,048	30,265,674	32,414,114	32,414,114	-6.63	-0.00
Parent*-Child (15-17)	2,299	18,806,509	19,404,038	19,632,355	-4.21	-1.16
Parent*-Child (12-20)	5,846	37,818,927	40,338,441	40,338,441	-6.25	-0.00
Sibling (12-14)-Sibling (15-17)*	1,421	3,913,308	4,202,039	4,202,039	-6.87	-0.00
Sibling (12-17)-Sibling (18-25)*	1,626	6,078,816	6,220,297	6,220,297	-2.27	-0.00
Spouse-Spouse/Partner-Partner	3,363	73,276,172	73,908,755	73,908,755	-0.86	0.00
Spouse-Spouse/Partner-Partner with Children (Younger Than 18)	1,727	27,636,272	29,656,613	29,656,613	-6.81	0.00

Table K.1 2016 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	4,221	56,243,224	56,243,224	56,243,224	-0.00	-0.00
Three	5,609	58,720,399	58,720,399	58,720,399	-0.00	-0.00
Four or More	8,017	115,901,236	115,901,236	115,901,236	-0.00	-0.00
Census Region						
Northeast	3,159	41,273,152	41,273,152	41,273,152	-0.00	-0.00
South	6,023	84,030,953	84,030,953	84,030,953	-0.00	-0.00
Midwest	4,255	45,228,444	45,228,444	45,228,444	0.00	-0.00
West	4,410	60,332,310	60,332,310	60,332,310	0.00	0.00
Quarter						
Quarter 1	4,317	57,623,403	57,623,403	57,623,403	-0.00	-0.00
Quarter 2	4,500	57,266,510	57,266,510	57,266,510	-0.00	-0.00
Quarter 3	4,575	58,120,170	58,120,170	58,120,170	-0.00	-0.00
Quarter 4	4,455	57,854,775	57,854,775	57,854,776	-0.00	-0.00
% Hispanic or Latino in Segment						
50-100%	1,751	27,374,561	27,374,561	27,374,561	-0.00	-0.00
10-<50%	4,646	70,871,800	70,871,800	70,871,800	-0.00	-0.00
<10%	11,450	132,618,498	132,618,498	132,618,498	-0.00	-0.00
% Black or African American in Segment						
50-100%	1,366	15,919,008	15,919,008	15,919,008	0.00	-0.00
10-<50%	3,522	47,916,948	47,916,948	47,916,948	-0.00	-0.00
<10%	12,959	167,028,903	167,028,903	167,028,903	-0.00	-0.00
% Owner-Occupied DUs in Segment						
50-100%	13,239	183,235,848	183,235,848	183,235,848	-0.00	-0.00
10-<50%	3,609	43,800,108	43,800,108	43,800,108	-0.00	-0.00
<10%	999	3,828,903	3,828,903	3,828,903	-0.00	-0.00
Combined Median Rent/Housing Value						
1st Quintile	2,940	29,361,028	29,361,028	29,361,028	0.00	-0.00
2nd Quintile	4,202	47,274,595	47,274,595	47,274,595	-0.00	-0.00
3rd Quintile	4,164	54,079,588	54,079,588	54,079,588	-0.00	-0.00
4th Quintile	3,855	55,793,549	55,793,549	55,793,549	-0.00	-0.00
5th Quintile	2,686	44,356,099	44,356,099	44,356,099	-0.00	-0.00
Population Density						
Large MSA	7,632	131,701,869	131,701,869	131,701,869	-0.00	-0.00
Medium to Small MSA	8,839	87,514,614	87,514,614	87,514,614	-0.00	-0.00
Non-MSA, Urban	437	3,681,367	3,681,367	3,681,367	0.00	-0.00
Non-MSA, Rural	939	7,967,009	7,967,009	7,967,009	0.00	-0.00
Group Quarters						
Group	168	653,710	653,710	653,710	0.00	-0.00
Non-Group	17,679	230,211,149	230,211,149	230,211,149	-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 2016 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	28,033	15	546	942	1,326	8,510	1.41	17	1,724	3,657	8,227	2,116,210	11.73	3	1,627	3,675	8,420	301,034	4.28
Pair Age Group																			
12-17, 12-17	3,199	15	453	829	1,207	6,487	1.45	17	925	1,618	2,790	29,859	2.09	3	552	1,453	3,001	21,497	2.30
12-17, 18-25	2,548	41	557	969	1,354	5,787	1.43	76	1,180	2,032	3,749	68,289	2.42	53	1,146	2,156	4,102	21,243	1.91
