SAMHSA
Substance Abuse and Mental Health Services Administration

2018 National Survey on Drug Use and Health (NSDUH): Methodological Resource Book

Section 12: Questionnaire Dwelling Unit-Level and Person Pair-Level Sampling Weight Calibration

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

May 2020

2018 National Survey on Drug Use and Health (NSDUH): Methodological Resource Book

Acknowledgments

This report was prepared for the Substance Abuse and Mental Health Services Administration (SAMHSA), U.S. Department of Health and Human Services (HHS), under Contract No. HHSS283201700002C with RTI International. Kathryn Piscopo served as the government project officer and as the contracting officer representative.

Electronic Access

This publication may be downloaded at https://www.samhsa.gov/data/.

Recommended Citation

Recommended Citation: Center for Behavioral Health Statistics and Quality. (2020). *2018 National Survey on Drug Use and Health (NSDUH): Methodological Resource Book, Section 12, Questionnaire dwelling unit-level and person pair-level sampling weight calibration*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Originating Office

Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration, 5600 Fishers Lane, Room 15-E09D, Rockville, MD 20857. For questions about this report, please e-mail CBHSQrequest@samhsa.hhs.gov.

U.S. Department of Health and Human Services
Substance Abuse and Mental Health Services Administration
Center for Behavioral Health Statistics and Quality
Populations Survey Branch

May 2020

Table of Contents

Chapter **Page**

Table of Contents (continued)

Chapter **Page**

Appendix

List of Tables

Table **Page**

List of Tables (continued)

Table **Page**

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, 2020a) 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 (2020a) 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 2018. 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, 2020b), 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-stage interview design[1] of the survey was capitalized upon to estimate these controls from the large screener sample at the screening interview stage. The dwelling unit 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

[1] The screening interview involves listing all household members along with their basic demographic information. Immediately after completion of the screening, 0, 1, or 2 people in the household will be selected to complete the actual questionnaire interview. The first stage is screening, and the second stage is interviewing.

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).[2] 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).

[2] 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),[3] 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 2018 NSDUHs. The coordinated design facilitated 50 percent overlap in third-stage units (area segments) within each successive 2-year period from 2014 through 2018. This designed sample overlap slightly increased the precision of estimates of year-to-year trends because of the expected small but positive correlation resulting from the overlapping sampled area segments between successive survey years. The 50 percent overlap of segments significantly reduced segment listing costs because only one-half of the segments needed to be listed for the 2015 through 2018 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 2014-2018 survey sample was designed to yield

- 4,560 completed interviews in California;

[3] This report presents information from the 2018 National Survey on Drug Use and Health (NSDUH). Prior to 2002, the survey was called the National Household Survey on Drug Abuse (NHSDA).

- 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 2014-2018 design, was the incorporation of census block groups at the second stage of selection to potentially reduce sampling variance and facilitate moving to an address-based sampling design in the future, if desired. NSDUH is sometimes referred to as a two-stage interview sample where the first interview stage 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 interview stage. In the 2014-2018 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 2014-2018 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-stage interview sample design) for calibration at levels (such as pair and QDU) for which suitable U.S. Census Bureau–based controls were not available. The second interview stage 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 (2014-2018). 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 2014-2018 NSDUH Sample Sizes

Sample Unit		2014	2015	2016	2017	2018
SDU	Selected	154,533	165,328	173,149	184,266	193,456
	Completed	127,605	132,210	135,188	138,061	141,879
QDU	Selected	64,796	66,721	67,574	68,889	70,048
	Completed	49,672	50,119	50,095	50,328	50,373
Pair	Selected	26,844	27,778	28,033	28,778	29,063
	Completed	18,229	17,954	17,847	17,704	17,418
Person	Selected	91,640	94,499	95,607	97,667	99,111
	Completed	67,901	68,073	67,942	68,032	67,791

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-stage interview 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], 2020a). 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 potential 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

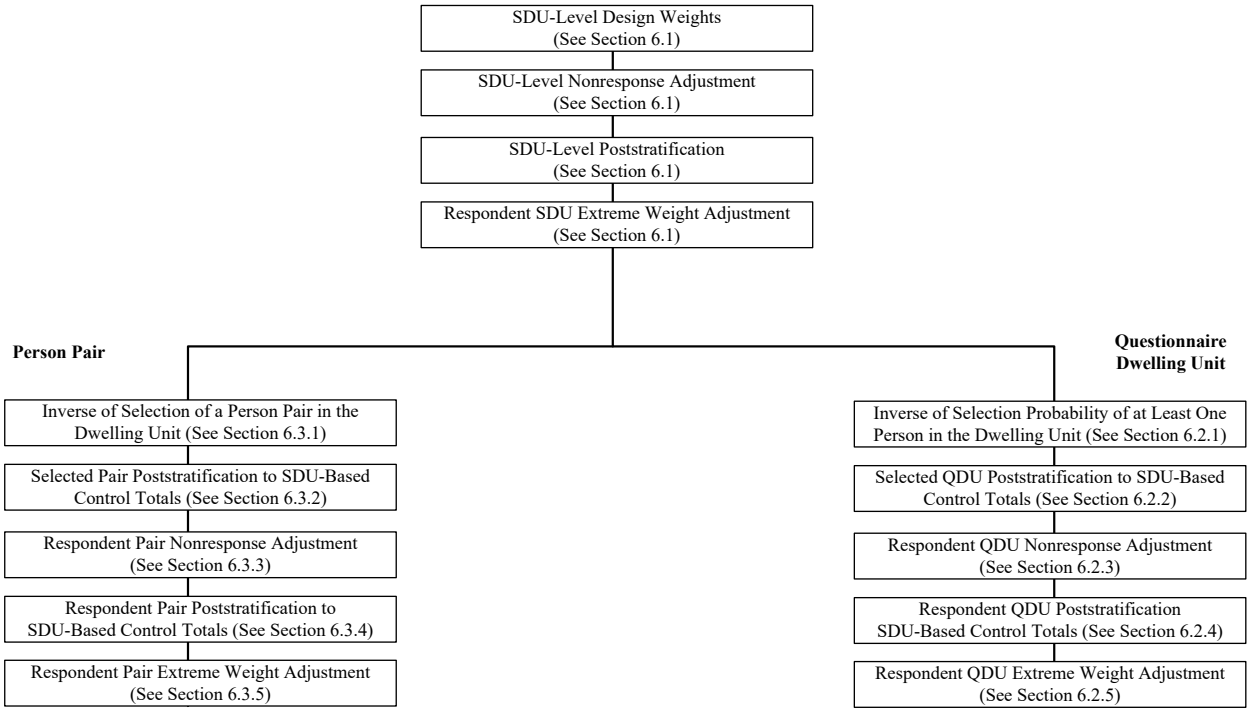
Some of the multiplicities, including counts of all possible pairs in a household for a given domain, were used for poststratification. Details are provided in Chapter 4. Additional information on the imputation of pair relationships, multiplicity factors, and household-level person counts for poststratification can be found in the NSDUH Methodological Resource Book editing and imputation report (CBHSQ, 2020a). Special consideration is required for analysis of pair-level data, and details can be found in *How To Prepare and Analyze Pair Data in the National Survey on Drug Use and Health* (CBHSQ, 2017).

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

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

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

Exhibit 1.1 QDU and Pair Sampling Weight Calibration Steps



In dealing with extreme weights, it is assumed that they arise due to design (due to an imperfect frame, assignment of very small selection probabilities to some units, or a big weight adjustment factor after calibration) so that they make the sample representative of the population and, hence, do not introduce bias. The only problem is that they may lead to highly unstable estimates similar to the problem of Basu's circus elephants[4] (Hajek, 1971). So, we need to perform some treatment (such as winsorization[5]) 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

[4] 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.

[5] 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 2018 survey was the same as the one used in the 1999-2017 surveys. The domains were defined as the cross-classification of state, pair age,[6] 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.

[6] 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 2002-2017 National Surveys on Drug Use and Health (NSDUHs),[7] the 2018 NSDUH had a two-stage interview 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, 2019).

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 2018, 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 a modification of 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.

[7] This report presents information from the 2018 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 2014-2018 NSDUHs

	2014			2015			2016			2017			2018		
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	154,533	127,605	82.57	165,328	132,210	79.97	173,149	135,188	78.08	184,266	138,061	74.92	193,456	141,879	73.34
QDUs															
Total QDUs	64,796	49,672	76.66	66,721	50,119	75.12	67,574	50,095	74.13	68,889	50,328	73.06	70,048	50,373	71.91
People															
Total People	91,640	67,901	74.10	94,499	68,073	72.04	95,607	67,942	71.06	97,667	68,032	69.66	99,111	67,791	68.40
12-17	21,392	17,007	79.50	21,859	16,911	77.36	22,323	17,081	76.52	22,750	17,026	74.84	22,962	16,820	73.25
18-25	21,726	16,449	75.71	23,211	17,097	73.66	22,836	16,435	71.97	23,707	16,469	69.47	24,363	16,561	67.98
26-34	14,004	10,252	73.21	14,720	10,446	70.96	15,022	10,528	70.08	15,140	10,416	68.80	15,241	10,224	67.08
35-49	19,065	13,590	71.28	19,341	13,304	68.79	19,988	13,572	67.90	20,280	13,639	67.25	20,958	13,889	66.27
50+	15,453	10,603	68.61	15,368	10,315	67.12	15,438	10,326	66.89	15,790	10,482	66.38	15,587	10,297	66.06
Non-Pairs[4]															
Total Non-Pairs	90,443	31,443	N/A	104,432	32,165	N/A	107,155	32,248	N/A	109,283	32,624	N/A	112,816	32,955	N/A
0,0	62,809	N/A	N/A	65,489	N/A	N/A	67,614	N/A	N/A	69,172	N/A	N/A	71,831	N/A	N/A
Total Singletons	37,952	31,443	82.85	38,943	32,165	82.60	39,541	32,248	81.56	40,111	32,624	81.33	40,985	32,955	80.41
0, 12-17	4,850	4,704	96.99	5,244	5,014	95.61	5,144	4,997	97.14	5,155	4,997	96.94	5,391	5,018	93.08
0, 18-25	7,250	6,647	91.68	7,583	7,102	93.66	7,647	6,895	90.17	7,858	7,079	90.09	7,895	7,019	88.90
0, 26-34	7,460	6,034	80.88	7,726	6,166	79.81	8,045	6,270	77.94	7,987	6,247	78.21	8,231	6,269	76.16
0, 35-49	8,074	6,450	79.89	8,093	6,320	78.09	8,442	6,596	78.13	8,601	6,679	77.65	8,878	6,923	77.98
0, 50+	10,318	7,608	73.74	10,297	7,563	73.45	10,263	7,490	72.98	10,510	7,622	72.52	10,590	7,726	72.96
Pairs															
Total Pairs[5]	26,844	18,229	67.91	27,778	17,954	64.63	28,033	17,847	63.66	28,778	17,704	61.52	29,063	17,418	59.93
12-17, 12-17	3,070	2,407	78.40	2,962	2,253	76.06	3,199	2,386	74.59	3,261	2,368	72.62	3,296	2,349	71.27
12-17, 18-25	2,443	1,832	74.99	2,571	1,795	69.82	2,548	1,774	69.62	2,679	1,758	65.62	2,680	1,764	65.82
12-17, 26-34	1,297	941	72.55	1,299	939	72.29	1,281	883	68.93	1,338	894	66.82	1,243	826	66.45
12-17, 35-49	5,530	3,940	71.25	5,654	3,888	68.77	5,829	3,930	67.42	5,845	3,870	66.21	5,939	3,836	64.59
12-17, 50+	1,132	776	68.55	1,167	769	65.90	1,123	725	64.56	1,211	771	63.67	1,117	678	60.70
18-25, 18-25	3,743	2,585	69.06	4,043	2,654	65.64	3,958	2,512	63.47	4,167	2,467	59.20	4,381	2,548	58.16
18-25, 26-34	1,378	870	63.13	1,577	975	61.83	1,429	886	62.00	1,443	839	58.14	1,557	851	54.66
18-25, 35-49	1,906	1,180	61.91	2,092	1,186	56.69	2,013	1,134	56.33	2,084	1,176	56.43	2,187	1,179	53.91
18-25, 50+	1,263	750	59.38	1,302	731	56.14	1,283	722	56.27	1,309	683	52.18	1,282	652	50.86
26-34, 26-34	1,356	865	63.79	1,492	870	58.31	1,518	905	59.62	1,551	871	56.16	1,495	827	55.32
26-34, 35-49	737	442	59.97	716	408	56.98	788	445	56.47	810	445	54.94	788	408	51.78
26-34, 50+	420	235	55.95	418	218	52.15	443	234	52.82	460	249	54.13	432	216	50.00
35-49, 35-49	1,160	658	56.72	1,158	635	54.84	1,213	627	51.69	1,233	628	50.93	1,338	661	49.40
35-49, 50+	498	262	52.61	470	232	49.36	490	213	43.47	474	213	44.94	490	221	45.10
50+, 50+	911	486	53.35	857	401	46.79	918	471	51.31	913	472	51.70	838	402	47.97

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

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

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 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 nonresponse (nr) via weighting classes, (2) for poststratification (ps) via raking ratio adjustments, and (3) for extreme values (ev) via winsorization. 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, 2020b).

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-stage interview 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, 2020b).

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

[8] 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.

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.[9] Similar to previous National Surveys on Drug Use and Health (NSDUHs), for the GEM modeling used in the 2018 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.

   a. Youth, Young Adult, Adult;

[9] Deep stratification refers to the stratification that was used in the sample design. In the case of the 2018 survey, deep stratification refers to the cross-classification of state sampling region by age group.

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 2018 survey. This DU gave rise to a selected pair of respondents, one aged 81 and the other aged 57. The selection probability in this DU was 0.14578 for a respondent aged 50 or older. Using the formula (2.6) in Chapter 2, the pair selection probability was computed to be 0.000492242. Therefore, the inverse of the selection probability, the pair-level design weight, was 2,031.52. 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 1999-2017 surveys, the extreme weight definition was based on the following hierarchy of domains:

1. Pair age group[10] (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.

[10] 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 2018 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 2014-2018 NSDUH sample design plans slightly modified the 2005-2013 approach, such that the basic sampling plan involved five stages of selection: (1) selection of census tracts, (2) selection of census block groups within each state sampling (SS) region, (3) the selection of subareas or segments (comprising U.S. Census Bureau blocks) within SS regions, (4) the selection of dwelling units (DUs) within these subareas, and finally, (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 2018 NSDUH Methodological Resource Book (MRB) sample design report (Center for Behavioral Health Statistics and Quality [CBHSQ], 2019).

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 screener dwelling units (SDU) and QDU, person, and pair levels for the 2018 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 2018 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 2014-2018 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 2002-2018 survey pair data, the built-in ev control feature was used for each adjustment step.

Exhibit 6.1 Summary of 2018 NSDUH QDU Sample Weight Components

Screener Dwelling Unit Level

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

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

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, 2020b). 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 2018 NSDUH Person Pair Sample Weight Components

Screener Dwelling Unit Level

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

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

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, 2020b). 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	2014		2015		2016		2017		2018	
	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs
QDU										
Northeast	12,950	9,664	13,519	9,777	13,414	9,552	14,037	9,915	14,264	9,753
South	21,448	16,680	21,887	16,708	22,287	16,810	22,628	16,901	22,716	16,931
Midwest	15,276	11,618	15,808	11,698	16,025	11,768	16,282	11,760	16,550	11,810
West	15,122	11,710	15,507	11,936	15,848	11,965	15,942	11,752	16,518	11,879
Total	64,796	49,672	66,721	50,119	67,574	50,095	68,889	50,328	70,048	50,373
Model Group	**2014**		**2015**		**2016**		**2017**		**2018**	
	Selected Pairs	Completed Pairs	Selected Pairs	Completed Pairs	Selected Pairs	Completed Pairs	Selected Pairs	Completed Pairs	Selected Pairs	Completed Pairs
Pair										
Northeast + South	13,969	9,436	14,502	9,309	14,543	9,182	15,072	9,284	15,024	9,072
Midwest + West	12,875	8,793	13,276	8,645	13,490	8,665	13,706	8,420	14,039	8,346
Total	26,844	18,229	27,778	17,954	28,033	17,847	28,778	17,704	29,063	17,418

QDU = questionnaire dwelling unit.

6.1 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 sample weights reflect the probability of selecting the DUs. These components were derived from (1) the probability of selecting the census tract and (2) census block groups within each SS region, (3) the probability of selecting the geographic segment within each SS region, (4) a quarter segment weight adjustment, (5) a subsegmentation inflation factor, (6) the probability of selecting a DU from within each counted and listed sampled segment, (7) the probability of inclusion of added DUs, and (8) DU percent release adjustment. The three remaining weight components, #9 through #11, are GEM calibration adjustments accounting for (9) DU nonresponse at the screening level, (10) DU poststratification to census controls, and (11) DU-level ev adjustment, although in 2017, ev adjustment at this stage was deemed unnecessary, and thus Weight Component #11 was set to one for all respondent DUs. The person-level, QDU-level, and person pair-level weights use the product of the above 11 weight components as the common initial weight before further adjustments. For more detailed information on Weight Components #1 through #3 and #5 through #8, refer to the 2018 NSDUH MRB sample design report (CBHSQ, 2019), and for more detail on Weight Components #4 and #9 through #11, see the 2018 NSDUH MRB person-level sampling weight calibration report (CBHSQ, 2020b).

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 (2014).

Weight Component #2, a 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 (2014).

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 (2020b).

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 formula (2.1) or formula (2.6) in Chapter 2. This weight component is the inverse of the selection probability discussed above.

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

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

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

6.3.3 Pair Weight Component #14: Respondent Pair Nonresponse Adjustment

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

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

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

This final poststratification utilized the same set of initially proposed predictor variables as previous adjustment steps. In addition, 10 pair relationship domain-level indicator variables were added to the set of covariates. The control totals for GEM calibration were derived from the SDU sample of all possible pairs of eligible people, as was done for Weight Component #13. The calibration control totals for these 10 domains used household-level person counts and the final QDU weights. As mentioned in the introduction, use of these household-level count totals for pair relationship domains in GEM calibration provided Hajek-type weight adjustment in the interest of obtaining more stable estimates. In setting up calibration covariates, multiplicity factors were needed. These factors, as discussed in the introduction, are used in constructing estimates for person-level parameters based on pair-related drug behavior. The factors depend on the pair domains of interest. For a selected set of pair domains, multiplicity factors are provided along with the pair-level analysis weights. See Chapter 11 in the NSDUH MRB editing and imputation report (CBHSQ, 2020a) 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 2018 National Survey on Drug Use and Health for various national domains. This table also shows the weighted response rates.[11] Most domains reflect the overall 70.08 percent response rate, with most rates between 70 and 75 percent, although the highest response rate is 93.25 percent, from the Group category of the Group Quarters variable. The lowest response rate came from the 12-17 age group of Household Type, with 54.14 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 2018 survey for various national domains. Because of the nature of the pair data, the response rates were lower in all domains examined than at the QDU level, with an overall response rate of 54.40 percent. The response rates range from a low of 38.65 percent in the Pair Race/Ethnicity category Other to a high of 77.34 percent from the Group category of the Group Quarters variable. This broad 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), and the highest response rate being in the group level of the variable group quarters may be driven by this variable including 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. Before 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 2018 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

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

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, 2020b). 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 2018 QDU domain Three of Household Size, indicates a sample size of 12,493 with a total design-based weight of 19,479,364 and a census total of 19,491,262 with an initial slippage rate of -0.64 percent, which would imply a common weight adjustment approximately equal to 1.006445, if this were the only calibration control. Similarly, looking at pair data in Table K.1, the pair domain category of Pair Age Group 26-34, 50+ has a sample size of 216, a design-based weight of 14,148,155, and a census total of 13,803,151, showing an initial slippage of 2.50 percent. The resultant required adjustment would be approximately equal to 0.975615, 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 7.7b.