12-17, 26-34	1,281	37	519	874	1,284	5,665	1.44	114	1,526	2,647	4,270	34,597	2.06	105	1,434	2,566	4,414	53,454	2.25
12-17, 35-49	5,829	27	506	871	1,216	7,524	1.42	105	1,958	3,464	5,942	64,016	2.22	64	1,725	3,241	6,126	66,195	2.23
12-17, 50+	1,123	39	584	970	1,405	5,596	1.39	554	5,171	8,483	12,689	204,337	2.24	440	4,918	8,608	13,831	139,536	2.01
18-25, 18-25	3,958	22	464	961	1,360	8,339	1.49	53	1,088	2,090	3,630	274,875	4.31	24	919	2,102	4,289	27,016	2.11
18-25, 26-34	1,429	57	558	998	1,407	8,510	1.43	231	2,059	3,643	5,739	250,986	3.88	117	1,735	3,410	6,020	77,779	2.36
18-25, 35-49	2,013	36	636	987	1,357	4,856	1.34	117	3,020	5,283	9,055	179,164	2.41	111	3,004	5,608	10,228	55,521	2.01
18-25, 50+	1,283	70	773	1,127	1,552	5,433	1.31	628	7,445	11,666	18,015	339,865	2.50	466	7,267	12,014	19,942	125,111	1.77
26-34, 26-34	1,518	44	617	974	1,327	6,896	1.34	252	3,738	6,399	9,294	351,341	3.66	312	2,884	5,143	8,475	121,164	2.43
26-34, 35-49	788	43	600	985	1,307	4,721	1.40	560	4,743	7,924	11,511	115,521	2.20	535	4,722	7,971	13,146	159,422	2.49
26-34, 50+	443	75	619	1,104	1,503	4,547	1.33	1,592	13,654	22,686	32,615	1,115,479	5.11	1,176	12,116	23,614	39,120	301,034	1.93
35-49, 35-49	1,213	55	531	925	1,287	5,405	1.39	432	4,657	8,846	13,749	481,715	6.03	279	4,233	8,245	15,205	175,256	3.02
35-49, 50+	490	64	635	993	1,343	4,187	1.36	1,518	12,710	20,989	32,588	1,298,482	6.61	1,695	14,337	26,058	43,991	212,562	1.94
50+, 50+	918	82	686	1,029	1,415	3,242	1.30	2,797	24,480	38,627	51,923	2,116,210	5.53	3,078	23,130	38,700	57,558	258,468	1.60
Pair Race/Ethnicity																			
Hispanic or Latino	4,909	18	619	994	1,468	8,510	1.43	20	1,995	4,046	8,434	1,115,479	8.47	4	1,747	3,955	8,981	301,034	4.33
Black or African American	2,845	38	799	1,083	1,416	4,502	1.26	72	2,188	4,116	8,463	406,796	4.79	14	1,962	4,386	9,256	157,307	3.55
White	15,501	36	488	913	1,225	4,965	1.38	66	1,620	3,448	8,051	2,116,210	15.25	35	1,589	3,457	7,922	258,468	4.42
Other	2,054	15	278	863	1,551	6,985	1.64	17	1,283	3,391	8,499	407,753	5.23	3	1,238	3,580	9,111	209,426	4.40
White & Black or African American	273	50	683	1,000	1,402	3,363	1.31	62	1,950	4,059	8,206	541,423	14.73	24	1,910	4,437	9,676	114,806	3.52
White & Hispanic or Latino	1,133	40	611	996	1,418	5,665	1.41	105	1,940	4,194	9,459	363,734	6.49	54	1,453	3,836	8,872	195,743	4.23
White & Other	864	31	268	736	1,190	4,620	1.60	97	1,378	2,968	6,743	274,875	5.37	32	1,202	3,284	8,331	247,868	4.34
Black or African American & Hispanic or Latino	139	47	632	1,062	1,633	6,062	1.62	178	2,060	4,238	7,488	111,802	3.99	200	2,550	5,154	10,611	116,233	3.65
Black or African American & Other	141	43	681	946	1,326	5,346	1.39	53	2,333	3,728	7,198	115,521	4.11	28	2,373	4,327	8,687	74,144	2.55
Hispanic or Latino & Other	174	39	311	777	1,343	4,547	1.68	113	1,217	2,974	6,492	198,470	7.36	64	1,489	3,908	7,775	179,653	5.32

Table L.1 2016 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	5,972	24	491	933	1,330	8,339	1.46	53	1,552	3,292	7,461	232,677	3.70	24	1,502	3,457	8,003	152,944	3.37
Female, Female	6,088	22	571	945	1,350	6,985	1.40	96	1,707	3,471	7,471	2,116,210	17.65	32	1,501	3,399	7,508	171,121	3.79