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: 2018

Domain	*n*	Baseline (B)[1]	Final (F)[2]	(B-F)/F% (Estimate)	(B-F)/F% (SE)
Households with Family Income					
$0-<$10,000	3,882	9,027,340 (260,264)	9,022,130 (260,136)	0.06	0.05
$10,000-<$20,000	5,255	14,507,932 (370,491)	14,505,274 (370,561)	0.02	-0.02
$20,000-<$30,000	5,166	13,462,767 (293,038)	13,467,634 (293,853)	-0.04	-0.28
$30,000-<$40,000	4,804	12,412,529 (281,788)	12,413,540 (281,961)	-0.01	-0.06
$40,000-<$50,000	5,023	12,528,921 (257,537)	12,526,487 (257,741)	0.02	-0.08
$50,000-<$75,000	7,807	19,625,458 (382,260)	19,622,968 (382,293)	0.01	-0.01
$75,000-<$100,000	6,028	14,691,263 (287,458)	14,696,136 (287,650)	-0.03	-0.07
$100,000+	12,408	30,096,009 (601,440)	30,098,049 (602,593)	-0.01	-0.19
Households with Number of Youths (<18)					
0	23,817	84,472,851 (994,477)	84,472,648 (995,332)	0.00	-0.09
1	10,565	17,452,785 (256,037)	17,454,765 (256,512)	-0.01	-0.19
2	9,448	15,256,139 (255,794)	15,259,151 (256,249)	-0.02	-0.18
3	4,203	6,163,532 (134,083)	6,159,433 (134,147)	0.07	-0.05
4+	2,340	3,006,912 (83,364)	3,006,223 (83,478)	0.02	-0.14
Households with Insurance Coverage					
Yes	45,433	115,195,930 (1,042,369)	115,200,435 (1,043,557)	-0.00	-0.11
No	4,940	11,156,289 (247,195)	11,151,784 (247,206)	0.04	-0.00
Households with Number of Older Adults (65+)					
0	42,525	90,099,679 (866,580)	90,104,551 (867,809)	-0.01	-0.14
1	4,978	22,739,317 (497,665)	22,732,361 (497,477)	0.03	0.04
2	2,808	13,301,933 (371,667)	13,303,972 (371,794)	-0.02	-0.03
3+	62	211,290 (33,863)	211,335 (33,914)	-0.02	-0.15

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: 2018

Drug	Mother User	*n*	Baseline[1]	Final[2]
Alcohol				
Lifetime	Yes	2,598	27.67 (1.57)	27.66 (1.58)
	No	398	15.69 (3.04)	16.60 (3.43)
	Overall	2,996	25.76 (1.45)	25.87 (1.48)
Past Year	Yes	2,156	23.95 (1.73)	23.81 (1.74)
	No	840	12.84 (1.87)	13.48 (2.09)
	Overall	2,996	20.65 (1.38)	20.72 (1.42)
Past Month	Yes	1,681	11.31 (1.35)	11.19 (1.36)
	No	1,315	5.43 (0.93)	5.52 (0.96)
	Overall	2,996	8.67 (0.86)	8.62 (0.87)
Cigarettes				
Lifetime	Yes	1,793	12.75 (1.37)	12.75 (1.38)
	No	1,203	4.85 (0.91)	4.96 (0.94)
	Overall	2,996	9.20 (0.87)	9.23 (0.88)
Past Year	Yes	685	7.82 (1.46)	7.83 (1.45)
	No	2,311	4.03 (0.62)	4.03 (0.63)
	Overall	2,996	4.75 (0.58)	4.75 (0.58)
Past Month	Yes	596	3.44 (1.03)	3.43 (1.03)
	No	2,400	2.23 (0.50)	2.18 (0.49)
	Overall	2,996	2.43 (0.45)	2.39 (0.44)

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: 2018

Drug	Father User	*n*	Baseline[1]	Final[2]
Alcohol				
Lifetime	Yes	1,723	24.63 (1.85)	24.76 (1.85)
	No	153	17.69 (5.61)	16.51 (5.05)
	Overall	1,876	24.01 (1.75)	24.02 (1.74)
Past Year	Yes	1,457	22.32 (1.98)	22.38 (1.96)
	No	419	12.72 (3.22)	12.40 (3.12)
	Overall	1,876	20.15 (1.69)	20.13 (1.67)
Past Month	Yes	1,207	9.61 (1.42)	9.79 (1.43)
	No	669	5.45 (1.80)	5.51 (1.81)
	Overall	1,876	8.07 (1.12)	8.20 (1.12)
Cigarettes				
Lifetime	Yes	1,287	9.72 (1.45)	9.80 (1.44)
	No	589	1.33 (0.51)	1.28 (0.48)
	Overall	1,876	6.86 (0.98)	6.89 (0.98)
Past Year	Yes	436	6.26 (1.78)	6.32 (1.81)
	No	1,440	3.11 (0.73)	3.15 (0.72)
	Overall	1,876	3.73 (0.69)	3.78 (0.69)
Past Month	Yes	363	2.42 (1.46)	2.42 (1.44)
	No	1,513	1.37 (0.53)	1.40 (0.52)
	Overall	1,876	1.54 (0.50)	1.56 (0.50)

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: 2018

Drug	Mother User	*n*	Baseline[1]	Final[2]
Any Illicit				
Lifetime	Yes	1,497	26.58 (1.94)	26.29 (1.94)
	No	1,499	17.44 (1.81)	17.79 (1.87)
	Overall	2,996	21.70 (1.33)	21.74 (1.35)
Past Year	Yes	476	21.49 (3.66)	21.42 (3.72)
	No	2,520	15.05 (1.31)	15.17 (1.34)
	Overall	2,996	16.04 (1.22)	16.12 (1.25)
Past Month	Yes	240	18.54 (4.33)	17.96 (4.21)
	No	2,756	7.93 (1.01)	7.94 (1.01)
	Overall	2,996	8.64 (0.99)	8.60 (0.99)
Marijuana				
Lifetime	Yes	1,389	19.54 (1.82)	19.28 (1.83)
	No	1,607	11.51 (1.61)	11.77 (1.66)
	Overall	2,996	14.97 (1.19)	14.99 (1.22)
Past Year	Yes	364	19.49 (4.36)	19.47 (4.47)
	No	2,632	11.85 (1.19)	11.88 (1.21)
	Overall	2,996	12.71 (1.15)	12.73 (1.18)
Past Month	Yes	201	14.58 (4.05)	13.85 (3.82)
	No	2,795	7.30 (1.01)	7.30 (1.01)
	Overall	2,996	7.70 (0.98)	7.66 (0.97)

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: 2018

Drug	Father User	*n*	Baseline[1]	Final[2]
Any Illicit				
Lifetime	Yes	1,073	28.83 (2.41)	29.05 (2.41)
	No	803	16.54 (2.13)	16.29 (2.08)
	Overall	1,876	23.39 (1.64)	23.41 (1.62)
Past Year	Yes	284	19.84 (4.94)	19.90 (4.81)
	No	1,592	13.85 (1.44)	13.89 (1.44)
	Overall	1,876	14.61 (1.42)	14.65 (1.40)
Past Month	Yes	170	10.17 (3.96)	11.02 (4.70)
	No	1,706	7.22 (1.27)	7.07 (1.22)
	Overall	1,876	7.44 (1.21)	7.36 (1.18)
Marijuana				
Lifetime	Yes	974	15.77 (2.13)	15.92 (2.11)
	No	902	9.39 (1.67)	9.24 (1.61)
	Overall	1,876	12.64 (1.33)	12.63 (1.30)
Past Year	Yes	225	17.44 (5.89)	17.39 (5.71)
	No	1,651	10.15 (1.32)	10.19 (1.30)
	Overall	1,876	10.87 (1.33)	10.90 (1.31)
Past Month	Yes	149	10.75 (4.67)	11.65 (5.57)
	No	1,727	6.07 (1.24)	5.95 (1.18)
	Overall	1,876	6.34 (1.19)	6.29 (1.16)

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: 2018

Drug	Parent Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Alcohol				
Lifetime	Yes	2,735	26.74 (1.67)	26.81 (1.66)
	No	2,036	24.34 (1.62)	24.33 (1.64)
	Overall	4,771	25.70 (1.20)	25.73 (1.20)
Past Year	Yes	2,735	22.53 (1.62)	22.59 (1.61)
	No	2,036	19.37 (1.54)	19.36 (1.55)
	Overall	4,771	21.16 (1.16)	21.18 (1.16)
Past Month	Yes	2,735	9.53 (1.21)	9.54 (1.20)
	No	2,036	8.46 (1.08)	8.42 (1.08)
	Overall	4,771	9.06 (0.83)	9.05 (0.83)
Cigarettes				
Lifetime	Yes	2,735	8.80 (1.15)	8.79 (1.14)
	No	2,036	8.77 (0.99)	8.75 (0.99)
	Overall	4,771	8.78 (0.78)	8.78 (0.78)
Past Year	Yes	2,735	4.92 (0.70)	4.89 (0.68)
	No	2,036	4.22 (0.70)	4.27 (0.71)
	Overall	4,771	4.61 (0.50)	4.62 (0.50)
Past Month	Yes	2,735	2.48 (0.62)	2.40 (0.59)
	No	2,036	2.37 (0.58)	2.41 (0.58)
	Overall	4,771	2.43 (0.42)	2.40 (0.41)

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

Table 7.5 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: 2018

Drug	Parent Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Any Illicit				
Lifetime	Yes	2,735	22.01 (1.54)	21.93 (1.53)
	No	2,036	25.15 (1.67)	25.32 (1.69)
	Overall	4,771	23.37 (1.12)	23.40 (1.13)
Past Year	Yes	2,735	16.14 (1.44)	16.11 (1.42)
	No	2,036	16.26 (1.41)	16.43 (1.43)
	Overall	4,771	16.19 (1.03)	16.25 (1.04)
Past Month	Yes	2,735	9.22 (1.27)	9.13 (1.26)
	No	2,036	8.06 (1.03)	8.09 (1.02)
	Overall	4,771	8.72 (0.87)	8.68 (0.86)
Marijuana				
Lifetime	Yes	2,735	15.34 (1.43)	15.27 (1.41)
	No	2,036	14.83 (1.39)	14.93 (1.41)
	Overall	4,771	15.12 (1.01)	15.12 (1.01)
Past Year	Yes	2,735	12.81 (1.36)	12.78 (1.35)
	No	2,036	12.38 (1.33)	12.48 (1.35)
	Overall	4,771	12.63 (0.98)	12.65 (0.98)
Past Month	Yes	2,735	8.14 (1.25)	8.06 (1.24)
	No	2,036	7.07 (1.02)	7.11 (1.00)
	Overall	4,771	7.68 (0.86)	7.65 (0.85)

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 Child about the Dangers of Tobacco, Alcohol, or Drug Use within the Past 12 Months: 2018

Drug	Mother Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Alcohol				
Lifetime	0 times	229	19.25 (4.09)	18.87 (4.02)
	1-2 times	522	21.26 (2.78)	21.12 (2.79)
	A few times	764	22.21 (2.46)	22.94 (2.66)
	Many times	1,315	30.83 (2.30)	30.83 (2.30)
	Overall	2,830	25.97 (1.49)	26.14 (1.51)
Past Year	0 times	229	13.27 (3.53)	13.01 (3.44)
	1-2 times	522	16.57 (2.45)	16.41 (2.43)
	A few times	764	18.59 (2.37)	19.33 (2.60)
	Many times	1,315	24.54 (2.31)	24.40 (2.31)
	Overall	2,830	20.70 (1.42)	20.81 (1.46)
Past Month	0 times	229	6.69 (3.10)	6.42 (2.98)
	1-2 times	522	8.23 (1.83)	7.88 (1.74)
	A few times	764	8.25 (1.46)	8.46 (1.51)
	Many times	1,315	9.87 (1.49)	9.81 (1.51)
	Overall	2,830	8.90 (0.90)	8.86 (0.91)
Cigarettes				
Lifetime	0 times	229	8.33 (3.39)	8.16 (3.32)
	1-2 times	522	6.48 (1.60)	6.17 (1.50)
	A few times	764	7.43 (1.38)	7.59 (1.43)
	Many times	1,315	11.71 (1.54)	11.85 (1.57)
	Overall	2,830	9.41 (0.91)	9.46 (0.92)
Past Year	0 times	229	1.00 (0.63)	1.07 (0.65)
	1-2 times	522	3.31 (1.30)	3.03 (1.16)
	A few times	764	4.50 (1.09)	4.55 (1.12)
	Many times	1,315	6.34 (1.03)	6.40 (1.04)
	Overall	2,830	4.92 (0.62)	4.93 (0.62)
Past Month	0 times	229	0.80 (0.61)	0.88 (0.63)
	1-2 times	522	1.41 (0.98)	1.13 (0.75)
	A few times	764	2.12 (0.62)	2.08 (0.62)
	Many times	1,315	3.46 (0.87)	3.47 (0.87)
	Overall	2,830	2.55 (0.48)	2.50 (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 7.6b 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: 2018

Drug	Father Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Alcohol				
Lifetime	0 times	239	13.48 (3.36)	13.96 (3.47)
	1-2 times	390	23.02 (3.75)	23.05 (3.68)
	A few times	524	33.23 (3.61)	33.34 (3.62)
	Many times	565	22.51 (2.97)	22.50 (2.98)
	Overall	1,718	24.62 (1.85)	24.72 (1.84)
Past Year	0 times	239	9.58 (3.01)	9.64 (2.99)
	1-2 times	390	19.14 (3.68)	19.15 (3.61)
	A few times	524	29.58 (3.65)	29.63 (3.65)
	Many times	565	17.98 (2.81)	18.12 (2.82)
	Overall	1,718	20.60 (1.78)	20.67 (1.77)
Past Month	0 times	239	3.70 (1.75)	3.54 (1.67)
	1-2 times	390	7.60 (2.41)	8.01 (2.52)
	A few times	524	11.62 (2.44)	11.82 (2.47)
	Many times	565	8.09 (2.33)	8.15 (2.34)
	Overall	1,718	8.42 (1.20)	8.58 (1.21)
Cigarettes				
Lifetime	0 times	239	4.00 (2.23)	3.94 (2.17)
	1-2 times	390	5.88 (1.44)	6.01 (1.48)
	A few times	524	9.91 (2.34)	10.04 (2.35)
	Many times	565	7.11 (1.89)	6.99 (1.85)
	Overall	1,718	7.24 (1.06)	7.26 (1.05)
Past Year	0 times	239	2.92 (2.15)	2.88 (2.09)
	1-2 times	390	3.01 (0.97)	3.07 (1.00)
	A few times	524	4.15 (1.05)	4.35 (1.12)
	Many times	565	4.64 (1.68)	4.60 (1.65)
	Overall	1,718	3.86 (0.74)	3.91 (0.74)
Past Month	0 times	239	0.20 (0.13)	0.26 (0.19)
	1-2 times	390	0.81 (0.45)	0.88 (0.51)
	A few times	524	1.42 (0.55)	1.45 (0.56)
	Many times	565	2.96 (1.58)	2.94 (1.55)
	Overall	1,718	1.59 (0.54)	1.61 (0.54)

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: 2018

Drug	Mother Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Any Illicit				
Lifetime	0 times	229	14.53 (3.41)	14.07 (3.31)
	1-2 times	522	14.41 (2.16)	14.20 (2.11)
	A few times	764	18.27 (2.43)	18.61 (2.59)
	Many times	1,315	27.33 (2.24)	27.30 (2.24)
	Overall	2,830	21.71 (1.37)	21.74 (1.40)
Past Year	0 times	229	8.52 (3.00)	8.37 (2.91)
	1-2 times	522	10.50 (2.02)	10.22 (1.96)
	A few times	764	12.85 (2.04)	13.32 (2.26)
	Many times	1,315	21.12 (2.10)	21.02 (2.09)
	Overall	2,830	16.10 (1.27)	16.15 (1.30)
Past Month	0 times	229	6.62 (2.89)	6.56 (2.80)
	1-2 times	522	6.21 (1.74)	5.98 (1.65)
	A few times	764	5.98 (1.37)	6.01 (1.37)
	Many times	1,315	11.86 (1.82)	11.83 (1.82)
	Overall	2,830	8.89 (1.05)	8.86 (1.05)
Marijuana				
Lifetime	0 times	229	7.45 (2.93)	7.36 (2.85)
	1-2 times	522	9.61 (1.96)	9.37 (1.88)
	A few times	764	10.86 (1.90)	11.37 (2.16)
	Many times	1,315	21.14 (2.14)	21.01 (2.13)
	Overall	2,830	15.33 (1.26)	15.38 (1.28)
Past Year	0 times	229	6.85 (2.88)	6.74 (2.80)
	1-2 times	522	8.33 (1.91)	8.01 (1.83)
	A few times	764	9.72 (1.86)	10.22 (2.12)
	Many times	1,315	17.81 (2.05)	17.66 (2.04)
	Overall	2,830	13.16 (1.22)	13.18 (1.24)
Past Month	0 times	229	5.98 (2.86)	5.94 (2.78)
	1-2 times	522	5.47 (1.71)	5.15 (1.60)
	A few times	764	5.25 (1.32)	5.31 (1.32)
	Many times	1,315	10.62 (1.81)	10.59 (1.80)
	Overall	2,830	7.92 (1.03)	7.88 (1.03)

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: 2018

Drug	Father Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Any Illicit				
Lifetime	0 times	239	18.54 (4.00)	18.68 (4.04)
	1-2 times	390	27.88 (3.95)	27.80 (3.89)
	A few times	524	24.21 (3.06)	24.37 (3.08)
	Many times	565	21.75 (2.91)	21.86 (2.91)
	Overall	1,718	23.47 (1.71)	23.57 (1.70)
Past Year	0 times	239	10.06 (2.75)	9.87 (2.70)
	1-2 times	390	17.67 (3.42)	17.86 (3.36)
	A few times	524	15.40 (2.62)	15.70 (2.66)
	Many times	565	12.29 (2.41)	12.32 (2.40)
	Overall	1,718	14.18 (1.45)	14.31 (1.45)
Past Month	0 times	239	3.25 (1.20)	3.44 (1.33)
	1-2 times	390	8.43 (3.08)	8.34 (2.97)
	A few times	524	8.21 (2.27)	8.33 (2.28)
	Many times	565	7.03 (2.10)	6.92 (2.06)
	Overall	1,718	7.17 (1.23)	7.19 (1.22)
Marijuana				
Lifetime	0 times	239	7.62 (2.47)	7.62 (2.46)
	1-2 times	390	13.90 (3.22)	13.93 (3.12)
	A few times	524	14.10 (2.58)	14.28 (2.60)
	Many times	565	11.04 (2.34)	11.07 (2.32)
	Overall	1,718	12.15 (1.35)	12.23 (1.33)
Past Year	0 times	239	6.41 (2.37)	6.43 (2.36)
	1-2 times	390	12.14 (3.21)	12.23 (3.13)
	A few times	524	11.31 (2.48)	11.49 (2.50)
	Many times	565	9.90 (2.26)	10.02 (2.25)
	Overall	1,718	10.35 (1.35)	10.47 (1.35)
Past Month	0 times	239	1.65 (0.85)	1.85 (1.04)
	1-2 times	390	7.11 (3.07)	7.13 (2.96)
	A few times	524	6.91 (2.23)	7.04 (2.24)
	Many times	565	6.78 (2.10)	6.66 (2.05)
	Overall	1,718	6.16 (1.22)	6.20 (1.20)

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

References

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

Center for Behavioral Health Statistics and Quality. (2014). *2013 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. (2017). *How to prepare and analyze pair data in the National Survey on Drug Use and Health.* Rockville, MD: Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (2019). *2018 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. (2020a). *2018 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. (2020b). *2018 National Survey on Drug Use and Health: Methodological resource book (Section 11, Person-level sampling weight calibration)*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

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

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

Chromy, J. R., & Singh, A. C. (2001). Estimation for person-pair drug-related characteristics in the presence of pair multiplicities and extreme sampling weights. In *Proceedings of the 2001 Joint Statistical Meetings*, *American Statistical Association, Survey Research Methods Section, Atlanta, GA* [CD-ROM]. Alexandria, VA: American Statistical Association. Retrieved from https://www.amstat.org/ASA/Membership/Sections-and-Interest-Groups.aspx

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

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

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

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

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

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

List of Contributors

This methodological report was prepared by the Substance Abuse and Mental Health Services Administration (SAMHSA), Center for Behavioral Health Statistics and Quality, and by RTI International (a registered trademark and a trade name of Research Triangle Institute). Work by RTI was performed under Contract No. HHSS2832017000002C. Kathryn Piscopo served as government project officer and as the contracting officer representative, and David Hunter served as the RTI project director.

The contributor to this report at SAMHSA was Yang Cheng.

Contributors to this report at RTI included Patrick Chen, Harper Gordek, Jeff Laufenberg, and Matthew Westlake. Also at RTI, report and web production staff included Teresa Bass, Debbie Bond, and Margaret Johnson.

This page intentionally left blank

Appendix A: Technical Details about the Generalized Exponential Model

This page intentionally left blank

Appendix A: Technical Details about the Generalized Exponential Model

A.1 Distance Function

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

A.3 Newton-Raphson Steps

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

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

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

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

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

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

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

A.4 Scaled Constrained Exponential Model

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

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

This page intentionally left blank

Appendix B: Derivation of Poststratification Control Totals

This page intentionally left blank

Appendix B: Derivation of Poststratification Control Totals

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

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

B.1 Derivation of QDU-Level Poststratification Controls

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

B.1.1 Person Level

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

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

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

B.1.1.2 Adjusting SDU Data to the Control Totals

In the person-level weighting, the SDU weights were poststratified to meet control totals based on the population estimates received from the census. For NSDUH weighting, a generalized exponential model (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, 2020a) 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 2018, 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 2018 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 2018 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	1.93	4.82	1.32	2.4688	243	(0.55, 2.00)	(0.56, 2.00)
	1.76	4.02	0.71	2.3882	243	(0.34, 2.10)	(0.35, 2.07)
						(0.90, 1.63)	(0.90, 1.63)
res.qdu.nr	1.74	4.03	0.79	2.4412	243	(1.00, 1.90)	(1.00, 1.90)
	1.61	4.04	0.53	2.5425	243	(1.00, 3.30)	(1.00, 3.30)
						(1.40, 3.38)	(1.40, 3.38)
res.qdu.ps	1.61	4.04	0.53	2.5425	243	(0.95, 1.20)	(0.99, 1.20)
	1.65	4.25	0.30	2.5418	243	(0.86, 1.20)	(0.87, 1.16)
						(0.99, 1.10)	(0.99, 1.01)

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 2018 Distribution of Weight Adjustment Factors and Weight Products (Model Group 1: Northeast)

	SDU Weight	QDU Design Weight		sel.qdu.ps[1]		res.qdu.nr[1]		res.qdu.ps[1]	
Statistics	**1-11**	**duwght12**	**1-12**	**duwght13**	**1-13**	**duwght14**	**1-14**	**duwght15**	**1-15**
Minimum	20	1.00	33	0.35	33	0.55	51	0.65	51
1%	78	1.00	81	0.68	82	1.01	107	0.95	109
5%	111	1.00	155	0.83	157	1.07	204	0.98	203
10%	158	1.00	217	0.88	214	1.15	298	0.99	298
25%	280	1.00	495	0.94	477	1.28	625	1.00	624
Median	748	1.30	1,002	1.00	1,007	1.43	1,342	1.00	1,342
75%	989	2.03	1,740	1.07	1,774	1.59	2,616	1.01	2,615
90%	1,290	4.81	3,446	1.15	3,445	1.77	5,125	1.01	5,159
95%	1,528	6.10	5,004	1.22	5,014	1.93	7,722	1.02	7,712
99%	2,513	9.71	9,179	1.41	9,072	2.43	14,242	1.05	14,168
Maximum	6,632	13.27	37,296	2.44	26,362	3.38	34,015	1.17	34,496
n	14,264	-	14,264	-	14,264	-	9,753	-	9,753
Mean	742	2.11	1,529	1.01	1,534	1.46	2,244	1.00	2,244
Max/Mean	8	-	23	-	20	-	19	-	17

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

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

Model Group 1 Overview

Selected Questionnaire Dwelling Unit-Level Poststratification

All 243 proposed effects, were kept in the model.

Respondent Questionnaire Dwelling Unit-Level Nonresponse

This step used the same set of 243 proposed effects as the selected questionnaire dwelling unit-level poststratification.

Respondent Questionnaire Dwelling Unit-Level Poststratification

This step used the same set of 243 proposed effects as the selected questionnaire dwelling unit-level poststratification and the respondent questionnaire dwelling unit-level nonresponse.

Exhibit C.1.1 Covariates for 2018 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 2018 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 1: Northeast

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

Exhibit C.1.3 Covariates for 2018 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 1: Northeast

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

Appendix C.2: Model Group 2: Midwest

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

This page intentionally left blank

Table C.2a 2018 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	1.51	3.06	0.76	2.0676	300	(0.71, 2.61)	(0.72, 2.61)
	1.27	2.40	0.40	2.0948	300	(0.64, 2.61)	(0.65, 2.61)
						(0.90, 1.42)	(0.90, 1.42)
res.qdu.nr	0.99	1.64	0.29	2.1279	300	(1.00, 1.50)	(1.00, 1.50)
	1.11	1.73	0.21	2.1952	300	(1.00, 3.12)	(1.00, 3.12)
						(1.40, 2.03)	(1.40, 2.03)
res.qdu.ps	1.11	1.73	0.21	2.1952	300	(0.98, 1.20)	(0.99, 1.20)
	1.07	1.57	0.10	2.1957	300	(0.88, 1.20)	(0.89, 1.14)
						(0.99, 1.02)	(0.99, 1.02)

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 $1 + \left[\left(n - 1\right)/n\right] * CV^2$, 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 2018 Distribution of Weight Adjustment Factors and Weight Products (Model Group 2: Midwest)

	SDU Weight	QDU Design Weight		sel.qdu.ps[1]		res.qdu.nr[1]		res.qdu.ps[1]	
Statistics	**1-11**	**duwght12**	**1-12**	**duwght13**	**1-13**	**duwght14**	**1-14**	**duwght15**	**1-15**
Minimum	26	1.00	26	0.27	20	0.47	22	0.56	21
1%	88	1.00	103	0.75	100	1.00	136	0.97	137
5%	139	1.00	204	0.83	204	1.08	289	0.99	289
10%	196	1.00	332	0.88	327	1.16	432	1.00	433
25%	552	1.00	729	0.93	714	1.27	918	1.00	918
Median	824	1.26	1,121	0.99	1,121	1.38	1,508	1.00	1,508
75%	1,053	2.02	1,847	1.06	1,854	1.50	2,580	1.00	2,583
90%	1,371	4.29	3,629	1.14	3,593	1.66	5,241	1.01	5,257
95%	1,609	5.97	5,157	1.20	5,260	1.77	7,730	1.01	7,758
99%	2,324	8.29	8,511	1.35	8,649	2.12	12,244	1.03	12,257
Maximum	8,426	10.92	17,243	2.61	19,150	3.12	27,120	1.42	27,169
n	16,550	-	16,550	-	16,550	-	11,810	-	11,810
Mean	839	2.01	1,646	1.00	1,651	1.41	2,314	1.10	2,314
Max/Mean	10	-	10	-	12	-	12	-	12

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

This step used the same set of 300 proposed effects as the selected questionnaire dwelling unit-level poststratification.

Respondent Questionnaire Dwelling Unit-Level Poststratification

This step used the same set of 300 proposed effects as the selected questionnaire dwelling unit-level poststratification and the respondent questionnaire dwelling unit-level nonresponse.

Exhibit C.2.1 Covariates for 2018 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 2018 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 2: Midwest

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

Exhibit C.2.3 Covariates for 2018 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 2: Midwest

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

Appendix C.3: Model Group 3: South

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

This page intentionally left blank

Table C.3a 2018 QDU Weight GEM Modeling Summary (Model Group 3: South)

Modeling Step[1]	Extreme Weight Proportions			UWE[2]	# Covariates[3]	Bounds[4]	
	% Unweighted	% Weighted	% Outwinsor			Nominal	Realized
sel.qdu.ps	0.96	1.69	0.33	1.9366	339	(0.66, 2.30)	(0.67, 2.30)
	0.83	1.43	0.20	1.9567	339	(0.50, 2.57)	(0.50, 2.57)
						(0.90, 1.29)	(0.90, 1.29)
res.qdu.nr	0.82	1.66	0.23	1.9696	339	(1.00, 1.40)	(1.00, 1.40)
	0.45	0.90	0.11	2.0772	338	(1.00, 3.72)	(1.00, 3.72)
						(1.20, 2.29)	(1.20, 2.29)
res.qdu.ps	0.45	0.90	0.11	2.0772	339	(0.98, 1.15)	(0.99, 1.15)
	0.54	1.11	0.07	2.0767	338	(0.94, 1.15)	(0.95, 1.13)
						(0.99, 1.03)	(0.99, 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 2018 Distribution of Weight Adjustment Factors and Weight Products (Model Group 3: South)

	SDU Weight	QDU Design Weight		sel.qdu.ps[1]		res.qdu.nr[1]		res.qdu.ps[1]	
Statistics	**1-11**	**duwght12**	**1-12**	**duwght13**	**1-13**	**duwght14**	**1-14**	**duwght15**	**1-15**
Minimum	20	1.00	20	0.44	17	0.64	21	0.59	21
1%	54	1.00	76	0.76	76	1.02	88	0.98	88
5%	113	1.00	212	0.85	211	1.08	273	0.99	273
10%	260	1.00	384	0.88	381	1.13	512	1.00	513
25%	705	1.00	904	0.94	892	1.22	1,130	1.00	1,131
Median	1,065	1.27	1,501	1.00	1,498	1.33	1,921	1.00	1,924
75%	1,457	2.22	2,546	1.06	2,561	1.44	3,376	1.00	3,377
90%	1,886	4.26	4,702	1.12	4,717	1.56	6,434	1.00	6,432
95%	2,206	5.72	6,493	1.17	6,589	1.65	9,203	1.01	9,199
99%	2,951	8.90	10,185	1.31	10,394	1.95	14,760	1.03	14,768
Maximum	8,909	10.45	22,096	2.57	25,254	3.72	32,749	1.15	32,780
n	22,716	-	22,716	-	22,716	-	16,931	-	16,931
Mean	1,106	2.03	2,129	1.00	2,136	1.34	2,866	1.00	2,866
Max/Mean	8	-	10	-	12	-	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 3 Overview

Selected Questionnaire Dwelling Unit-Level Poststratification

All 339 proposed effects were kept in the model.

Respondent Questionnaire Dwelling Unit-Level Nonresponse

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

Respondent Questionnaire Dwelling Unit-Level Poststratification

This step used the same set of 338 effects as the respondent questionnaire dwelling unit-level nonresponse.

Exhibit C.3.1 Covariates for 2018 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 2018 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**	**212**	
Age × Race (3 Levels)	5 × 3	8	8	All levels present.
Age × Hispanicity	5 × 2	4	4	All levels present.
Age × Gender	5 × 2	4	4	All levels present.
Race (3 Levels) × Hispanicity	3 × 2	2	2	All levels present.
Race (3 Levels) × Gender	3 × 2	2	2	All levels present.
Hispanicity × Gender	2 × 2	1	1	All levels present.
State × Age	17 × 5	64	64	All levels present.
State × Race	17 × 5	64	63	Coll (3,3) & (3,4); conv.
State × Gender	17 × 2	16	16	All levels present.
State × Hispanicity	17 × 2	16	16	All levels present.
% Black or African American × % Owner-Occupied	3 × 3	4	4	All levels present.
% Black or African American × Rent/Housing	3 × 5	8	8	All levels present.
% Hispanicity × % Owner-Occupied	3 × 3	4	4	All levels present.
% Hispanicity × Rent/Housing	3 × 5	8	8	All levels present.
% Owner-Occupied × Rent/Housing	3 × 5	8	8	All levels present.
Three-Factor Effects		**50**	**50**	
Race (3 Levels) × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Hispanicity	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Race (3 Levels)	3 × 5 × 3	16	16	All levels present.
State/Region × Hispanicity × Gender	3 × 2 × 2	2	2	All levels present.
State/Region × Race (3 Levels) × Hispanicity	3 × 3 × 2	4	4	All levels present.
State/Region × Race (3 Levels) × Gender	3 × 3 × 2	4	4	All levels present.
Total		**339**	**338**	

Exhibit C.3.3 Covariates for 2018 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 3: South

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

Appendix C.4: Model Group 4: West

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

This page intentionally left blank

Table C.4a 2018 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	1.28	3.00	0.79	2.2974	270	(0.65, 1.10)	(0.65, 1.10)
	0.98	2.09	0.25	2.3212	266	(0.55, 2.68)	(0.56, 2.68)
						(0.95, 1.05)	(0.95, 1.05)
res.qdu.nr	1.07	2.36	0.34	2.3903	270	(1.00, 2.30)	(1.00, 2.30)
	1.11	2.73	0.29	2.5545	266	(1.00, 3.63)	(1.00, 3.63)
						(1.30, 1.50)	(1.30, 1.30)
res.qdu.ps	1.11	2.73	0.29	2.5545	270	(0.90,1.70)	(0.97, 1.70)
	1.12	2.60	0.12	2.5505	266	(0.88, 1.70)	(.89, 1.27)
						(0.95, 1.10)	(0.95, 1.00)

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

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

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

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

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

Table C.4b 2018 Distribution of Weight Adjustment Factors and Weight Products (Model Group 4: West)

	SDU Weight	QDU Design Weight		sel.qdu.ps[1]		res.qdu.nr[1]		res.qdu.ps[1]	
Statistics	**1-11**	**duwght12**	**1-12**	**duwght13**	**1-13**	**duwght14**	**1-14**	**duwght15**	**1-15**
Minimum	24	1.00	24	0.33	28	0.69	28	0.72	28
1%	93	1.00	103	0.75	101	1.01	116	0.95	115
5%	122	1.00	148	0.84	150	1.08	176	0.99	177
10%	143	1.00	201	0.88	202	1.14	250	0.99	251
25%	284	1.00	467	0.94	463	1.24	580	1.00	580
Median	822	1.24	1,225	1.00	1,215	1.36	1,518	1.00	1,519
75%	1,438	2.06	2,136	1.07	2,116	1.50	2,905	1.00	2,907
90%	1,438	3.63	3,759	1.14	3,751	1.67	5,411	1.01	5,397
95%	1,794	5.35	5,510	1.20	5,683	1.78	8,478	1.02	8,464
99%	2,036	8.27	9,873	1.38	10,166	2.08	15,171	1.03	15,238
Maximum	7,417	12.08	37,543	2.68	24,178	3.63	39,662	1.66	33,675
n	16,518	-	16,518	-	16,518	-	11,879	-	11,879
Mean	936	1.91	1,718	1.01	1,732	1.39	2,409	1.00	2,409
Max/Mean	8	-	22	-	14	-	16	-	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, 266 were kept in the model. All one-factor effects were maintained in full. Two-factor effects were modified for percent Black or African American × Rent/Housing, combining 50-100 percent and 10-<50 percent for levels 1 and 4 of percent Black or African American. Also combined were 10-<50 percent and 50-100 percent Black or African American for the 0-<10 percent and 10-<50 percent levels of percent Owner-Occupied. All three-factor effects were maintained in full.

Respondent Questionnaire Dwelling Unit-Level Nonresponse

This step used the same set of 266 effects as the selected questionnaire dwelling unit-level poststratification.

Respondent Questionnaire Dwelling Unit-Level Poststratification

This step used the same set of 266 effects as the selected questionnaire dwelling unit-level poststratification and respondent questionnaire dwelling unit-level nonresponse.

Exhibit C.4.1 Covariates for 2018 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**	**169**	
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. (2,1) & (2,2), (3,1) & (3,2); sing.
% Black or African American × Rent/Housing	3 × 5	8	6	Coll. (1,1) & (1,2); zero. Coll. (4,1) & (4,2); 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**	**266**	

Exhibit C.4.2 Covariates for 2018 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 2018 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 and respondent questionnaire dwelling unit-level nonresponse.