Male, Female	15,973	15	556	945	1,315	8,510	1.40	17	1,793	3,892	8,788	2,116,026	11.57	3	1,722	3,865	8,962	301,034	4.47
Household Size																			
Two	7,064	43	537	929	1,271	6,062	1.39	53	1,787	3,991	8,677	172,363	3.12	24	1,445	3,410	8,089	123,071	3.47
Three	8,816	15	533	919	1,295	8,510	1.43	17	1,498	2,983	5,618	2,116,210	32.89	3	1,467	3,067	6,327	247,868	5.10
Four or More	12,153	22	561	969	1,385	8,339	1.42	59	1,863	4,332	9,724	1,115,479	6.27	12	1,875	4,598	10,245	301,034	4.16
Census Region																			
Northeast	5,368	35	326	845	1,105	8,510	1.47	68	1,439	3,081	7,697	906,065	8.93	27	1,302	3,294	7,914	247,868	4.65
South	9,175	15	760	1,075	1,528	8,339	1.33	17	2,257	4,355	9,143	1,298,482	8.18	3	2,131	4,387	9,317	212,562	3.91
Midwest	6,624	27	599	876	1,122	4,475	1.26	108	1,619	3,137	6,857	2,116,210	29.75	108	1,611	3,090	7,076	205,424	4.00
West	6,866	29	295	909	1,502	6,896	1.55	53	1,409	3,679	8,954	561,680	5.60	24	1,259	3,687	9,361	301,034	4.56
Quarter																			
Quarter1	6,606	18	577	981	1,422	8,510	1.42	20	1,872	3,880	8,500	2,116,026	16.39	4	1,796	3,969	8,643	258,468	4.32
Quarter2	7,211	27	530	925	1,301	8,339	1.39	53	1,648	3,505	8,067	2,116,210	15.40	20	1,535	3,527	8,183	301,034	4.34
Quarter3	7,183	22	547	905	1,282	7,524	1.39	70	1,710	3,579	8,079	1,115,479	8.02	19	1,589	3,585	8,315	247,868	4.26
Quarter4	7,033	15	533	952	1,327	6,774	1.44	17	1,679	3,650	8,274	561,680	6.15	3	1,614	3,725	8,581	246,640	4.19
% Hispanic or Latino in Segment																			
50-100%	2,667	52	737	1,179	1,547	8,339	1.29	146	2,536	4,790	10,302	1,115,479	9.39	40	2,133	4,863	11,008	246,640	3.85
10-<50%	7,316	18	715	1,080	1,576	7,524	1.38	20	2,246	4,526	9,857	428,609	4.84	4	2,054	4,497	9,998	301,034	4.10
<10%	18,050	15	386	870	1,183	8,510	1.41	17	1,489	3,137	7,312	2,116,210	16.37	3	1,434	3,238	7,508	247,868	4.36
% Black or African American in Segment																			
50-100%	1,902	15	768	1,044	1,355	8,510	1.29	17	2,101	4,059	8,432	541,423	6.17	3	1,925	4,175	9,231	157,307	3.50
10-<50%	5,436	18	745	1,039	1,447	7,524	1.33	20	2,210	4,234	8,891	1,298,482	8.05	4	1,890	4,241	9,091	207,566	3.82
<10%	20,695	27	427	899	1,285	8,339	1.45	53	1,574	3,470	8,001	2,116,210	13.34	24	1,522	3,514	8,117	301,034	4.50
% Owner-Occupied DUs[1] in Segment																			
50-100%	20,952	24	531	929	1,308	7,524	1.40	59	1,726	3,694	8,341	2,116,210	13.34	12	1,760	3,907	8,885	301,034	4.23
10-<50%	5,589	15	602	984	1,357	8,510	1.42	17	1,807	3,727	8,268	541,423	4.90	3	1,710	3,770	8,356	181,853	3.86
<10%	1,492	43	519	962	1,403	6,062	1.51	53	1,407	3,000	6,356	229,027	4.61	14	535	1,226	2,885	55,853	3.84
Combined Median Rent/Housing Value																			
1st Quintile	4,349	29	363	779	1,107	5,697	1.46	105	1,320	2,857	6,478	1,298,482	16.55	47	1,286	2,997	7,014	212,562	4.40
2nd Quintile	6,323	37	475	902	1,247	6,487	1.41	53	1,574	3,344	7,458	1,115,479	9.85	24	1,379	3,323	7,534	246,640	4.49
3rd Quintile	6,515	27	554	960	1,356	8,510	1.42	72	1,754	3,739	8,309	481,715	5.20	14	1,580	3,637	8,366	195,743	3.97
4th Quintile	6,310	15	617	962	1,386	7,524	1.41	17	1,879	4,014	8,801	428,609	5.82	3	1,794	3,974	8,962	301,034	4.49