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

This page intentionally left blank

Table D.1 2018 NSDUH QDU-Level Response Rates

Domain	Selected QDU	Respondent QDU	% Interview Response Rate[1]
Total	70,048	50,373	70.08
Census Region			
Northeast	14,264	9,753	66.58
South	22,716	16,931	72.73
Midwest	16,550	11,810	69.93
West	16,518	11,879	68.41
Quarter			
Quarter 1	16,464	11,898	70.80
Quarter 2	18,700	13,457	70.08
Quarter 3	17,653	12,763	70.50
Quarter 4	17,231	12,255	68.94
Household Type			
12-17, 18-25, 26+	5,244	4,097	77.98
12-17, 18-25	63	50	80.76
12-17, 26+	15,469	11,843	76.90
18-25, 26+	11,895	8,541	71.60
12-17	16	8	54.14
18-25	5,519	4,152	75.09
26+	31,842	21,682	68.01
Race/Ethnicity of Householder			
Hispanic or Latino White	10,069	7,378	71.00
Hispanic or Latino Black or African American	268	199	78.62
Hispanic or Latino Other	653	508	72.49
Non-Hispanic or Latino White	44,886	31,759	69.24
Non-Hispanic or Latino Black or African American	8,110	6,271	75.24
Non-Hispanic or Latino Other	6,062	4,258	66.14
% Hispanic or Latino in Segment			
50-100%	5,251	3,872	71.13
10-<50%	17,621	12,685	69.52
<10%	47,176	33,816	70.20
% Black or African American in Segment			
50-100%	4,944	3,849	75.62
10-<50%	13,958	10,269	71.80
<10%	51,146	36,255	68.97
% Owner-Occupied DUs in Segment			
50-100%	51,498	36,731	69.59
10-<50%	14,841	10,871	71.06
<10%	3,709	2,771	72.61
Combined Median Rent/Housing Value			
1st Quintile	11,376	8,561	74.02
2nd Quintile	15,287	11,254	71.65
3rd Quintile	16,646	11,930	70.24
4th Quintile	15,067	10,608	68.61
5th Quintile	11,672	8,020	66.88
Population Density			
Large MSA	30,559	21,519	68.68
Medium to Small MSA	34,000	24,877	71.71
Non-MSA, Urban	1,706	1,280	73.56
Non-MSA, Rural	3,783	2,697	70.12
Group Quarters			
Group	555	516	93.25
Non-Group	69,493	49,857	69.94
Household Size			
One	9,371	6,759	71.07
Two	29,245	20,364	68.06
Three	17,228	12,493	70.98
Four or More	14,204	10,757	74.63

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 2018 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	70,048	1.88	4.26	1.10	1.36	2.82	0.70	1.16	2.24	0.34
Census Region										
Northeast	14,264	2.69	7.24	2.12	1.93	4.82	1.32	1.76	4.02	0.71
South	22,716	1.47	2.88	0.60	0.96	1.69	0.33	0.83	1.43	0.20
Midwest	16,550	1.90	4.32	1.10	1.51	3.06	0.76	1.27	2.40	0.40
West	16,518	1.73	4.41	1.21	1.28	3.00	0.79	0.98	2.09	0.25
Quarter										
Quarter 1	16,464	2.39	5.15	1.38	1.64	3.20	0.84	1.60	2.72	0.40
Quarter 2	18,700	1.41	3.22	0.85	1.02	2.16	0.46	0.84	1.73	0.32
Quarter 3	17,653	2.05	4.75	1.18	1.43	3.14	0.76	1.05	2.16	0.32
Quarter 4	17,231	1.74	3.90	0.98	1.39	2.80	0.72	1.20	2.34	0.33
Household Type										
12-17, 18-25, 26+	5,244	1.62	4.41	1.28	1.62	4.41	1.28	1.49	3.45	0.84
12-17, 18-25	63	1.59	0.76	0.32	0.00	0.00	0.00	0.00	0.00	0.00
12-17, 26+	15,469	1.73	3.99	1.06	1.76	4.06	1.07	1.49	3.79	0.66
18-25, 26+	11,895	1.77	4.62	1.27	1.61	4.28	1.18	1.47	3.48	0.63
12-17	16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18-25	5,519	1.92	4.49	1.21	1.72	3.94	1.09	1.79	3.87	0.45
26+	31,842	2.04	4.19	1.00	0.97	2.25	0.51	0.72	1.63	0.21
Race/Ethnicity of Householder										
Hispanic or Latino White	10,069	1.05	1.85	0.48	0.58	1.12	0.31	0.64	1.31	0.19
Hispanic or Latino Black or African American	268	52.61	76.07	30.86	36.94	51.29	18.31	34.70	46.24	9.58
Hispanic or Latino Other	653	20.52	44.76	13.72	12.56	26.12	8.25	9.19	18.33	3.04
Non-Hispanic or Latino White	44,886	0.88	1.98	0.44	0.80	1.71	0.35	0.62	1.18	0.13
Non-Hispanic or Latino Black or African American	8,110	3.27	5.43	1.14	2.12	3.04	0.71	1.85	2.80	0.49
Non-Hispanic or Latino Other	6,062	4.55	9.52	1.86	3.00	6.61	1.37	2.72	5.63	0.92

Table E.1 2018 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%	5,251	1.94	4.56	1.37	1.31	3.08	0.84	1.71	3.66	0.56
10-<50%	17,621	2.18	4.75	1.29	1.53	3.04	0.85	1.25	2.36	0.37
<10%	47,176	1.76	3.96	0.95	1.30	2.69	0.61	1.06	2.01	0.31
% Black or African American in Segment										
50-100%	4,944	3.05	6.78	1.48	1.92	3.61	0.79	1.62	2.67	0.47
10-<50%	13,958	2.72	5.68	1.48	1.77	3.44	0.95	1.55	2.78	0.49
<10%	51,146	1.54	3.52	0.93	1.19	2.55	0.61	1.01	2.02	0.29
% Owner-Occupied DUs in Segment										
50-100%	51,498	1.26	2.70	0.70	1.06	2.04	0.45	0.82	1.53	0.23
10-<50%	14,841	3.44	7.77	2.13	2.06	4.80	1.33	1.95	3.95	0.64
<10%	3,709	4.26	9.21	1.86	2.72	5.14	1.41	2.67	4.60	0.61
Combined Median Rent/Housing Value										
1st Quintile	11,376	1.84	3.76	0.86	1.01	1.59	0.33	0.76	1.35	0.21
2nd Quintile	15,287	1.83	3.36	0.79	1.31	2.19	0.45	1.22	2.04	0.39
3rd Quintile	16,646	1.66	3.89	0.96	1.22	2.46	0.61	0.93	1.86	0.29
4th Quintile	15,067	1.65	4.04	1.21	1.27	2.95	0.83	1.16	2.25	0.34
5th Quintile	11,672	2.61	6.26	1.61	2.08	4.76	1.19	1.79	3.56	0.47
Population Density										
Large MSA[1]	30,559	2.49	5.63	1.52	1.92	4.06	1.07	1.69	3.09	0.47
Medium to Small MSA[1]	34,000	1.46	2.76	0.63	0.96	1.49	0.30	0.79	1.36	0.22
Non-MSA,[1] Urban	1,706	1.29	2.38	0.62	0.59	0.95	0.23	0.70	1.30	0.20
Non-MSA,[1] Rural	3,783	1.03	1.04	0.12	0.79	1.26	0.17	0.34	0.39	0.02
Group Quarters										
Group	555	3.96	6.81	1.05	3.96	6.02	0.71	3.60	5.77	0.43
Non-Group	69,493	1.87	4.24	1.10	1.34	2.80	0.70	1.14	2.21	0.34
Household Size										
One	9,371	1.77	4.00	0.91	0.50	1.23	0.31	0.45	0.83	0.11
Two	29,245	1.84	4.05	1.08	1.35	3.11	0.72	1.09	2.32	0.31
Three	17,228	1.79	4.01	0.99	1.50	3.44	0.90	1.32	3.22	0.56
Four or More	14,204	2.15	5.12	1.37	1.77	4.58	1.24	1.57	3.85	0.70

[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 2018 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	50,373	1.10	2.20	0.36	0.99	2.04	0.25	1.02	2.09	0.13
Census Region										
Northeast	9,753	1.74	4.03	0.79	1.61	4.04	0.53	1.65	4.25	0.30
South	16,931	0.82	1.66	0.23	0.45	0.90	0.11	0.54	1.11	0.07
Midwest	11,810	0.99	1.64	0.29	1.11	1.73	0.21	1.07	1.57	0.10
West	11,879	1.07	2.36	0.34	1.11	2.73	0.29	1.12	2.60	0.12
Quarter										
Quarter 1	11,898	1.65	3.32	0.48	1.23	2.23	0.26	1.26	2.28	0.15
Quarter 2	13,457	0.72	1.48	0.31	0.79	1.62	0.25	0.79	1.65	0.13
Quarter 3	12,763	0.92	1.91	0.31	1.00	2.40	0.26	1.13	2.61	0.13
Quarter 4	12,255	1.15	2.10	0.34	0.96	1.93	0.21	0.91	1.83	0.10
Household Type										
12-17, 18-25, 26+	4,097	1.39	3.06	0.68	1.66	3.86	0.62	1.64	3.58	0.27
12-17, 18-25	50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12-17, 26+	11,843	1.53	3.63	0.62	1.15	2.91	0.39	1.20	3.03	0.20
18-25, 26+	8,541	1.45	3.60	0.61	1.28	3.31	0.42	1.35	3.49	0.26
12-17	8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18-25	4,152	1.57	3.19	0.36	1.71	3.93	0.59	1.73	3.52	0.24
26+	21,682	0.58	1.60	0.25	0.52	1.48	0.15	0.54	1.55	0.08
Race/Ethnicity of Householder										
Hispanic or Latino White	7,378	0.66	1.44	0.26	0.62	1.62	0.18	0.61	1.71	0.09
Hispanic or Latino Black or African American	199	35.18	46.12	10.33	18.59	22.78	2.14	21.61	27.95	1.25
Hispanic or Latino Other	508	8.86	16.61	2.76	6.30	12.79	1.75	7.09	15.99	1.23
Non-Hispanic or Latino White	31,759	0.55	1.16	0.15	0.52	0.90	0.10	0.48	0.81	0.04
Non-Hispanic or Latino Black or African American	6,271	1.79	2.60	0.49	1.05	1.95	0.19	1.12	1.98	0.08
Non-Hispanic or Latino Other	4,258	2.37	5.06	0.71	3.55	9.56	1.38	3.92	9.79	0.76
% Hispanic or Latino in Segment										
50-100%	3,872	1.68	3.61	0.58	1.19	2.50	0.28	1.24	2.88	0.15
10-<50%	12,685	1.17	2.17	0.38	1.21	2.57	0.28	1.25	2.57	0.15
<10%	33,816	1.00	2.04	0.32	0.88	1.75	0.23	0.90	1.77	0.11

Table E.2 2018 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%	3,849	1.45	2.52	0.51	1.17	2.46	0.27	1.22	2.54	0.14
10-<50%	10,269	1.42	2.54	0.46	1.20	2.43	0.29	1.29	2.62	0.15
<10%	36,255	0.97	2.06	0.31	0.91	1.88	0.23	0.92	1.88	0.12
% Owner-Occupied DUs in Segment										
50-100%	36,731	0.79	1.54	0.25	0.75	1.43	0.18	0.76	1.47	0.10
10-<50%	10,871	1.80	3.86	0.65	1.33	3.09	0.37	1.34	3.09	0.15
<10%	2,771	2.42	4.08	0.59	2.81	5.91	0.66	3.14	6.35	0.40
Combined Median Rent/Housing Value										
1st Quintile	8,561	0.82	1.57	0.23	0.58	1.14	0.14	0.55	1.02	0.08
2nd Quintile	11,254	1.14	2.36	0.48	0.92	1.66	0.18	0.95	1.66	0.09
3rd Quintile	11,930	0.97	1.84	0.30	0.98	1.92	0.22	0.99	2.03	0.12
4th Quintile	10,608	0.96	1.96	0.33	0.90	1.89	0.22	0.95	1.96	0.13
5th Quintile	8,020	1.70	3.29	0.43	1.62	3.47	0.46	1.73	3.60	0.21
Population Density										
Large MSA[1]	21,519	1.59	3.12	0.50	1.37	2.77	0.32	1.43	2.90	0.17
Medium to Small MSA[1]	24,877	0.78	1.32	0.22	0.75	1.30	0.16	0.78	1.26	0.07
Non-MSA,[1] Urban	1,280	0.55	0.76	0.21	0.23	0.31	0.09	0.23	0.33	0.10
Non-MSA,[1] Rural	2,697	0.30	0.30	0.03	0.48	0.89	0.19	0.30	0.73	0.08
Group Quarters										
Group	516	3.10	5.17	0.48	0.58	0.74	0.04	0.97	1.11	0.02
Non-Group	49,857	1.08	2.18	0.36	0.99	2.05	0.25	1.02	2.09	0.13
Household Size										
One	6,759	0.52	0.95	0.14	0.40	0.93	0.09	0.34	0.96	0.04
Two	20,364	0.91	2.23	0.33	0.91	2.14	0.25	0.92	2.11	0.12
Three	12,493	1.34	3.09	0.59	1.03	2.62	0.33	1.13	2.88	0.21
Four or More	10,757	1.52	3.71	0.66	1.44	3.46	0.47	1.50	3.56	0.25

[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 2018 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	50,373	126,352,219	126,352,219	126,352,219	0.00	0.00
Census Region						
Northeast	9,753	21,883,077	21,883,077	21,883,077	0.00	-0.00
South	16,931	48,528,458	48,528,458	48,528,458	0.00	-0.00
Midwest	11,810	27,328,021	27,328,021	27,328,021	-0.00	-0.00
West	11,879	28,612,663	28,612,663	28,612,663	-0.00	0.00
Quarter						
Quarter 1	11,898	31,353,493	31,353,493	31,353,493	0.00	-0.00
Quarter 2	13,457	31,845,863	31,845,863	31,845,863	-0.00	0.00
Quarter 3	12,763	31,520,252	31,520,252	31,520,252	0.00	0.00
Quarter 4	12,255	31,632,611	31,632,611	31,632,611	-0.00	0.00
Household Type						
12-17, 18-25, 26+	4,097	5,135,464	5,135,464	5,135,464	0.00	0.00
12-17, 18-25	50	59,429	59,429	59,429	0.00	0.00
12-17, 26+	11,843	13,508,994	13,508,994	13,508,994	0.00	0.00
18-25, 26+	8,541	13,571,747	13,571,747	13,571,747	0.00	0.00
12-17	8	11,777	11,777	11,777	0.00	0.00
18-25	4,152	5,777,633	5,777,633	5,777,633	-0.00	-0.00
26+	21,682	88,287,174	88,287,174	88,287,174	0.00	0.00
Race/Ethnicity of Householder						
Hispanic or Latino White	7,378	15,643,617	15,643,617	15,643,617	-0.00	0.00
Hispanic or Latino Black or African American	199	938,356	938,356	938,356	-0.00	0.00
Hispanic or Latino Other	508	1,296,791	1,296,791	1,296,791	-0.00	0.00
Non-Hispanic or Latino White	31,759	83,426,262	83,426,262	83,426,262	0.00	0.00
Non-Hispanic or Latino Black or African American	6,271	15,740,220	15,740,220	15,740,220	0.00	-0.00
Non-Hispanic or Latino Other	4,258	9,306,973	9,306,973	9,306,973	0.00	0.00
% Hispanic or Latino in Segment						
50-100%	3,872	9,888,926	9,888,926	9,888,926	-0.00	0.00
10-<50%	12,685	36,213,751	36,213,751	36,213,751	-0.00	0.00
<10%	33,816	80,249,541	80,249,541	80,249,541	0.00	0.00
% Black or African American in Segment						
50-100%	3,849	9,377,526	9,377,526	9,377,526	0.00	-0.00
10-<50%	10,269	27,448,586	27,448,586	27,448,586	0.00	0.00
<10%	36,255	89,526,106	89,526,106	89,526,106	0.00	0.00
% Owner-Occupied DUs in Segment						
50-100%	36,731	91,475,353	91,475,353	91,475,353	0.00	0.00
10-<50%	10,871	28,143,153	28,143,153	28,143,153	-0.00	0.00
<10%	2,771	6,733,713	6,733,713	6,733,713	0.00	0.00

Table F.1 2018 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	8,561	18,442,361	18,442,361	18,442,361	0.00	-0.00
2nd Quintile	11,254	26,583,193	26,583,193	26,583,193	-0.00	0.00
3rd Quintile	11,930	28,836,296	28,836,296	28,836,296	0.00	0.00
4th Quintile	10,608	28,379,056	28,379,056	28,379,056	0.00	0.00
5th Quintile	8,020	24,111,312	24,111,312	24,111,312	0.00	0.00
Population Density						
Large MSA	21,519	66,818,029	66,818,029	66,818,029	-0.00	0.00
Medium to Small MSA	24,877	51,770,877	51,770,877	51,770,877	0.00	0.00
Non-MSA, Urban	1,280	2,547,394	2,547,394	2,547,394	0.00	0.00
Non-MSA, Rural	2,697	5,215,918	5,215,918	5,215,918	0.00	-0.00
Group Quarters						
Group	516	733,803	733,803	733,803	0.00	0.00
Non-Group	49,857	125,618,416	125,618,416	125,618,416	0.00	0.00
Household Size						
One	6,759	34,332,894	34,340,550	34,133,348	0.58	0.61
Two	20,364	57,783,574	57,759,608	57,980,733	-0.34	-0.38
Three	12,493	19,479,364	19,491,262	19,604,907	-0.64	-0.58
Four or More	10,757	14,756,387	14,760,799	14,633,230	0.84	0.87

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 2018 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	70,048	20	476	872	1,265	8,909	1.45	20	670	1,234	2,130	37,543	2.15	17	661	1,232	2,137	26,362	2.16
Census Region																			
Northeast	14,264	20	280	748	989	6,632	1.49	33	495	1,002	1,740	37,296	2.47	33	477	1,007	1,774	26,362	2.39
South	22,716	20	705	1,065	1,457	8,909	1.34	20	904	1,501	2,546	22,096	1.94	17	892	1,498	2,561	25,254	1.96
Midwest	16,550	26	552	824	1,053	8,426	1.32	26	729	1,121	1,847	17,243	2.07	20	714	1,121	1,854	19,150	2.09
West	16,518	24	284	822	1,438	7,417	1.59	24	467	1,225	2,136	37,543	2.30	28	463	1,215	2,116	24,178	2.32
Quarter																			
Quarter 1	16,464	20	543	959	1,331	8,909	1.42	24	736	1,316	2,265	22,096	2.07	19	740	1,333	2,281	25,254	2.08
Quarter 2	18,700	20	458	815	1,157	7,727	1.43	20	639	1,159	1,971	31,467	2.14	17	633	1,158	1,981	25,446	2.17
Quarter 3	17,653	24	468	859	1,283	8,426	1.47	24	657	1,240	2,127	37,543	2.19	18	646	1,224	2,102	24,275	2.17
Quarter 4	17,231	25	464	865	1,299	7,417	1.45	26	662	1,232	2,191	37,296	2.19	20	642	1,236	2,194	26,362	2.21
Household Type																			
12-17, 18-25, 26+	5,244	23	532	908	1,316	7,624	1.44	23	532	908	1,316	7,624	1.44	21	511	909	1,328	7,416	1.43
12-17, 18-25	63	82	433	863	1,249	2,481	1.41	82	433	864	1,249	2,482	1.41	73	379	875	1,318	3,097	1.46
12-17, 26+	15,469	20	397	799	1,189	8,426	1.51	20	399	803	1,194	8,426	1.51	17	390	806	1,199	8,337	1.50
18-25, 26+	11,895	24	545	939	1,320	8,909	1.41	24	618	1,076	1,508	10,029	1.40	28	605	1,089	1,531	7,160	1.39
12-17	16	97	252	657	1,262	1,664	1.50	97	254	666	1,275	1,681	1.50	80	209	674	1,188	1,811	1.56
18-25	5,519	25	456	862	1,270	6,992	1.44	25	504	980	1,450	8,520	1.44	19	491	975	1,444	4,868	1.43
26+	31,842	20	485	879	1,268	7,712	1.43	35	1,150	2,081	3,656	37,543	1.78	36	1,131	2,053	3,662	26,362	1.79
Race/Ethnicity of Householder																			
Hispanic or Latino White	10,069	24	579	970	1,391	8,602	1.31	33	707	1,235	1,880	31,467	1.90	32	698	1,222	1,885	25,446	1.88
Hispanic or Latino Black or African American	268	43	818	1,938	3,059	8,909	1.68	43	1,015	2,449	5,054	37,543	2.36	33	1,115	2,323	4,091	26,362	2.36
Hispanic or Latino Other	653	36	256	789	1,817	7,727	2.09	36	370	1,121	2,671	37,296	3.01	31	361	1,094	2,593	23,999	2.69
Non-Hispanic or Latino White	44,886	20	421	840	1,192	5,657	1.42	20	659	1,221	2,186	26,330	2.17	17	655	1,223	2,196	23,451	2.20
Non-Hispanic or Latino Black or African American	8,110	20	703	1,000	1,367	7,417	1.33	35	844	1,390	2,314	24,229	1.97	33	839	1,385	2,320	16,173	1.95
Non-Hispanic or Latino Other	6,062	24	263	766	1,332	8,426	1.66	24	423	1,078	1,961	35,908	2.26	20	413	1,074	2,012	24,275	2.27
% Hispanic or Latino in Segment																			
50-100%	5,251	91	711	1,159	1,518	8,602	1.29	91	878	1,435	2,154	22,096	1.80	79	892	1,462	2,225	23,266	1.77
10-<50%	17,621	24	634	1,058	1,522	8,909	1.39	24	834	1,467	2,473	37,543	2.03	19	820	1,463	2,448	26,362	2.02
<10%	47,176	20	348	807	1,125	8,426	1.46	20	573	1,130	1,963	35,908	2.24	17	568	1,129	1,990	25,446	2.26

Table G.1 2018 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%	4,944	24	673	976	1,277	6,632	1.35	24	809	1,318	2,212	28,324	2.10	19	794	1,319	2,227	26,362	2.10
10-<50%	13,958	23	671	989	1,399	8,909	1.36	23	846	1,411	2,358	26,330	1.97	21	844	1,399	2,347	25,254	1.94
<10%	51,146	20	385	828	1,220	8,602	1.48	20	599	1,177	2,055	37,543	2.21	17	592	1,175	2,070	25,446	2.24
% Owner-Occupied DUs[1] in Segment																			
50-100%	51,498	20	453	849	1,215	8,909	1.44	24	657	1,202	2,084	37,296	2.14	19	645	1,201	2,086	24,178	2.17
10-<50%	14,841	23	538	951	1,365	8,426	1.45	23	719	1,325	2,259	37,543	2.18	20	707	1,327	2,275	26,362	2.16
<10%	3,709	20	533	970	1,443	6,672	1.46	20	696	1,317	2,251	24,229	2.13	17	707	1,343	2,280	16,727	2.03
Combined Median Rent/Housing Value																			
1st Quintile	11,376	24	344	763	1,089	7,712	1.44	24	559	1,085	1,877	22,096	2.18	20	550	1,082	1,896	23,266	2.20
2nd Quintile	15,287	36	452	840	1,188	8,602	1.39	36	644	1,168	1,992	28,324	2.19	37	641	1,167	2,028	26,362	2.23
3rd Quintile	16,646	25	448	854	1,260	7,794	1.45	33	634	1,201	2,037	24,229	2.17	32	627	1,195	2,047	18,935	2.18
4th Quintile	15,067	24	532	933	1,346	8,909	1.44	24	712	1,323	2,268	37,543	2.12	19	703	1,323	2,260	23,999	2.10
5th Quintile	11,672	20	628	996	1,452	8,263	1.44	20	808	1,413	2,481	35,908	2.07	17	797	1,416	2,510	25,446	2.06
Population Density																			
Large MSA[1]	30,559	20	791	1,069	1,470	8,909	1.31	20	986	1,538	2,521	37,543	1.92	17	989	1,554	2,552	26,362	1.91
Medium to Small MSA[1]	34,000	24	295	719	1,079	7,712	1.50	24	458	988	1,778	22,096	2.33	28	451	983	1,772	25,254	2.35
Non-MSA,[1] Urban	1,706	34	212	655	1,029	3,285	1.59	58	367	931	1,696	14,720	2.45	53	371	937	1,701	19,150	2.53
Non-MSA,[1] Rural	3,783	20	187	517	935	2,965	1.61	26	315	831	1,574	12,760	2.52	20	310	824	1,582	13,165	2.58
Group Quarters																			
Group	555	45	302	794	1,304	3,216	1.50	47	401	1,037	1,567	17,345	2.55	49	367	980	1,597	15,910	2.46
Non-Group	69,493	20	478	873	1,264	8,909	1.45	20	672	1,235	2,134	37,543	2.15	17	663	1,234	2,142	26,362	2.16
Household Size																			
One	9,371	30	446	844	1,221	7,712	1.42	81	1,147	2,469	5,085	37,543	1.86	78	1,154	2,491	5,138	26,362	1.86
Two	29,245	20	477	868	1,249	8,602	1.44	31	796	1,483	2,585	24,229	1.76	31	783	1,483	2,580	24,178	1.78
Three	17,228	20	479	879	1,259	7,727	1.44	20	530	978	1,497	18,878	1.61	17	512	983	1,518	9,780	1.60
Four or More	14,204	23	497	900	1,336	8,909	1.48	23	507	925	1,396	9,223	1.51	18	488	919	1,396	8,337	1.50

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

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

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

Table G.2 2018 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	50,373	17	646	1,203	2,065	26,362	2.19	21	843	1,629	2,938	39,662	2.30	21	843	1,630	2,937	34,496	2.30
Census Region																			
Northeast	9,753	36	460	970	1,690	26,362	2.44	51	625	1,342	2,616	34,015	2.54	51	624	1,342	2,615	34,496	2.54
South	16,931	17	887	1,477	2,477	25,254	1.97	21	1,130	1,921	3,376	32,749	2.08	21	1,131	1,924	3,377	32,780	2.08
Midwest	11,810	20	704	1,098	1,791	17,973	2.13	22	918	1,508	2,580	27,120	2.20	21	918	1,508	2,583	27,169	2.20
West	11,879	28	437	1,140	2,005	24,178	2.39	28	580	1,518	2,905	39,662	2.55	28	580	1,519	2,907	33,675	2.55
Quarter																			
Quarter 1	11,898	28	721	1,294	2,211	25,254	2.12	28	925	1,739	3,130	32,749	2.22	28	926	1,739	3,130	32,780	2.22
Quarter 2	13,457	17	614	1,134	1,899	25,446	2.21	22	799	1,550	2,723	29,721	2.30	22	799	1,552	2,723	30,716	2.30
Quarter 3	12,763	19	630	1,196	2,047	24,275	2.19	21	826	1,598	2,876	33,196	2.33	21	827	1,597	2,876	32,869	2.32
Quarter 4	12,255	20	637	1,209	2,119	26,362	2.22	22	837	1,653	3,040	39,662	2.34	21	837	1,652	3,037	34,496	2.34
Household Type																			
12-17, 18-25, 26+	4,097	21	504	905	1,329	4,925	1.42	24	633	1,150	1,705	8,049	1.44	24	632	1,151	1,710	5,681	1.43
12-17, 18-25	50	73	441	861	1,331	3,097	1.46	96	527	1,122	1,696	3,753	1.48	95	525	1,122	1,706	3,766	1.48
12-17, 26+	11,843	17	401	808	1,206	5,070	1.49	21	506	1,039	1,579	10,879	1.51	21	506	1,039	1,581	8,202	1.50
18-25, 26+	8,541	28	599	1,084	1,529	7,160	1.40	28	795	1,481	2,159	11,133	1.43	28	794	1,483	2,161	8,447	1.43
12-17	8	82	373	681	1,188	1,811	1.46	82	516	1,027	2,473	3,665	1.71	82	511	1,019	2,475	3,686	1.71
18-25	4,152	39	485	971	1,449	4,868	1.44	43	629	1,293	1,959	7,587	1.46	42	631	1,299	1,959	6,872	1.46
26+	21,682	36	1,118	2,036	3,648	26,362	1.80	59	1,580	2,973	5,390	39,662	1.82	59	1,578	2,971	5,391	34,496	1.82
Race/Ethnicity of Householder																			
Hispanic or Latino White	7,378	32	678	1,191	1,833	25,446	1.90	36	882	1,617	2,520	29,721	2.04	36	883	1,616	2,524	30,716	2.03
Hispanic or Latino Black or African American	199	107	1,112	2,558	4,091	26,362	2.46	120	1,539	3,059	5,092	31,916	2.45	122	1,566	3,035	5,097	31,966	2.46
Hispanic or Latino Other	508	31	343	986	2,377	17,950	2.62	43	401	1,286	3,273	30,888	2.97	43	403	1,297	3,283	31,166	2.98
Non-Hispanic or Latino White	31,759	17	642	1,195	2,127	23,451	2.22	21	850	1,649	3,077	32,955	2.30	21	849	1,650	3,075	32,869	2.30
Non-Hispanic or Latino Black or African American	6,271	36	833	1,353	2,258	15,729	1.95	45	1,015	1,691	2,965	26,998	2.10	46	1,018	1,694	2,959	26,971	2.10
Non-Hispanic or Latino Other	4,258	20	389	1,008	1,873	24,275	2.32	22	503	1,337	2,722	39,662	2.66	21	506	1,349	2,736	34,496	2.64
% Hispanic or Latino in Segment																			
50-100%	3,872	79	861	1,413	2,145	23,266	1.79	80	1,150	1,881	3,067	30,281	1.89	78	1,152	1,884	3,068	30,540	1.89
10-<50%	12,685	19	784	1,410	2,351	26,362	2.06	21	1,019	1,924	3,386	39,662	2.18	21	1,021	1,925	3,384	34,496	2.17
<10%	33,816	17	564	1,110	1,922	25,446	2.29	22	741	1,497	2,724	33,196	2.39	21	742	1,498	2,724	32,780	2.39

Table G.2 2018 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%	3,849	19	797	1,302	2,161	26,362	2.10	21	956	1,606	2,895	27,109	2.22	21	953	1,606	2,884	27,574	2.22
10-<50%	10,269	21	823	1,371	2,266	25,254	1.96	24	1,053	1,824	3,163	39,662	2.07	24	1,056	1,826	3,162	33,675	2.06
<10%	36,255	17	575	1,144	1,997	25,446	2.28	22	765	1,574	2,878	34,015	2.38	21	764	1,576	2,877	34,496	2.38
% Owner-Occupied DUs[1] in Segment																			
50-100%	36,731	19	636	1,176	2,020	24,178	2.19	21	834	1,601	2,896	32,955	2.29	21	834	1,601	2,896	32,869	2.29
10-<50%	10,871	20	675	1,276	2,182	26,362	2.22	22	865	1,703	3,069	39,662	2.34	21	867	1,704	3,062	34,496	2.33
<10%	2,771	17	675	1,304	2,187	16,727	2.08	22	857	1,673	2,994	30,281	2.27	22	858	1,675	2,966	30,540	2.27
Combined Median Rent/Housing Value																			
1st Quintile	8,561	20	560	1,073	1,853	23,266	2.20	22	713	1,389	2,501	26,588	2.27	21	714	1,390	2,496	26,804	2.27
2nd Quintile	11,254	37	625	1,143	1,940	26,362	2.25	45	814	1,529	2,694	31,138	2.32	44	815	1,530	2,693	31,124	2.32
3rd Quintile	11,930	32	618	1,174	1,989	17,950	2.20	36	802	1,603	2,813	30,888	2.31	36	801	1,604	2,804	31,166	2.31
4th Quintile	10,608	19	678	1,293	2,191	23,451	2.13	21	906	1,805	3,183	34,015	2.24	21	905	1,807	3,182	34,496	2.24
5th Quintile	8,020	17	768	1,363	2,422	25,446	2.12	22	1,038	1,960	3,603	39,662	2.23	22	1,038	1,962	3,598	33,675	2.22
Population Density																			
Large MSA[1]	21,519	17	964	1,513	2,461	26,362	1.94	21	1,291	2,107	3,628	39,662	2.04	21	1,295	2,108	3,625	34,496	2.04
Medium to Small MSA[1]	24,877	28	451	973	1,735	25,254	2.37	28	606	1,296	2,382	28,434	2.46	28	607	1,297	2,381	28,664	2.46
Non-MSA,[1] Urban	1,280	53	353	916	1,664	13,530	2.49	57	489	1,187	2,180	16,970	2.57	55	489	1,185	2,185	16,961	2.57
Non-MSA,[1] Rural	2,697	20	315	821	1,553	13,165	2.57	22	443	1,088	2,145	22,588	2.76	21	445	1,090	2,143	23,116	2.76
Group Quarters																			
Group	516	49	367	990	1,594	15,910	2.51	50	405	1,073	1,759	16,363	2.48	50	413	1,073	1,759	16,348	2.49
Non-Group	49,857	17	649	1,205	2,069	26,362	2.19	21	848	1,636	2,956	39,662	2.29	21	849	1,637	2,952	34,496	2.29
Household Size																			
One	6,759	78	1,122	2,440	5,021	26,362	1.88	82	1,529	3,375	7,159	39,662	1.93	82	1,531	3,375	7,166	34,496	1.93
Two	20,364	31	765	1,447	2,510	24,178	1.79	39	1,024	2,016	3,668	21,743	1.88	38	1,026	2,017	3,658	19,386	1.88
Three	12,493	17	509	971	1,494	9,780	1.58	22	670	1,296	2,044	21,783	1.69	22	669	1,298	2,048	22,542	1.69
Four or More	10,757	19	481	908	1,377	7,572	1.49	21	617	1,182	1,843	11,068	1.59	21	616	1,183	1,843	11,836	1.59

[1] DU = dwelling unit, MSA = metropolitan statistical area, NR = nonresponse adjustment, PS = poststratification adjustment, QDU = questionnaire dwelling unit, Res = respondent, SDU = screener dwelling unit, Sel = selected.

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

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

Appendix H: GEM Modeling Summary for the Pair Weights

This page intentionally left blank

Appendix H: GEM Modeling Summary for the Pair Weights

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

For 2018, 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 2018 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 2018 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	4.01	18.43	8.07	8.5146	213	(0.20, 1.10)	(0.20, 1.10)
	1.07	2.83	0.36	4.5984	205	(0.20, 2.49)	(0.20, 2.49)
						(0.90, 1.36)	(0.90, 1.36)
res.pr.nr	1.00	3.49	0.46	4.6210	213	(1.05, 2.00)	(1.05, 2.00)
	1.72	6.92	1.09	5.8993	213	(1.00, 4.74)	(1.00, 4.71)
						N/A	N/A
res.pr.ps	1.74	6.76	1.33	5.8993	223	(0.22, 1.20)	(0.22, 1.20)
	1.17	3.94	0.22	5.8557	215	(0.20, 1.57)	(0.20, 1.57)
						N/A	N/A
res.pr.ev	1.17	3.94	0.22	5.8557	223	(0.98, 1.17)	(0.99, 1.17)
	0.22	0.73	0.02	5.8192	215	(0.95, 1.17)	(0.96, 1.14)
						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 $1 + \lceil (n-1)/n \rceil * CV^2$, 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 2018 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	20	1.02	23	0.03	8	0.60	12	0.13	6	0.60	6
1%	65	1.13	178	0.20	109	1.01	134	0.28	89	0.92	88
5%	130	1.37	419	0.27	303	1.05	411	0.48	326	0.98	323
10%	203	1.52	751	0.46	531	1.10	721	0.63	625	0.98	622
25%	544	2.48	1,680	0.77	1,452	1.23	1,973	0.85	1,823	0.99	1,818
Median	921	3.81	3,543	1.03	3,618	1.45	5,027	1.03	4,895	1.00	4,886
75%	1,302	8.59	8,054	1.30	8,198	1.83	12,390	1.17	12,395	1.01	12,440
90%	1,789	17.35	16,854	1.71	18,642	2.40	30,576	1.29	31,224	1.02	31,426
95%	2,155	29.23	28,549	2.01	32,663	2.86	54,715	1.37	55,575	1.03	55,465
99%	3,043	58.64	67,546	2.32	77,684	3.93	154,831	1.48	159,203	1.05	159,290
Maximum	8,909	1,165.90	1,232,953	2.78	262,200	4.71	558,734	1.57	477,023	1.15	431,045
n	15,024	-	15,024	-	15,024	-	9,072	-	9,072	-	9,072
Mean	990	8.54	8,150	1.06	8,422	1.63	13,947	0.99	13,974	1.00	13,947
Max/Mean	9	-	151	-	31	-	40	-	32	-	31

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 2018 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 2018 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 2018 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 2018 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 2018 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	4.19	20.27	9.27	16.2517	212	(0.43, 1.80)	(0.44, 1.80)
	2.04	4.97	0.52	5.0067	200	(0.30, 1.98)	(0.31, 1.97)
						(0.90, 2.08)	(1.22, 2.08)
res.pr.nr	2.06	5.70	0.63	5.2308	212	(1.00, 2.30)	(1.00, 2.30)
	2.65	8.19	1.57	6.4613	212	(1.00, 5.00)	(1.00, 5.00)
						(1.80, 5.00)	(1.81, 1.81)
res.pr.ps	2.72	9.85	2.31	6.4613	222	(0.48, 1.30)	(0.48, 1.30)
	1.52	4.74	0.28	6.3561	210	(0.38, 1.54)	(0.39, 1.54)
						N/A	N/A
res.pr.ev	1.52	4.74	0.28	6.3561	222	(0.98, 1.20)	(0.99, 1.20)
	0.40	1.36	0.03	6.3227	210	(0.94, 1.20)	(0.95, 1.13)
						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 $1 + \lceil (n-1)/n \rceil * CV^2$, 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 2018 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	24	1.02	37	0.03	30	0.64	34	0.22	27	0.73	27
1%	87	1.02	189	0.34	139	1.00	176	0.44	162	0.90	158
5%	131	1.36	384	0.51	314	1.00	425	0.57	401	0.97	401
10%	171	1.52	653	0.64	567	1.08	758	0.69	708	0.98	705
25%	452	2.48	1,444	0.82	1,344	1.22	1,828	0.88	1,785	0.99	1,787
Median	833	3.91	3,109	1.02	3,143	1.46	4,407	1.05	4,467	1.00	4,474
75%	1,262	8.17	7,246	1.23	7,492	1.86	11,311	1.18	11,159	1.01	11,192
90%	1,664	16.12	15,013	1.44	16,558	2.50	26,646	1.29	26,775	1.02	26,853
95%	1,932	26.93	26,159	1.56	28,882	3.13	49,988	1.35	49,218	1.03	49,197
99%	2,863	54.98	62,265	1.82	71,921	4.91	140,356	1.47	142,996	1.06	140,896
Maximum	8,426	2,031.50	2,749,216	2.76	298,076	5.00	555,914	1.54	479,818	1.14	458,425
n	14,039	-	14,039	-	14,039	-	8,346	-	8,346	-	8,346
Mean	911	8.18	7,507	1.03	7,567	1.67	12,729	1.02	12,729	1.00	12,729
Max/Mean	9	-	366	-	39	-	44	-	38	-	36

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, all 212 proposed factors were retained in the final model.