5th Quintile	4,536	18	738	1,092	1,494	5,412	1.33	20	2,115	4,580	9,743	2,116,210	23.05	4	2,245	4,765	10,350	258,468	3.89

Table L.1 2016 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]	12,392	15	870	1,153	1,561	8,510	1.26	17	2,698	5,198	10,759	2,116,210	11.30	3	2,531	5,181	11,070	301,034	3.70
Medium to Small MSA[1]	13,515	27	316	761	1,101	6,985	1.50	53	1,248	2,759	6,158	1,298,482	10.07	24	1,178	2,808	6,565	247,868	4.65
Non-MSA,[1] Urban	652	29	261	704	1,060	2,926	1.49	104	1,102	2,399	5,200	204,337	5.51	96	1,159	2,681	5,693	105,272	3.97
Non-MSA,[1] Rural	1,474	43	199	569	932	4,477	1.61	125	957	2,184	5,125	351,341	7.91	70	896	2,093	5,039	212,562	5.73
Group Quarters																			
Group	220	43	184	396	1,328	3,997	2.03	53	553	1,018	3,745	28,847	3.12	24	301	834	4,207	18,510	2.96
Non-Group	27,813	15	554	944	1,326	8,510	1.41	17	1,738	3,681	8,249	2,116,210	11.70	3	1,646	3,699	8,445	301,034	4.27

[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 $1 + \left[(n-1)/n\right]^{*}CV^{2}$, where CV = coefficient of variation of weights.

Table L.2 2016 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	17,847	3	1,518	3,400	7,704	301,034	4.43	4	2,038	4,912	12,016	471,208	5.48
Pair Age Group													
12-17, 12-17	2,393	3	550	1,459	3,032	21,497	2.31	4	707	1,826	3,985	29,858	2.33
12-17, 18-25	1,770	53	1,143	2,184	4,007	19,604	1.93	61	1,556	3,016	5,899	29,853	1.97
12-17, 26-34	902	105	1,445	2,580	4,573	53,454	2.31	113	1,827	3,326	6,086	65,888	2.46
12-17, 35-49	3,929	64	1,709	3,241	6,103	66,195	2.25	68	2,211	4,443	8,956	91,944	2.45
12-17, 50+	711	511	4,573	8,283	13,341	139,536	2.12	600	6,160	12,560	21,637	139,306	2.12
18-25, 18-25	2,560	24	880	2,092	4,349	27,016	2.15	24	1,243	2,972	6,620	57,738	2.32
18-25, 26-34	856	117	1,744	3,464	5,967	77,779	2.43	152	2,538	5,484	10,389	103,750	2.41
18-25, 35-49	1,153	111	3,095	6,024	10,795	55,521	1.98	126	4,552	9,240	18,416	123,416	2.17
18-25, 50+	710	466	7,044	12,010	19,665	95,422	1.75	813	11,040	20,342	35,951	280,488	1.93
26-34, 26-34	906	435	2,761	4,974	8,318	114,651	2.36	605	4,086	7,524	13,554	293,831	3.31
26-34, 35-49	426	535	4,476	8,047	13,736	159,422	2.33	792	8,022	14,808	25,697	226,855	2.09
26-34, 50+	224	1,176	12,524	24,187	39,079	301,034	2.12	2,735	21,558	42,250	72,813	325,194	2.01
35-49, 35-49	625	485	3,676	7,635	13,380	175,256	3.45	748	6,792	14,093	28,802	457,845	3.64
35-49, 50+	218	2,013	13,820	23,537	43,082	196,685	1.86	5,148	29,748	61,357	101,843	389,219	1.83
50+, 50+	464	3,078	24,005	39,459	59,711	258,468	1.63	4,413	40,818	75,289	114,125	471,208	1.63
Pair Race/Ethnicity													
Hispanic or Latino	3,253	4	1,665	3,707	8,426	301,034	4.58	5	2,098	5,020	12,409	389,219	5.34
Black or African American	2,070	14	1,858	4,089	8,535	157,307	3.57	16	2,285	5,436	12,159	277,291	4.28
White	9,534	35	1,465	3,179	7,269	258,468	4.61	39	2,054	4,779	11,712	471,208	5.65
Other	1,208	3	1,033	2,757	6,976	180,314	4.92	4	1,343	4,419	12,594	328,035	6.75
White & Black or African American	176	24	1,942	4,449	9,755	112,048	3.46	24	2,497	5,899	15,556	181,457	4.16
White & Hispanic or Latino	742	54	1,417	3,766	8,356	177,700	3.69	54	1,801	4,983	12,741	333,258	4.74
White & Other	559	32	1,238	3,091	7,812	247,868	5.21	32	1,518	3,980	11,331	328,371	5.64