Respondent Pair-Level Poststratification

In the respondent pair-level poststratification step, 210 of 222 proposed factors were retained in the final model. All main and two-factor effects were retained to proposed levels. None of the 12 three-factor effects were kept in the model 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.2.1 Covariates for 2018 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 2018 NSDUH Pair Weights (res.pr.nr) Model Group 2: Midwest and West

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

Exhibit H.2.3 Covariates for 2018 NSDUH Pair Weights (res.pr.ps) Model Group 2: Midwest and West

Variables	Level	Proposed	Final	Comments
One-Factor Effects		**85**	**85**	
Intercept	1	1	1	All levels present.
State	26	24	24	All levels present.
Quarter	4	3	3	All levels present.
Population Density	4	3	3	All levels present.
Group Quarter	3	2	2	All levels present.
Household Size	3	2	2	All levels present.
Pair Age	15	14	14	All levels present.
Pair Gender	3	2	2	All levels present.
Pair Race/Ethnicity	10	9	9	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Rent/Housing	5	4	4	All levels present.
Segment % Black or African American	3	2	2	All levels present.
Segment % Hispanic or Latino	3	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
Pair Relationship	10	10	10	All levels present.
Two-Factor Effects		**125**	**125**	
Pair Race/Ethnicity (5 Levels) × Pair Age (6 Levels)	5 × 6	20	20	All levels present.
Pair Race/Ethnicity (5 Levels) × Pair Gender	5 × 3	8	8	All levels present.
Pair Gender × Pair Age (6 Levels)	3 × 6	10	10	All levels present.
State/Region × Pair Race/Ethnicity (5 Levels)	6 × 5	20	20	All levels present.
State/Region × Pair Age (6 Levels)	6 × 6	25	25	All levels present.
State/Region × Pair Gender	6 × 3	10	10	All levels present.
Rent/Housing × % Black or African American	5 × 3	8	8	All levels present.
Rent/Housing × % Hispanic or Latino	5 × 3	8	8	All levels present.
Rent/Housing × % Owner-Occupied	5 × 3	8	8	All levels present.
% Owner-Occupied × % Black or African American	3 × 3	4	4	All levels present.
% Owner-Occupied × % Hispanic or Latino	3 × 3	4	4	All levels present.
Three-Factor Effects		**12**	**0**	
Pair Race/Ethnicity (4 Levels) × Pair Gender × Pair Age (3 Levels)	4 × 3 × 3	12	0	Drop all; conv.
Total		**222**	**210**	

Exhibit H.2.4 Covariates for 2018 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 2018 NSDUH Person Pair-Level Response Rates

Domain	Selected Pairs	Respondent Pairs	% Interview Response Rate[1]
Total	29,063	17,418	54.40
Pair Age Group			
12-17, 12-17	3,296	2,356	71.92
12-17, 18-25	2,680	1,764	67.32
12-17, 26-34	1,243	825	66.62
12-17, 35-49	5,939	3,848	66.42
12-17, 50+	1,117	671	59.99
18-25, 18-25	4,381	2,582	57.70
18-25, 26-34	1,557	836	54.37
18-25, 35-49	2,187	1,202	55.51
18-25, 50+	1,282	647	48.59
26-34, 26-34	1,495	811	54.02
26-34, 35-49	788	392	50.07
26-34, 50+	432	203	45.82
35-49, 35-49	1,338	664	51.40
35-49, 50+	490	221	45.43
50+, 50+	838	396	47.60
Pair Race/Ethnicity			
Hispanic or Latino	5,291	3,181	55.44
Black or African American	2,883	1,995	65.52
White	15,601	9,159	53.13
Other	2,269	1,236	38.65
White & Black or African American	293	174	64.02
White & Hispanic or Latino	1,164	685	58.32
White & Other	1,007	638	62.68
Black or African American & Hispanic or Latino	169	108	59.69
Black or African American & Other	156	103	60.17
Hispanic or Latino & Other	230	139	55.12
Pair Gender			
Male, Male	6,165	3,559	51.46
Female, Female	6,256	4,056	59.86
Male, Female	16,642	9,803	53.66
Household Size			
Two	7,321	4,169	52.25
Three	9,122	5,442	52.61
Four or More	12,620	7,807	56.37

Table I.1 2018 NSDUH Person Pair-Level Response Rates (continued)

Domain	Selected Pairs	Respondent Pairs	% Interview Response Rate[1]
Census Region			
Northeast	5,811	3,186	47.19
South	9,213	5,886	58.16
Midwest	6,888	4,122	55.61
West	7,151	4,224	53.13
Quarter			
Quarter 1	6,906	4,233	55.06
Quarter 2	7,738	4,563	53.59
Quarter 3	7,226	4,272	53.59
Quarter 4	7,193	4,350	55.34
% Hispanic or Latino in Segment			
50-100%	2,680	1,633	53.67
10-<50%	7,633	4,456	52.98
<10%	18,750	11,329	55.30
% Black or African American in Segment			
50-100%	1,932	1,320	63.49
10-<50%	5,798	3,542	54.59
<10%	21,333	12,556	53.40
% Owner-Occupied DUs in Segment			
50-100%	21,881	13,055	54.40
10-<50%	5,780	3,533	54.44
<10%	1,402	830	53.57
Combined Median Rent/Housing Value			
1st Quintile	4,655	3,023	62.88
2nd Quintile	6,358	3,959	56.03
3rd Quintile	6,930	4,144	56.07
4th Quintile	6,334	3,615	50.91
5th Quintile	4,786	2,677	49.20
Population Density			
Large MSA	12,805	7,401	51.94
Medium to Small MSA	14,091	8,671	57.84
Non-MSA, Urban	658	424	57.77
Non-MSA, Rural	1,509	922	55.35
Group Quarters			
Group	255	195	77.34
Non-Group	28,808	17,223	54.33

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 2018 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	29,063	1.84	4.74	1.21	4.11	19.58	9.52	1.64	6.07	1.01
Pair Age Group										
12-17, 12-17	3,296	1.03	3.35	0.75	3.61	14.70	4.36	1.00	4.13	0.46
12-17, 18-25	2,680	1.04	3.57	1.21	6.83	25.96	11.08	1.12	3.32	0.39
12-17, 26-34	1,243	2.01	4.87	1.73	1.93	6.35	1.57	0.80	2.86	0.53
12-17, 35-49	5,939	1.41	3.65	0.88	1.67	8.25	2.40	0.69	1.85	0.30
12-17, 50+	1,117	1.79	4.64	1.07	1.34	9.57	4.10	0.27	1.54	0.16
18-25, 18-25	4,381	1.78	4.12	1.01	6.73	26.98	12.86	2.44	6.21	0.56
18-25, 26-34	1,557	3.08	6.67	1.76	4.17	10.76	2.70	2.83	4.22	0.45
18-25, 35-49	2,187	2.38	5.59	1.48	6.45	22.22	7.74	2.51	5.48	0.69
18-25, 50+	1,282	2.18	5.58	1.53	3.43	15.14	5.08	1.25	3.36	0.24
26-34, 26-34	1,495	3.41	7.54	1.50	2.41	7.94	2.62	2.14	5.42	0.91
26-34, 35-49	788	2.92	6.33	2.09	4.06	16.52	7.89	4.06	6.06	1.09
26-34, 50+	432	1.62	4.86	1.41	2.31	14.85	6.97	0.69	4.48	1.40
35-49, 35-49	1,338	2.17	6.55	1.30	4.19	22.29	10.58	1.72	2.90	0.42
35-49, 50+	490	2.24	5.65	1.65	4.29	28.23	15.60	2.45	13.01	2.73
50+, 50+	838	2.15	6.13	1.43	6.56	31.70	20.01	4.18	12.45	2.17
Pair Race/Ethnicity										
Hispanic or Latino	5,291	2.44	7.49	2.61	3.65	18.71	8.09	1.44	5.91	1.12
Black or African American	2,883	2.46	4.66	0.73	5.41	29.33	17.92	2.01	5.76	0.96
White	15,601	0.72	1.95	0.34	3.51	15.18	6.86	1.24	4.81	0.55
Other	2,269	4.41	10.51	2.37	6.08	34.46	18.56	2.86	14.13	3.74
White & Black or African American	293	3.75	3.18	0.54	6.48	9.74	4.17	4.44	8.60	0.81
White & Hispanic or Latino	1,164	2.41	5.01	1.36	4.12	20.83	9.93	3.78	6.49	1.00
White & Other	1,007	4.27	8.87	2.32	4.97	15.94	6.36	0.99	4.68	0.19
Black or African American & Hispanic or Latino	169	15.98	31.75	10.53	13.61	43.12	12.21	4.73	14.65	4.45
Black or African American & Other	156	3.21	4.89	0.63	3.85	13.26	5.51	0.00	0.00	0.00
Hispanic or Latino & Other	230	4.35	11.76	3.14	6.52	30.37	17.96	3.91	8.60	2.33
Pair Gender										
Male, Male	6,165	1.98	4.72	1.17	4.98	17.21	7.27	2.58	5.47	0.88
Female, Female	6,256	1.98	4.76	1.09	4.91	22.31	12.10	1.44	4.00	0.48
Male, Female	16,642	1.74	4.74	1.28	3.49	19.43	9.39	1.36	6.82	1.20
Household Size										
Two	7,321	1.80	4.65	1.12	0.94	3.43	0.84	0.38	1.17	0.16
Three	9,122	1.69	4.04	1.03	1.91	23.98	14.64	1.36	6.28	0.72
Four or More	12,620	1.98	5.26	1.39	7.54	25.67	11.34	2.57	8.41	1.59

Table J.1 2018 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	5,811	2.62	7.46	2.20	4.20	22.92	11.99	1.36	6.16	0.82
South	9,213	1.30	3.17	0.62	3.92	16.65	7.17	1.06	4.54	0.48
Midwest	6,888	1.67	4.00	1.06	4.31	21.68	12.00	2.09	6.19	1.08
West	7,151	2.08	6.06	1.64	4.10	19.60	9.09	2.17	8.04	1.84
Quarter										
Quarter 1	6,906	2.13	5.41	1.63	4.59	22.49	11.51	2.35	7.29	1.17
Quarter 2	7,738	1.46	3.59	0.79	3.57	18.45	10.07	1.19	3.95	0.55
Quarter 3	7,226	2.20	5.98	1.46	3.99	18.10	7.43	1.56	7.05	1.16
Quarter 4	7,193	1.63	3.95	0.96	4.37	19.21	9.03	1.52	5.94	1.17
% Hispanic or Latino in Segment										
50-100%	2,680	1.57	3.87	1.05	3.17	20.11	9.69	1.23	6.82	1.74
10-<50%	7,633	2.33	5.98	1.66	4.56	19.87	9.25	1.98	6.72	1.00
<10%	18,750	1.69	4.22	1.00	4.06	19.31	9.64	1.56	5.57	0.88
% Black or African American in Segment										
50-100%	1,932	3.11	7.71	1.89	6.68	36.88	22.15	2.90	7.24	1.63
10-<50%	5,798	2.47	6.16	1.59	4.09	22.20	11.13	1.72	5.82	0.69
<10%	21,333	1.56	3.96	1.02	3.89	16.66	7.49	1.50	6.02	1.05
% Owner-Occupied DUs[1] in Segment										
50-100%	21,881	1.36	3.34	0.83	3.83	17.59	8.02	1.52	6.01	1.02
10-<50%	5,780	3.25	8.51	2.31	5.05	24.98	13.84	2.40	6.52	0.97
<10%	1,402	3.64	8.56	2.07	4.64	28.41	15.26	0.36	3.86	1.24
Combined Median Rent/Housing Value										
1st Quintile	4,655	1.50	3.79	0.87	4.10	20.22	10.24	1.53	3.73	0.86
2nd Quintile	6,358	1.71	3.72	0.94	3.99	17.52	7.26	1.78	5.41	0.70
3rd Quintile	6,930	1.70	4.56	1.17	3.98	19.95	8.98	1.47	6.80	1.35
4th Quintile	6,334	1.72	4.51	1.30	4.22	18.63	8.64	1.47	7.95	1.32
5th Quintile	4,786	2.72	7.00	1.70	4.33	21.86	12.97	2.03	5.22	0.68
Population Density										
Large MSA[1]	12,805	2.50	6.18	1.66	4.57	23.02	12.03	2.07	7.85	1.31
Medium to Small MSA[1]	14,091	1.35	3.11	0.72	3.70	15.31	6.27	1.26	4.08	0.69
Non-MSA,[1] Urban	658	1.22	2.70	0.50	2.58	7.19	2.77	1.67	1.99	0.15
Non-MSA,[1] Rural	1,509	1.19	1.23	0.15	4.77	12.59	5.05	1.46	0.71	0.10
Group Quarters										
Group	255	3.92	6.81	1.04	8.24	26.67	8.28	3.14	10.39	1.77
Non-Group	28,808	1.83	4.72	1.22	4.08	19.56	9.53	1.62	6.06	1.01

[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 2018 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	17,418	1.58	6.50	1.12	2.20	8.53	1.85
Pair Age Group							
12-17, 12-17	2,356	0.81	3.56	0.57	0.42	1.42	0.19
12-17, 18-25	1,764	1.42	5.48	0.63	1.70	8.72	1.73
12-17, 26-34	825	1.09	4.51	0.61	1.33	6.43	2.11
12-17, 35-49	3,848	0.83	2.64	0.39	0.83	5.06	1.18
12-17, 50+	671	0.60	2.21	0.34	0.30	1.67	0.03
18-25, 18-25	2,582	2.29	6.42	0.74	3.21	12.97	2.58
18-25, 26-34	836	3.47	4.88	0.65	6.34	18.92	6.49
18-25, 35-49	1,202	2.91	5.50	0.68	6.16	9.89	1.53
18-25, 50+	647	2.16	7.12	1.03	2.47	5.07	0.62
26-34, 26-34	811	1.36	3.17	0.56	2.34	6.03	1.34
26-34, 35-49	392	1.53	2.76	0.62	2.81	6.48	1.18
26-34, 50+	203	1.48	9.92	2.58	1.97	9.17	0.87
35-49, 35-49	664	1.20	4.00	0.86	3.16	10.56	1.51
35-49, 50+	221	3.17	15.25	2.86	2.26	9.42	1.77
50+, 50+	396	3.79	12.19	2.22	3.03	12.04	3.57
Pair Race/Ethnicity							
Hispanic or Latino	3,181	1.48	8.04	1.63	1.98	7.97	2.09
Black or African American	1,995	2.01	6.25	1.05	1.20	3.78	0.48
White	9,159	1.16	4.97	0.62	1.79	6.94	1.21
Other	1,236	2.91	15.11	3.26	7.61	32.85	8.52
White & Black or African American	174	4.60	9.93	2.75	1.15	0.33	0.06
White & Hispanic or Latino	685	3.21	8.35	1.57	2.19	3.59	0.90
White & Other	638	0.94	1.58	0.15	0.63	0.49	0.03
Black or African American & Hispanic or Latino	108	3.70	12.07	5.04	7.41	6.98	1.24
Black or African American & Other	103	0.00	0.00	0.00	0.97	2.80	0.24
Hispanic or Latino & Other	139	5.04	12.38	2.47	5.76	8.11	1.95
Pair Gender							
Male, Male	3,559	2.11	4.42	0.77	3.29	10.59	1.91
Female, Female	4,056	1.70	7.54	0.85	1.60	5.32	0.81
Male, Female	9,803	1.35	6.73	1.30	2.05	8.87	2.12
Household Size							
Two	4,169	0.34	0.88	0.23	0.74	1.60	0.35
Three	5,442	1.21	6.39	0.97	2.21	10.44	2.17
Four or More	7,807	2.51	9.16	1.61	2.97	11.03	2.43

Table J.2 2018 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	3,186	1.44	7.96	1.22	2.13	11.61	2.16
South	5,886	0.87	3.94	0.55	1.55	6.01	1.05
Midwest	4,122	2.09	6.33	1.05	2.33	6.79	1.52
West	4,224	2.20	9.66	2.00	3.03	11.24	2.99
Quarter							
Quarter 1	4,233	2.24	6.92	1.37	3.05	8.02	1.67
Quarter 2	4,563	1.10	5.47	0.85	1.64	6.56	1.09
Quarter 3	4,272	1.73	7.27	0.97	1.85	9.93	2.55
Quarter 4	4,350	1.31	6.31	1.29	2.30	9.59	2.06
% Hispanic or Latino in Segment							
50-100%	1,633	1.47	8.86	1.75	2.76	10.64	2.92
10-<50%	4,456	2.24	7.50	1.11	2.92	9.46	1.97
<10%	11,329	1.34	5.53	1.01	1.84	7.61	1.57
% Black or African American in Segment							
50-100%	1,320	2.50	7.34	1.84	1.36	5.69	1.30
10-<50%	3,542	1.33	5.30	0.64	2.94	11.37	2.40
<10%	12,556	1.56	6.76	1.18	2.08	7.98	1.74
% Owner-Occupied DUs[1] in Segment							
50-100%	13,055	1.44	6.48	1.10	2.08	8.46	1.82
10-<50%	3,533	2.41	7.09	1.31	2.92	9.06	2.02
<10%	830	0.36	0.95	0.14	0.96	6.04	1.32
Combined Median Rent/Housing Value							
1st Quintile	3,023	1.69	7.08	1.72	1.19	4.84	1.32
2nd Quintile	3,959	1.69	5.42	0.76	2.40	5.61	0.96
3rd Quintile	4,144	1.40	9.10	1.67	1.59	8.45	2.18
4th Quintile	3,615	1.27	6.31	1.04	2.27	10.77	1.87
5th Quintile	2,677	2.02	4.01	0.39	3.88	11.41	2.70
Population Density							
Large MSA[1]	7,401	2.05	7.56	1.26	2.80	11.31	2.53
Medium to Small MSA[1]	8,671	1.21	5.80	1.07	1.74	5.27	1.03
Non-MSA,[1] Urban	424	1.18	1.29	0.20	1.18	2.22	0.50
Non-MSA,[1] Rural	922	1.52	1.00	0.12	2.17	1.94	0.30
Group Quarters							
Group	195	2.56	9.89	1.59	0.51	0.33	0.03
Non-Group	17,223	1.57	6.49	1.12	2.22	8.55	1.85