Black or African American & Hispanic or Latino	99	200	2,608	4,915	10,512	99,949	3.38	258	3,081	7,044	14,214	137,109	3.43
Black or African American & Other	97	28	2,573	4,455	9,656	74,144	2.58	28	2,985	5,494	11,981	115,811	3.08
Hispanic or Latino & Other	109	64	1,489	3,813	7,157	43,168	2.62	68	2,655	7,204	16,351	91,944	2.53
Pair Gender													
Male, Male	3,721	24	1,417	3,236	7,379	139,924	3.32	24	1,940	4,628	11,612	427,353	4.72
Female, Female	4,157	32	1,446	3,228	7,053	171,121	3.76	32	1,873	4,349	10,291	436,797	5.25
Male, Female	9,969	3	1,602	3,543	8,119	301,034	4.74	4	2,178	5,252	13,207	471,208	5.49
Household Size													
Two	4,221	24	1,321	3,020	7,172	123,071	3.83	24	1,804	4,399	11,925	294,632	4.49
Three	5,609	3	1,394	2,920	5,685	247,868	5.57	4	1,839	4,043	8,585	436,797	7.23
Four or More	8,017	12	1,750	4,286	9,450	301,034	4.11	18	2,405	6,297	14,622	471,208	5.10

Table L.2 2016 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	3,159	27	1,132	2,923	7,056	247,868	5.14	39	1,695	4,689	12,302	457,845	6.02
South	6,023	3	2,046	4,204	8,855	189,175	3.82	4	2,701	5,963	13,259	364,376	4.77
Midwest	4,255	108	1,530	2,997	6,459	205,424	4.23	155	2,018	4,207	9,932	436,797	5.14
West	4,410	24	1,162	3,170	7,948	301,034	5.01	24	1,500	4,448	12,562	471,208	6.12
Quarter													
Quarter1	4,317	4	1,712	3,643	7,885	258,468	4.44	5	2,273	5,291	11,875	389,219	5.55
Quarter2	4,500	20	1,437	3,203	7,385	301,034	4.46	21	1,967	4,692	11,748	436,797	5.32
Quarter3	4,575	27	1,452	3,359	7,691	247,868	4.39	41	1,945	4,681	12,156	457,845	5.41
Quarter4	4,455	3	1,493	3,368	7,837	246,640	4.43	4	2,031	4,986	12,275	471,208	5.62
% Hispanic or Latino in Segment													
50-100%	1,751	53	1,967	4,293	9,855	246,640	4.09	61	2,538	6,032	15,147	427,353	5.22
10-<50%	4,646	4	1,868	4,079	8,923	301,034	4.25	5	2,549	6,021	14,467	471,208	5.10
<10%	11,450	3	1,349	3,059	7,020	247,868	4.51	4	1,839	4,377	10,872	457,845	5.61
% Black or African American in Segment													
50-100%	1,366	3	1,815	3,911	8,623	157,307	3.33	4	2,259	5,350	12,519	229,206	4.01
10-<50%	3,522	4	1,777	3,850	8,562	152,704	3.68	5	2,372	5,567	13,540	427,353	4.69
<10%	12,959	24	1,420	3,228	7,387	301,034	4.79	24	1,946	4,685	11,622	471,208	5.83
% Owner-Occupied DUs[1] in Segment													
50-100%	13,239	12	1,674	3,631	8,132	301,034	4.40	18	2,265	5,330	12,949	471,208	5.31
10-<50%	3,609	3	1,557	3,434	7,643	181,853	3.90	4	2,055	4,786	11,490	389,219	5.33
<10%	999	14	504	1,187	2,654	55,853	4.02	16	642	1,619	3,879	97,757	4.79
Combined Median Rent/Housing Value													
1st Quintile	2,940	47	1,273	2,949	6,781	169,166	4.23	49	1,712	3,991	9,605	381,195	5.55
2nd Quintile	4,202	24	1,297	3,050	7,082	246,640	4.69	24	1,610	4,036	10,356	325,194	5.67
3rd Quintile	4,164	14	1,519	3,431	7,800	159,422	3.88	16	2,051	4,924	11,927	389,219	4.92
4th Quintile	3,855	3	1,675	3,588	7,904	301,034	4.71	4	2,379	5,572	13,756	471,208	5.38
5th Quintile	2,686	4	2,109	4,362	9,170	258,468	4.42	5	2,964	6,711	15,464	457,845	5.35
Population Density													
Large MSA[1]	7,632	3	2,337	4,712	9,831	301,034	3.94	4	3,246	7,165	16,676	471,208	4.78
Medium to Small MSA[1]	8,839	24	1,133	2,704	6,165	247,868	4.67	24	1,532	3,761	9,133	381,195	5.64
Non-MSA,[1] Urban	437	96	1,095	2,591	5,298	105,272	4.02	98	1,411	3,493	7,373	203,443	5.31
Non-MSA,[1] Rural	939	70	837	2,050	4,796	156,103	5.24	111	1,147	2,867	7,291	229,541	6.58
Group Quarters													
Group	168	24	301	775	2,733	17,453	3.16	24	348	943	3,604	23,907	3.34
Non-Group	17,679	3	1,544	3,436	7,744	301,034	4.42	4	2,063	4,955	12,085	471,208	5.45

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

Table L.3 2016 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	17,847	4	2,038	4,912	12,016	471,208	5.48	2	2,009	4,946	12,065	509,876	5.55	2	2,001	4,943	12,083	493,922	5.53
Pair Age Group																			
12-17, 12-17	2,386	4	709	1,826	4,004	29,858	2.32	2	608	1,785	4,061	25,880	2.41	2	606	1,777	4,057	25,112	2.41
12-17, 18-25	1,774	61	1,548	3,027	5,931	32,015	1.99	71	1,460	3,020	6,020	28,209	1.97	70	1,467	3,033	6,042	27,238	1.96
12-17, 26-34	883	113	1,800	3,290	6,031	65,888	2.48	93	1,853	3,530	6,279	71,655	2.50	91	1,837	3,515	6,261	72,259	2.53
12-17, 35-49	3,930	68	2,211	4,446	8,986	91,944	2.46	67	2,190	4,571	8,965	89,227	2.41	65	2,194	4,567	8,995	84,139	2.40
12-17, 50+	725	600	5,914	12,268	21,450	139,306	2.15	480	5,347	12,050	22,753	160,343	2.21	465	5,337	11,992	22,741	164,981	2.21
18-25, 18-25	2,512	24	1,253	2,987	6,659	57,738	2.35	16	1,253	3,028	6,970	31,429	2.21	16	1,260	3,017	7,010	28,241	2.19
18-25, 26-34	886	134	2,373	5,196	9,790	79,605	2.31	153	2,373	5,170	9,981	82,047	2.36	159	2,337	5,168	9,973	83,609	2.37
18-25, 35-49	1,134	126	4,605	9,386	18,687	123,416	2.16	79	4,514	9,480	19,017	86,633	2.08	77	4,512	9,551	19,178	86,417	2.06
18-25, 50+	722	729	10,661	19,654	35,379	280,488	1.96	685	9,928	18,642	35,710	193,530	1.96	684	9,872	18,740	35,421	168,474	1.95
26-34, 26-34	905	552	3,969	7,406	13,196	293,831	3.41	483	4,062	7,579	13,639	268,913	3.30	477	4,055	7,511	13,719	271,169	3.27
26-34, 35-49	445	748	7,128	14,434	25,064	226,855	2.15	433	7,160	14,214	25,408	232,407	2.10	430	7,188	14,058	25,619	230,909	2.10
26-34, 50+	234	2,544	21,368	42,250	72,084	325,194	2.01	1,742	18,500	39,889	75,106	331,550	2.06	1,698	18,561	39,917	73,319	334,463	2.07
35-49, 35-49	627	888	7,107	14,377	28,802	457,845	3.60	796	6,375	13,954	28,168	509,876	4.00	769	6,345	13,947	27,506	493,922	3.98
35-49, 50+	213	5,148	29,296	57,268	100,594	389,219	1.86	3,227	28,196	55,682	101,882	403,253	1.96	3,167	28,153	55,987	102,590	389,346	1.95
50+, 50+	471	1,030	40,390	75,342	113,516	471,208	1.63	967	38,369	74,022	115,395	422,499	1.61	965	38,739	73,876	115,381	380,154	1.60
Pair Race/Ethnicity																			
Hispanic or Latino	3,272	5	2,129	5,025	12,430	389,219	5.33	5	2,129	5,061	12,599	348,721	5.26	5	2,119	5,060	12,572	348,313	5.27
Black or African American	2,005	16	2,274	5,317	11,932	277,291	4.27	12	2,247	5,445	12,327	242,114	4.37	12	2,246	5,448	12,384	244,529	4.37
White	9,262	39	2,045	4,752	11,660	471,208	5.55	41	2,109	5,041	12,065	509,876	5.68	40	2,104	5,051	12,085	493,922	5.62
Other	1,157	4	1,470	4,740	13,802	328,035	6.65	2	1,527	4,799	14,214	331,550	6.38	2	1,522	4,801	14,173	329,271	6.40