[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 2018 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	17,418	2.23	8.28	1.82	1.35	4.40	0.28	0.32	1.11	0.06
Pair Age Group										
12-17, 12-17	2,349	0.43	1.42	0.18	0.38	1.54	0.06	0.00	0.00	0.00
12-17, 18-25	1,764	1.70	8.32	1.62	0.85	3.25	0.12	0.06	0.22	0.00
12-17, 26-34	826	1.33	6.37	2.12	0.48	1.73	0.08	0.36	1.75	0.03
12-17, 35-49	3,836	0.96	6.05	1.80	0.81	3.30	0.28	0.18	0.64	0.01
12-17, 50+	678	0.29	1.73	0.11	0.44	1.44	0.19	0.00	0.00	0.00
18-25, 18-25	2,548	3.49	14.03	2.83	2.67	6.49	0.32	0.12	0.67	0.01
18-25, 26-34	851	6.35	22.61	9.16	4.35	15.65	2.53	2.47	10.96	1.32
18-25, 35-49	1,179	5.94	10.69	2.38	2.88	8.35	0.59	1.02	1.46	0.03
18-25, 50+	652	2.15	5.48	0.71	1.84	5.09	0.31	1.07	4.04	0.10
26-34, 26-34	827	2.42	8.09	2.06	0.24	2.21	0.08	0.00	0.00	0.00
26-34, 35-49	408	2.94	9.46	1.66	1.23	3.50	0.41	0.25	0.45	0.02
26-34, 50+	216	0.93	3.42	0.09	0.00	0.00	0.00	0.00	0.00	0.00
35-49, 35-49	661	3.63	10.81	3.40	1.06	4.61	0.23	0.00	0.00	0.00
35-49, 50+	221	1.36	4.92	0.61	1.81	4.61	0.09	0.00	0.00	0.00
50+, 50+	402	2.74	10.97	1.70	1.24	4.72	0.11	0.25	0.80	0.03
Pair Race/Ethnicity										
Hispanic or Latino	3,198	1.91	7.88	1.80	1.34	3.95	0.44	0.31	1.06	0.19
Black or African American	1,958	1.33	2.93	0.80	0.72	1.29	0.04	0.10	0.36	0.00
White	8,863	1.70	6.13	1.03	0.93	2.46	0.09	0.00	0.00	0.00
Other	1,208	7.78	33.15	8.32	5.30	22.24	1.40	3.31	10.10	0.30
White & Black or African American	168	0.60	0.14	0.09	1.79	0.55	0.03	0.00	0.00	0.00
White & Hispanic or Latino	728	2.75	7.22	1.94	3.30	7.69	0.52	0.27	0.74	0.01
White & Other	813	1.48	3.87	1.09	0.00	0.00	0.00	0.00	0.00	0.00
Black or African American & Hispanic or Latino	132	9.09	9.63	1.32	2.27	5.77	1.17	1.52	1.05	0.01
Black or African American & Other	191	2.09	4.37	1.04	0.00	0.00	0.00	0.00	0.00	0.00
Hispanic or Latino & Other	159	5.03	14.27	6.91	1.89	2.89	0.15	0.00	0.00	0.00
Pair Gender										
Male, Male	3,554	3.35	9.65	1.81	1.58	3.69	0.34	0.37	1.49	0.13
Female, Female	4,061	1.63	5.55	0.92	1.31	4.07	0.27	0.37	1.20	0.10
Male, Female	9,803	2.08	8.68	2.08	1.30	4.69	0.27	0.29	0.98	0.03
Household Size										
Two	4,169	0.84	2.04	0.50	0.50	1.28	0.12	0.29	0.47	0.02
Three	5,442	2.26	11.21	2.35	1.14	5.53	0.35	0.40	1.76	0.12
Four or More	7,807	2.96	9.92	2.22	1.96	5.38	0.33	0.28	1.11	0.05

Table J.3 2018 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	3,186	2.13	9.58	1.72	1.73	7.31	0.27	0.09	1.01	0.03
South	5,886	1.56	5.52	1.19	0.88	2.36	0.21	0.31	0.66	0.02
Midwest	4,122	2.50	6.64	1.91	1.24	2.59	0.16	0.32	0.44	0.01
West	4,224	2.98	12.46	2.71	1.85	6.60	0.49	0.52	2.31	0.17
Quarter										
Quarter 1	4,233	3.19	9.11	2.05	1.61	4.23	0.30	0.21	0.48	0.13
Quarter 2	4,563	1.67	6.06	1.20	0.96	4.92	0.27	0.33	1.62	0.04
Quarter 3	4,272	1.87	10.33	2.21	1.17	4.27	0.27	0.35	1.24	0.03
Quarter 4	4,350	2.25	7.56	1.81	1.70	4.18	0.29	0.39	1.13	0.04
% Hispanic or Latino in Segment										
50-100%	1,633	2.82	10.01	2.04	1.71	5.86	0.54	0.80	1.52	0.15
10-<50%	4,456	2.92	10.07	2.28	1.82	6.13	0.43	0.49	1.91	0.11
<10%	11,329	1.88	6.97	1.53	1.12	3.17	0.15	0.19	0.60	0.02
% Black or African American in Segment										
50-100%	1,320	1.59	6.36	1.58	1.06	4.20	0.24	0.23	2.74	0.08
10-<50%	3,542	2.96	10.33	2.45	1.78	5.60	0.39	0.45	1.20	0.09
<10%	12,556	2.09	7.87	1.66	1.27	4.06	0.26	0.29	0.92	0.05
% Owner-Occupied DUs[1] in Segment										
50-100%	13,055	2.06	7.90	1.64	1.15	4.05	0.25	0.28	1.16	0.03
10-<50%	3,533	3.14	10.03	2.61	2.24	6.01	0.46	0.51	0.98	0.20
<10%	830	1.08	6.39	1.35	0.84	2.55	0.06	0.12	0.37	0.02
Combined Median Rent/Housing Value										
1st Quintile	3,023	1.16	5.50	1.20	0.50	1.13	0.08	0.23	0.39	0.01
2nd Quintile	3,959	2.42	5.26	1.01	1.29	2.87	0.21	0.20	0.63	0.08
3rd Quintile	4,144	1.76	8.80	1.76	0.94	4.09	0.31	0.19	0.91	0.09
4th Quintile	3,615	2.10	8.97	2.04	1.69	5.57	0.33	0.28	1.70	0.05
5th Quintile	2,677	4.07	11.81	2.88	2.61	7.13	0.40	0.86	1.64	0.06
Population Density										
Large MSA[1]	7,401	2.85	10.61	2.27	1.77	5.98	0.38	0.50	1.56	0.10
Medium to Small MSA[1]	8,671	1.74	5.36	1.24	1.07	2.31	0.16	0.17	0.59	0.01
Non-MSA,[1] Urban	424	1.42	3.30	1.35	0.47	0.43	0.00	0.00	0.00	0.00
Non-MSA,[1] Rural	922	2.28	4.46	1.05	1.08	3.08	0.22	0.43	0.14	0.01
Group Quarters										
Group	195	0.51	0.33	0.01	0.51	0.18	0.02	0.00	0.00	0.00
Non-Group	17,223	2.25	8.30	1.83	1.36	4.41	0.28	0.33	1.12	0.06

Table J.3 2018 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)	2,668	1.01	5.99	1.98	0.90	2.94	0.28	0.22	0.54	0.01
Parent-Child (12-17)	4,872	0.94	5.17	1.39	0.76	2.69	0.24	0.18	0.50	0.01
Parent-Child (12-20)	5,677	1.51	5.93	1.39	0.92	3.35	0.31	0.33	1.30	0.03
Sibling (12-14)-Sibling (15-17)	1,361	0.44	1.65	0.21	0.37	1.64	0.03	0.00	0.00	0.00
Sibling (12-17)-Sibling (18-25)	1,588	1.39	6.43	1.18	0.82	3.16	0.12	0.06	0.24	0.00
Spouse-Spouse/Partner-Partner	3,271	1.90	9.13	1.94	0.92	4.46	0.20	0.24	0.69	0.03
Spouse-Spouse/Partner-Partner with Children (Younger Than 18)	1,601	1.87	14.75	3.13	1.31	8.24	0.42	0.25	1.40	0.05

[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 2018 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	17,418	232,761,691	232,761,691	232,761,691	0.00	-0.00
Pair Age Group						
12-17, 12-17	2,349	7,338,186	7,337,076	7,337,076	0.02	0.00
12-17, 18-25	1,764	8,285,429	8,316,147	8,316,147	-0.37	0.00
12-17, 26-34	826	4,898,147	4,853,005	4,853,005	0.93	0.00
12-17, 35-49	3,836	29,599,532	29,432,248	29,432,248	0.57	0.00
12-17, 50+	678	12,963,301	13,143,221	13,143,221	-1.37	0.00
18-25, 18-25	2,548	12,458,432	12,518,448	12,518,448	-0.48	-0.00
18-25, 26-34	851	7,599,132	7,389,307	7,389,307	2.84	-0.00
18-25, 35-49	1,179	16,448,990	16,390,927	16,390,927	0.35	0.00
18-25, 50+	652	19,079,260	19,367,243	19,367,243	-1.49	-0.00
26-34, 26-34	827	11,481,522	11,799,201	11,799,201	-2.69	0.00
26-34, 35-49	408	9,028,709	9,007,881	9,007,881	0.23	0.00
26-34, 50+	216	14,148,155	13,803,151	13,803,151	2.50	0.00
35-49, 35-49	661	18,387,997	18,533,203	18,533,203	-0.78	0.00
35-49, 50+	221	17,770,170	17,972,875	17,972,875	-1.13	0.00
50+, 50+	402	43,274,728	42,897,759	42,897,759	0.88	0.00
Pair Race/Ethnicity						
Hispanic or Latino	3,198	43,164,709	43,451,104	43,451,104	-0.66	0.00
Black or African American	1,958	24,754,749	24,967,604	24,967,604	-0.85	0.00
White	8,863	117,860,970	121,467,712	121,467,712	-2.97	0.00
Other	1,208	18,880,586	19,313,393	19,313,393	-2.24	-0.00
White & Black or African American	168	2,462,216	2,421,203	2,421,203	1.69	0.00
White & Hispanic or Latino	728	9,460,598	9,850,352	9,850,352	-3.96	0.00
White & Other	813	9,898,679	6,838,411	6,838,411	44.75	0.00
Black or African American & Hispanic or Latino	132	2,214,724	1,618,637	1,618,637	36.83	0.00
Black or African American & Other	191	2,029,113	1,153,602	1,153,602	75.89	0.00
Hispanic or Latino & Other	159	2,035,348	1,679,672	1,679,672	21.18	0.00
Pair Gender						
Male, Male	3,554	41,268,611	41,665,870	41,665,870	-0.95	-0.00
Female, Female	4,061	42,526,678	42,327,520	42,327,520	0.47	-0.00
Male, Female	9,803	148,966,401	148,768,301	148,768,301	0.13	-0.00
Pair Relationship Domain[3,4,5]						
Parent-Child (12-14)*	2,668	11,504,178	12,814,471	12,814,471	-10.23	0.00
Parent-Child (12-17)*	4,872	23,418,459	25,400,959	25,400,959	-7.80	0.00
Parent-Child (15-17)*	2,204	11,914,281	12,586,488	12,586,488	-5.34	0.00
Parent-Child (12-20)*	5,677	32,274,384	34,453,295	34,453,295	-6.32	0.00
Parent*-Child (12-14)	2,668	17,785,790	19,710,529	19,710,529	-9.77	0.00
Parent*-Child (12-17)	4,872	30,109,271	32,717,326	32,717,326	-7.97	0.00
Parent*-Child (15-17)	2,204	18,814,749	19,566,949	19,358,117	-2.81	1.08
Parent*-Child (12-20)	5,677	38,321,515	40,315,974	40,315,974	-4.95	0.00
Sibling (12-14)-Sibling (15-17)*	1,361	3,895,059	4,103,683	4,103,683	-5.08	0.00
Sibling (12-17)-Sibling (18-25)*	1,588	6,068,472	6,467,752	6,467,752	-6.17	0.00
Spouse-Spouse/Partner-Partner	3,271	77,459,294	76,167,650	76,167,650	1.70	-0.00
Spouse-Spouse/Partner-Partner with Children (Younger Than 18)	1,601	26,848,952	29,979,755	29,979,755	-10.44	0.00

Table K.1 2018 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	4,169	57,980,733	57,980,733	57,980,733	-0.00	0.00
Three	5,442	58,814,722	58,814,722	58,814,722	0.00	-0.00
Four or More	7,807	115,966,236	115,966,236	115,966,236	0.00	-0.00
Census Region						
Northeast	3,186	41,354,011	41,354,011	41,354,011	0.00	-0.00
South	5,886	85,172,707	85,172,707	85,172,707	0.00	-0.00
Midwest	4,122	45,213,911	45,213,911	45,213,911	0.00	-0.00
West	4,224	61,021,061	61,021,061	61,021,061	-0.00	-0.00
Quarter						
Quarter 1	4,233	58,809,570	58,809,570	58,809,570	0.00	-0.00
Quarter 2	4,563	57,423,511	57,423,511	57,423,511	0.00	-0.00
Quarter 3	4,272	58,619,894	58,619,894	58,619,894	0.00	-0.00
Quarter 4	4,350	57,908,715	57,908,715	57,908,715	-0.00	-0.00
% Hispanic or Latino in Segment						
50-100%	1,633	26,643,520	26,643,520	26,643,520	-0.00	0.00
10-<50%	4,456	72,158,889	72,158,889	72,158,889	-0.00	-0.00
<10%	11,329	133,959,282	133,959,282	133,959,282	0.00	-0.00
% Black or African American in Segment						
50-100%	1,320	17,131,039	17,131,039	17,131,039	-0.00	0.00
10-<50%	3,542	49,477,317	49,477,317	49,477,317	0.00	-0.00
<10%	12,556	166,153,335	166,153,335	166,153,335	0.00	-0.00
% Owner-Occupied DUs in Segment						
50-100%	13,055	184,257,419	184,257,419	184,257,419	0.00	-0.00
10-<50%	3,533	44,523,184	44,523,184	44,523,184	0.00	-0.00
<10%	830	3,981,088	3,981,088	3,981,088	-0.00	0.00
Combined Median Rent/Housing Value						
1st Quintile	3,023	31,137,182	31,137,182	31,137,182	0.00	0.00
2nd Quintile	3,959	46,832,918	46,832,918	46,832,918	-0.00	-0.00
3rd Quintile	4,144	53,791,126	53,791,126	53,791,126	0.00	-0.00
4th Quintile	3,615	55,135,454	55,135,454	55,135,454	0.00	0.00
5th Quintile	2,677	45,865,011	45,865,011	45,865,011	0.00	-0.00
Population Density						
Large MSA	7,401	132,381,325	132,381,325	132,381,325	-0.00	-0.00
Medium to Small MSA	8,671	88,077,300	88,077,300	88,077,300	0.00	-0.00
Non-MSA, Urban	424	3,787,672	3,787,672	3,787,672	-0.00	0.00
Non-MSA, Rural	922	8,515,394	8,515,394	8,515,394	0.00	-0.00
Group Quarters						
Group	195	619,095	619,095	619,095	-0.00	-0.00
Non-Group	17,223	232,142,596	232,142,596	232,142,596	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 2018 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	29,063	20	494	879	1,283	8,909	1.46	23	1,560	3,344	7,660	2,749,216	11.96	8	1,393	3,384	7,898	298,076	4.79
Pair Age Group																			
12-17, 12-17	3,296	20	388	771	1,205	6,391	1.54	23	790	1,551	2,687	36,535	2.30	8	490	1,271	2,967	21,213	2.44
12-17, 18-25	2,680	46	556	919	1,316	7,624	1.42	71	1,164	1,940	3,531	242,233	4.54	28	1,193	2,165	4,058	24,450	1.95
12-17, 26-34	1,243	38	431	838	1,239	7,083	1.54	100	1,440	2,630	4,071	36,581	2.23	64	1,318	2,600	4,595	64,341	2.54
12-17, 35-49	5,939	26	436	797	1,169	8,426	1.49	118	1,764	3,270	5,797	92,225	2.43	69	1,474	3,077	5,977	59,499	2.39
12-17, 50+	1,117	34	480	911	1,381	5,248	1.48	349	4,493	8,578	13,298	229,168	2.49	288	4,009	8,068	14,725	112,119	2.09
18-25, 18-25	4,381	24	528	926	1,349	6,012	1.41	47	1,036	1,857	3,194	263,545	5.13	49	810	1,833	3,869	22,517	2.09
18-25, 26-34	1,557	40	535	958	1,353	8,909	1.46	135	1,857	3,528	5,641	57,646	2.08	108	1,466	3,257	6,296	52,700	2.20
18-25, 35-49	2,187	35	555	909	1,256	7,794	1.40	128	2,727	4,721	8,269	141,898	2.71	79	2,289	4,670	9,205	62,841	2.20
18-25, 50+	1,282	24	644	1,029	1,432	6,128	1.38	784	6,562	10,467	15,803	248,235	2.30	546	6,075	11,384	18,905	113,909	1.89
26-34, 26-34	1,495	51	532	911	1,294	6,469	1.46	415	3,417	6,421	9,823	237,306	2.29	381	2,943	5,560	9,841	95,302	2.27
26-34, 35-49	788	25	502	872	1,283	6,340	1.47	257	4,311	7,613	12,010	420,438	4.98	230	3,437	7,496	13,361	138,634	2.69
26-34, 50+	432	40	547	962	1,373	4,656	1.39	1,183	12,035	22,242	31,887	449,040	2.86	807	12,883	23,925	40,371	245,817	1.94
35-49, 35-49	1,338	29	472	857	1,243	6,736	1.52	331	4,541	7,963	13,277	443,630	5.23	319	3,675	7,598	13,265	181,356	3.51
35-49, 50+	490	43	568	950	1,350	6,632	1.45	1,531	12,797	22,725	34,084	1,232,953	5.55	817	14,184	26,798	47,381	262,200	2.06
50+, 50+	838	29	571	955	1,440	6,672	1.43	1,616	22,782	39,570	53,721	2,749,216	5.75	986	23,583	44,667	65,887	298,076	1.61
Pair Race/Ethnicity																			
Hispanic or Latino	5,291	24	564	954	1,428	8,909	1.47	37	1,822	3,798	8,409	687,333	6.33	16	1,593	3,621	8,498	249,785	4.33
Black or African American	2,883	25	728	1,027	1,407	5,467	1.31	48	1,974	3,716	8,197	2,749,216	37.88	14	1,684	3,929	8,627	262,200	4.60
White	15,601	20	445	848	1,190	5,040	1.42	23	1,480	3,136	7,328	796,065	7.44	8	1,337	3,206	7,533	250,258	4.79
Other	2,269	26	275	754	1,365	8,426	1.69	90	1,205	3,082	7,533	1,232,953	17.98	21	1,021	3,176	8,014	298,076	6.32
White & Black or African American	293	35	585	950	1,361	3,456	1.40	69	1,729	3,604	7,192	76,080	3.46	67	1,818	4,367	8,241	113,909	4.11
White & Hispanic or Latino	1,164	34	475	871	1,313	6,711	1.50	123	1,475	3,393	7,630	338,664	6.45	38	1,568	3,758	9,277	223,863	4.32
White & Other	1,007	24	347	742	1,215	4,115	1.58	64	1,478	3,222	6,866	302,385	5.39	15	1,163	2,983	6,711	167,267	4.81
Black or African American & Hispanic or Latino	169	82	707	1,158	1,845	6,672	1.57	117	2,113	4,693	11,259	200,628	4.78	69	1,951	4,590	11,481	112,052	3.18
Black or African American & Other	156	44	574	922	1,296	2,901	1.36	205	1,835	4,343	9,640	106,662	3.95	147	1,235	2,916	7,370	82,987	4.35
Hispanic or Latino & Other	230	39	264	796	1,271	5,966	1.70	59	1,187	2,821	7,569	242,233	7.60	58	1,263	3,052	7,634	127,548	4.69