White & Black or African American	180	24	2,149	4,993	12,376	181,457	4.81	31	2,430	6,073	14,874	194,992	4.14	30	2,413	6,086	14,963	195,248	4.13
White & Hispanic or Latino	752	54	1,882	5,045	13,071	427,353	6.02	70	1,949	5,206	12,713	352,709	5.36	68	1,947	5,228	12,736	353,870	5.40
White & Other	749	32	1,750	4,224	11,633	328,371	5.65	42	1,208	2,992	7,707	200,103	5.46	41	1,199	2,980	7,706	204,440	5.50
Black or African American & Hispanic or Latino	144	183	1,934	5,997	13,178	137,109	3.45	102	1,731	5,275	11,015	136,892	3.76	103	1,739	5,302	11,080	137,829	3.77
Black or African American & Other	172	21	2,522	5,496	13,943	150,154	3.79	7	944	2,459	6,242	64,501	3.96	7	936	2,438	6,248	64,812	3.98
Hispanic or Latino & Other	154	68	2,061	6,640	15,698	142,099	3.34	67	1,357	4,324	10,733	80,300	3.16	65	1,370	4,310	10,852	80,516	3.18

Table L.3 2016 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	3,714	24	1,949	4,636	11,621	427,353	4.72	16	1,881	4,687	11,651	326,090	4.65	16	1,877	4,686	11,645	326,337	4.65
Female, Female	4,148	32	1,873	4,348	10,296	436,797	5.25	27	1,896	4,412	10,293	361,713	5.04	27	1,885	4,396	10,327	352,103	5.03
Male, Female	9,985	4	2,178	5,246	13,196	471,208	5.49	2	2,140	5,345	13,052	509,876	5.65	2	2,131	5,340	13,078	493,922	5.62
Household Size																			
Two	4,221	24	1,804	4,399	11,925	294,632	4.49	16	1,811	4,520	11,774	262,070	4.57	16	1,802	4,513	11,821	262,082	4.58
Three	5,609	4	1,839	4,043	8,585	436,797	7.23	2	1,767	4,078	8,909	395,535	6.97	2	1,764	4,078	8,943	388,975	6.94
Four or More	8,017	18	2,405	6,297	14,622	471,208	5.10	11	2,345	6,240	14,513	509,876	5.29	11	2,339	6,264	14,531	493,922	5.25
Census Region																			
Northeast	3,159	39	1,695	4,689	12,302	457,845	6.02	32	1,604	4,566	12,510	509,876	6.36	33	1,614	4,563	12,535	493,922	6.34
South	6,023	4	2,701	5,963	13,259	364,376	4.77	2	2,679	5,973	13,410	417,868	4.84	2	2,673	5,988	13,448	420,577	4.83
Midwest	4,255	155	2,018	4,207	9,932	436,797	5.14	95	2,066	4,369	9,841	364,500	5.20	90	2,058	4,370	9,883	352,103	5.17
West	4,410	24	1,500	4,448	12,562	471,208	6.12	16	1,505	4,520	12,469	422,499	6.03	16	1,499	4,507	12,504	374,658	5.98
Quarter																			
Quarter1	4,317	5	2,273	5,291	11,875	389,219	5.55	5	2,160	5,308	11,753	403,253	5.87	5	2,153	5,309	11,821	389,346	5.86
Quarter2	4,500	21	1,967	4,692	11,748	436,797	5.32	7	1,935	4,805	11,801	393,195	5.31	7	1,933	4,767	11,794	384,235	5.30
Quarter3	4,575	41	1,945	4,681	12,156	457,845	5.41	32	1,962	4,708	12,403	509,876	5.35	33	1,947	4,703	12,452	493,922	5.35
Quarter4	4,455	4	2,031	4,986	12,275	471,208	5.62	2	2,011	4,978	12,306	422,499	5.65	2	2,009	4,970	12,371	420,577	5.58
% Hispanic or Latino in Segment																			
50-100%	1,751	61	2,538	6,032	15,147	427,353	5.22	48	2,579	6,276	15,794	333,943	4.77	48	2,564	6,288	15,853	335,801	4.80
10-<50%	4,646	5	2,549	6,021	14,467	471,208	5.10	5	2,453	6,013	14,443	422,499	5.19	5	2,447	6,021	14,475	420,577	5.16
<10%	11,450	4	1,839	4,377	10,872	457,845	5.61	2	1,800	4,399	10,658	509,876	5.81	2	1,799	4,399	10,633	493,922	5.78
% Black or African American in Segment																			
50-100%	1,366	4	2,259	5,350	12,519	229,206	4.01	2	2,111	5,151	12,184	249,145	4.21	2	2,086	5,149	12,165	246,870	4.21
10-<50%	3,522	5	2,372	5,567	13,540	427,353	4.69	5	2,259	5,596	13,471	395,535	4.67	5	2,259	5,581	13,510	388,975	4.67
<10%	12,959	24	1,946	4,685	11,622	471,208	5.83	16	1,940	4,763	11,640	509,876	5.92	16	1,944	4,756	11,654	493,922	5.89
% Owner-Occupied DUs[1] in Segment																			
50-100%	13,239	18	2,265	5,330	12,949	471,208	5.31	11	2,230	5,360	12,937	509,876	5.40	11	2,230	5,359	12,963	493,922	5.37
10-<50%	3,609	4	2,055	4,786	11,490	389,219	5.33	2	2,038	4,781	11,789	321,330	5.32	2	2,044	4,790	11,772	328,207	5.31
<10%	999	16	642	1,619	3,879	97,757	4.79	12	599	1,536	3,958	93,946	4.50	12	600	1,522	3,931	93,993	4.52
Combined Median Rent/Housing Value																			
1st Quintile	2,940	49	1,712	3,991	9,605	381,195	5.55	27	1,669	4,011	9,507	395,535	5.65	27	1,665	4,019	9,468	388,975	5.63
2nd Quintile	4,202	24	1,610	4,036	10,356	325,194	5.67	16	1,588	3,966	10,094	384,899	6.09	16	1,581	3,973	10,207	352,188	6.09
3rd Quintile	4,164	16	2,051	4,924	11,927	389,219	4.92	7	2,053	5,110	12,145	326,090	4.89	7	2,056	5,128	12,172	326,337	4.89
4th Quintile	3,855	4	2,379	5,572	13,756	471,208	5.38	2	2,337	5,713	13,687	422,499	5.17	2	2,329	5,730	13,743	374,658	5.15
5th Quintile	2,686	5	2,964	6,711	15,464	457,845	5.35	5	2,930	6,781	15,400	509,876	5.56	5	2,918	6,767	15,396	493,922	5.50

Table L.3 2016 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]	7,632	4	3,246	7,165	16,676	471,208	4.78	2	3,155	7,251	16,762	509,876	4.83	2	3,144	7,260	16,810	493,922	4.80
Medium to Small MSA[1]	8,839	24	1,532	3,761	9,133	381,195	5.64	16	1,507	3,781	9,071	395,535	5.79	16	1,507	3,776	9,069	388,975	5.79
Non-MSA,[1] Urban	437	98	1,411	3,493	7,373	203,443	5.31	54	1,463	3,612	7,399	190,744	5.29	53	1,462	3,560	7,416	191,207	5.32
Non-MSA,[1] Rural	939	111	1,147	2,867	7,291	229,541	6.58	40	1,179	3,103	7,232	219,994	6.27	38	1,190	3,085	7,271	222,665	6.17
Group Quarters																			
Group	168	24	348	943	3,604	23,907	3.34	16	326	963	3,474	27,666	3.47	16	320	980	3,407	26,821	3.48
Non-Group	17,679	4	2,063	4,955	12,085	471,208	5.45	2	2,040	4,985	12,139	509,876	5.53	2	2,043	4,985	12,165	493,922	5.50
Pair Relationship Domain[4]																			
Parent-Child (12-14)	2,749	113	2,055	4,220	8,603	125,580	2.78	137	2,245	4,555	9,151	145,142	2.82	136	2,241	4,544	9,174	147,754	2.81
Parent-Child (12-17)	5,048	68	2,251	4,557	9,831	125,580	2.71	67	2,394	4,858	10,119	145,142	2.70	65	2,381	4,860	10,131	147,754	2.69
Parent-Child (12-20)	5,846	68	2,472	5,150	11,450	148,501	2.78	67	2,613	5,448	12,026	161,954	2.75	65	2,606	5,443	12,037	157,575	2.74
Sibling (12-14)-Sibling (15-17)	1,421	4	716	1,743	3,992	29,858	2.33	2	664	1,868	4,375	25,364	2.34	2	662	1,871	4,375	25,112	2.34
Sibling (12-17)-Sibling (18-25)	1,626	61	1,556	3,027	5,854	29,853	1.99	71	1,480	3,050	6,060	28,209	1.97	70	1,476	3,068	6,089	27,238	1.96
Spouse-Spouse/ Partner-Partner	3,363	64	2,552	6,692	18,436	471,208	4.93	52	2,574	6,841	18,641	509,876	5.08	52	2,560	6,815	18,657	493,922	5.04
Spouse-Spouse/ Partner-Partner with Children (Younger Than 18)	1,727	67	2,562	6,235	14,288	457,845	6.42	76	2,770	6,691	15,152	509,876	6.70	75	2,779	6,666	15,143	493,922	6.61

[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.