Table L.1 2018 NSDUH Selected Pair-Level Weight Summary Statistics (continued)

Domain	*n*	SDU-Level Weights[1] (SDUWT: WT1*...*WT11)						Before sel.pr.ps[1] (SDUWT*PRWT12)						After sel.pr.ps[1] (SDUWT*PRWT12*PRWT13)					
		Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
Pair Gender																			
Male, Male	6,165	24	489	875	1,295	7,083	1.47	31	1,477	3,107	6,933	758,206	6.51	8	1,407	3,274	7,644	236,007	3.91
Female, Female	6,256	24	490	889	1,298	8,426	1.45	47	1,568	3,222	6,992	2,749,216	30.07	13	1,402	3,210	7,196	181,356	4.12
Male, Female	16,642	20	496	875	1,271	8,909	1.46	23	1,583	3,522	8,189	1,232,953	8.59	14	1,377	3,519	8,329	298,076	5.00
Household Size																			
Two	7,321	29	501	878	1,256	6,736	1.43	59	1,496	3,514	8,348	200,628	3.54	28	1,066	2,834	7,607	160,363	4.32
Three	9,122	20	479	859	1,241	6,632	1.45	23	1,382	2,767	5,353	2,749,216	30.60	14	1,328	2,858	5,989	249,785	5.68
Four or More	12,620	24	501	894	1,332	8,909	1.48	31	1,735	3,880	9,049	1,232,953	8.29	8	1,746	4,259	9,664	298,076	4.49
Census Region																			
Northeast	5,811	25	296	734	994	6,632	1.51	47	1,223	2,749	6,897	1,232,953	12.71	16	1,022	2,838	7,020	245,817	5.05
South	9,213	20	712	1,074	1,510	8,909	1.35	23	2,046	4,046	8,930	758,206	6.64	8	1,778	4,090	9,169	262,200	4.33
Midwest	6,888	26	550	818	1,074	8,426	1.34	82	1,487	2,832	6,296	2,749,216	30.20	55	1,425	2,939	6,482	250,258	4.72
West	7,151	24	325	885	1,442	7,083	1.58	37	1,375	3,498	8,646	796,065	7.38	30	1,268	3,439	8,817	298,076	5.04
Quarter																			
Quarter1	6,906	24	561	962	1,346	8,909	1.45	48	1,735	3,583	8,157	1,232,953	9.97	14	1,596	3,757	8,580	298,076	4.53
Quarter2	7,738	20	469	818	1,184	6,469	1.44	23	1,473	3,094	7,114	2,749,216	23.49	13	1,302	3,108	7,349	214,963	4.72
Quarter3	7,226	24	474	868	1,300	8,426	1.50	31	1,539	3,390	7,645	758,206	6.85	8	1,357	3,392	7,886	262,200	4.88
Quarter4	7,193	25	495	870	1,307	6,711	1.45	72	1,537	3,338	7,760	796,065	8.18	15	1,351	3,362	7,843	274,487	4.98
% Hispanic or Latino in Segment																			
50-100%	2,680	91	704	1,154	1,531	6,736	1.29	134	2,333	4,531	9,927	687,333	6.90	32	2,015	4,490	10,387	298,076	4.46
10-<50%	7,633	24	631	1,047	1,525	8,909	1.42	34	2,030	4,165	9,499	1,232,953	8.02	10	1,797	4,104	9,663	262,200	4.22
<10%	18,750	20	373	805	1,134	8,426	1.47	23	1,357	2,885	6,710	2,749,216	15.69	8	1,199	2,954	6,955	292,298	5.09
% Black or African American in Segment																			
50-100%	1,932	24	714	1,020	1,311	6,632	1.33	34	1,988	3,749	8,719	2,749,216	43.81	10	1,700	4,013	8,672	262,200	4.39
10-<50%	5,798	25	652	976	1,427	8,909	1.39	50	1,912	3,771	8,338	1,232,953	11.17	13	1,590	3,758	8,711	224,209	4.41
<10%	21,333	20	410	834	1,238	7,624	1.49	23	1,443	3,169	7,362	796,065	6.92	8	1,325	3,216	7,624	298,076	4.94
% Owner-Occupied DUs[1] in Segment																			
50-100%	21,881	24	468	855	1,240	8,909	1.46	34	1,555	3,351	7,675	796,065	7.24	10	1,477	3,542	8,258	298,076	4.75
10-<50%	5,780	25	555	956	1,393	8,426	1.45	68	1,601	3,390	7,761	2,749,216	26.81	14	1,534	3,520	7,947	223,863	4.34
<10%	1,402	20	557	980	1,449	6,672	1.46	23	1,458	3,094	6,901	1,232,953	23.96	8	528	1,206	2,942	114,767	6.37
Combined Median Rent/Housing Value																			
1st Quintile	4,655	24	345	761	1,103	6,736	1.47	65	1,258	2,660	6,411	1,232,953	14.78	32	1,072	2,683	6,714	292,298	5.04
2nd Quintile	6,358	36	458	849	1,213	7,083	1.41	69	1,442	3,126	7,043	449,040	6.24	27	1,367	3,235	7,342	250,258	4.72
3rd Quintile	6,930	25	489	872	1,294	7,794	1.46	48	1,532	3,248	7,244	796,065	8.44	14	1,269	3,141	7,361	249,785	5.19
4th Quintile	6,334	24	530	935	1,371	8,909	1.47	34	1,767	3,741	8,494	473,564	6.41	10	1,570	3,646	8,555	298,076	4.78
5th Quintile	4,786	20	644	995	1,457	7,624	1.43	23	1,881	4,180	9,363	2,749,216	23.38	8	1,857	4,385	9,981	274,487	4.06

Table L.1 2018 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]	12,805	20	791	1,086	1,500	8,909	1.32	23	2,392	4,712	10,105	2,749,216	12.68	8	2,219	4,810	10,529	298,076	4.06
Medium to Small MSA[1]	14,091	24	304	710	1,086	6,736	1.52	47	1,147	2,541	5,766	684,810	7.15	16	1,002	2,497	5,948	292,298	5.45
Non-MSA,[1] Urban	658	61	263	677	1,044	3,285	1.59	104	881	2,351	5,743	115,335	4.11	51	765	2,207	5,571	111,681	4.77
Non-MSA,[1] Rural	1,509	26	197	534	958	2,965	1.62	37	864	2,125	4,865	338,664	7.26	30	772	2,164	5,335	150,792	5.58
Group Quarters																			
Group	255	45	370	794	1,278	2,985	1.45	59	669	1,622	2,808	27,205	2.40	55	540	1,253	2,936	17,914	2.60
Non-Group	28,808	20	495	880	1,283	8,909	1.46	23	1,570	3,370	7,706	2,749,216	11.91	8	1,409	3,413	7,950	298,076	4.77

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

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

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

Table L.2 2018 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	17,418	10	1,334	3,146	7,349	292,298	4.90	12	1,901	4,720	11,895	558,734	6.16
Pair Age Group													
12-17, 12-17	2,356	10	491	1,269	2,981	21,213	2.46	12	660	1,792	4,216	29,638	2.38
12-17, 18-25	1,764	28	1,237	2,214	4,074	24,450	1.93	28	1,657	3,206	6,092	60,861	2.16
12-17, 26-34	825	64	1,352	2,608	4,569	64,341	2.55	64	1,764	3,502	6,553	102,392	2.81
12-17, 35-49	3,848	86	1,523	3,140	6,153	59,499	2.37	86	2,059	4,404	9,029	138,107	2.71
12-17, 50+	671	329	3,993	7,863	15,325	112,119	2.12	349	5,979	12,376	24,559	153,151	2.25
18-25, 18-25	2,582	55	805	1,780	3,726	22,517	2.11	68	1,167	2,787	6,479	61,764	2.43
18-25, 26-34	836	169	1,455	3,274	6,342	52,700	2.26	175	2,149	4,829	10,339	222,956	3.39
18-25, 35-49	1,202	79	2,305	4,669	9,295	62,841	2.19	108	3,984	8,369	17,891	157,521	2.24
18-25, 50+	647	546	5,646	10,601	18,222	113,909	1.99	961	10,422	20,081	38,868	202,548	1.99
26-34, 26-34	811	423	2,867	5,500	9,993	95,302	2.27	510	4,266	9,064	16,786	224,351	2.83
26-34, 35-49	392	230	3,671	7,774	13,872	110,378	2.53	389	6,331	13,809	28,037	280,883	2.62
26-34, 50+	203	807	11,356	21,969	35,057	245,817	2.17	1,564	22,375	47,176	86,081	417,067	2.06
35-49, 35-49	664	319	3,921	7,859	13,664	181,356	3.53	472	7,137	13,990	26,365	381,283	3.59
35-49, 50+	221	960	13,055	27,294	47,320	246,090	2.07	1,906	27,395	54,683	112,119	492,838	1.97
50+, 50+	396	986	22,960	46,354	67,496	292,298	1.60	2,159	42,175	92,698	141,163	558,734	1.73
Pair Race/Ethnicity													
Hispanic or Latino	3,181	21	1,491	3,374	7,820	249,785	4.58	23	2,157	5,156	12,640	553,006	5.84
Black or African American	1,995	14	1,644	3,801	8,310	245,817	4.53	14	2,031	5,009	11,505	365,234	5.11
White	9,159	10	1,304	2,963	6,909	246,090	4.99	12	1,886	4,561	11,462	558,734	6.42
Other	1,236	21	901	2,513	6,335	292,298	7.07	33	1,067	3,818	12,756	555,914	7.64
White & Black or African American	174	67	1,703	4,537	9,932	113,909	3.76	111	2,657	6,933	15,212	177,596	4.00
White & Hispanic or Latino	685	102	1,614	3,790	9,655	223,863	4.59	126	2,335	6,005	15,417	318,588	5.03
White & Other	638	15	1,185	2,966	6,401	107,095	4.53	15	1,632	4,011	9,803	183,243	5.15
Black or African American & Hispanic or Latino	108	69	1,617	4,164	13,165	53,031	2.50	116	3,128	8,489	21,139	79,930	2.20
Black or African American & Other	103	227	1,211	2,731	7,179	82,273	4.66	279	1,660	3,778	9,354	142,068	4.91
Hispanic or Latino & Other	139	58	1,100	2,736	6,717	81,503	4.51	138	1,674	4,460	11,541	146,709	4.67
Pair Gender													
Male, Male	3,559	10	1,319	3,010	6,916	202,382	3.88	12	1,810	4,521	11,399	417,067	5.35
Female, Female	4,056	13	1,354	3,055	6,685	181,356	4.15	14	1,786	4,356	9,925	301,088	5.15
Male, Female	9,803	14	1,334	3,262	7,815	292,298	5.15	14	1,969	4,994	13,075	558,734	6.32
Household Size													
Two	4,169	28	1,000	2,473	7,100	132,700	4.62	28	1,411	3,626	11,870	294,827	5.51
Three	5,442	14	1,292	2,714	5,487	249,785	5.93	14	1,836	4,022	8,775	558,734	8.04
Four or More	7,807	10	1,671	3,968	8,950	292,298	4.47	12	2,395	6,127	14,183	555,914	5.57

Table L.2 2018 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	3,186	21	890	2,425	6,082	245,817	5.60	23	1,353	4,032	10,990	558,734	7.42
South	5,886	10	1,743	3,898	8,544	246,090	4.22	12	2,382	5,522	13,029	496,482	5.22
Midwest	4,122	55	1,402	2,787	6,055	224,209	4.80	64	2,036	4,295	9,934	437,903	5.74
West	4,224	30	1,121	3,141	7,763	292,298	5.39	34	1,599	4,595	13,166	555,914	6.68
Quarter													
Quarter1	4,233	14	1,520	3,521	7,849	292,298	4.49	14	2,195	5,244	12,836	555,914	5.39
Quarter2	4,563	13	1,231	2,861	6,799	214,963	4.89	15	1,750	4,293	11,073	553,006	6.53
Quarter3	4,272	10	1,317	3,129	7,225	249,785	5.00	12	1,925	4,797	12,214	558,734	6.53
Quarter4	4,350	15	1,312	3,121	7,416	245,817	5.19	15	1,810	4,602	11,549	492,838	6.19
% Hispanic or Latino in Segment													
50-100%	1,633	32	1,792	4,091	9,236	249,785	4.57	39	2,698	6,464	15,749	553,006	5.68
10-<50%	4,456	10	1,668	3,781	8,880	223,863	4.34	12	2,397	5,930	15,136	496,482	5.39
<10%	11,329	13	1,171	2,809	6,473	292,298	5.18	14	1,664	4,148	10,367	558,734	6.55
% Black or African American in Segment													
50-100%	1,320	10	1,648	3,905	8,251	245,817	4.30	12	2,032	5,083	11,950	376,847	5.07
10-<50%	3,542	13	1,520	3,462	7,975	224,209	4.45	14	2,176	5,315	13,211	545,341	5.77
<10%	12,556	13	1,269	2,993	7,029	292,298	5.11	15	1,806	4,535	11,483	558,734	6.38
% Owner-Occupied DUs[1] in Segment													
50-100%	13,055	10	1,435	3,320	7,666	292,298	4.85	12	2,020	4,926	12,348	558,734	6.16
10-<50%	3,533	14	1,455	3,188	7,344	223,863	4.42	14	2,052	4,973	12,362	553,006	5.30
<10%	830	16	473	1,031	2,481	110,602	7.20	16	655	1,674	4,282	211,353	8.04
Combined Median Rent/Housing Value													
1st Quintile	3,023	32	1,058	2,567	6,456	292,298	5.42	64	1,432	3,594	9,273	555,914	6.53
2nd Quintile	3,959	27	1,350	3,124	6,941	245,817	4.52	31	1,956	4,659	11,275	381,604	5.25
3rd Quintile	4,144	14	1,224	2,999	6,953	249,785	5.53	14	1,771	4,377	10,738	553,006	7.20
4th Quintile	3,615	10	1,463	3,257	7,578	223,863	5.01	12	2,027	5,153	13,135	558,734	6.38
5th Quintile	2,677	13	1,768	4,041	9,175	173,440	3.80	14	2,657	6,640	16,835	545,341	4.78
Population Density													
Large MSA[1]	7,401	10	2,060	4,403	9,509	249,785	4.10	12	3,060	7,209	16,687	558,734	5.22
Medium to Small MSA[1]	8,671	21	992	2,432	5,679	292,298	5.73	23	1,434	3,559	8,897	555,914	6.79
Non-MSA,[1] Urban	424	51	691	2,182	5,149	111,681	4.54	60	958	3,111	7,428	230,317	6.14
Non-MSA,[1] Rural	922	30	835	2,129	5,209	131,314	5.11	34	1,243	3,309	8,062	276,299	6.48
Group Quarters													
Group	195	55	570	1,253	2,936	17,914	2.64	68	733	1,678	4,248	30,000	2.39
Non-Group	17,223	10	1,358	3,184	7,395	292,298	4.87	12	1,928	4,764	12,026	558,734	6.12

[1] This step used demographic variables from screener data for all selected person pairs; DU = dwelling unit, MSA = metropolitan statistical area, NR = nonresponse adjustment, PR = pair, Res = respondent, SDU = screener dwelling unit.
[2] Q1 and Q3 refer to the first and third quartile of the weight distribution.
[3] Unequal weighting effect (UWE) is defined as $1 + [(n - 1)/n]*CV^2$, where CV = coefficient of variation of weights.

Table L.3 2018 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	17,418	12	1,901	4,720	11,895	558,734	6.16	6	1,808	4,702	11,874	479,818	6.09	6	1,804	4,704	11,877	458,425	6.05
Pair Age Group																			
12-17, 12-17	2,349	12	661	1,791	4,224	29,638	2.39	6	556	1,704	4,146	30,636	2.55	6	553	1,702	4,191	28,462	2.54
12-17, 18-25	1,764	28	1,656	3,210	6,059	60,861	2.14	18	1,605	3,306	6,066	33,276	2.01	17	1,600	3,302	6,059	33,018	2.01
12-17, 26-34	826	54	1,822	3,454	6,660	102,392	2.80	64	1,669	3,777	7,017	62,622	2.48	61	1,670	3,756	7,088	63,613	2.49
12-17, 35-49	3,836	86	2,060	4,412	9,022	153,151	2.81	81	2,082	4,494	9,257	106,750	2.60	80	2,077	4,515	9,292	98,219	2.58
12-17, 50+	678	349	5,726	11,896	24,363	140,356	2.23	120	4,716	11,853	23,242	183,985	2.47	116	4,665	11,806	23,147	185,582	2.48
18-25, 18-25	2,548	68	1,171	2,790	6,499	61,764	2.46	27	1,045	2,695	6,847	33,251	2.33	27	1,038	2,713	6,860	33,502	2.31
18-25, 26-34	851	157	2,087	4,777	10,325	222,956	3.67	193	2,060	4,694	10,626	111,742	3.02	191	2,010	4,765	10,661	96,636	2.93
18-25, 35-49	1,179	108	3,991	8,388	18,093	158,494	2.30	92	3,595	8,257	18,247	111,084	2.28	90	3,624	8,391	18,432	101,032	2.24
18-25, 50+	652	588	9,894	19,415	37,342	202,548	2.02	419	9,188	19,578	38,965	213,664	2.10	411	9,066	19,529	38,603	217,557	2.10
26-34, 26-34	827	510	4,180	8,897	16,391	224,351	2.76	264	3,876	8,618	16,976	183,631	2.77	258	3,814	8,569	16,920	185,052	2.77
26-34, 35-49	408	261	6,030	13,413	26,741	280,883	2.68	177	5,543	13,653	26,112	339,114	2.94	173	5,570	13,468	26,097	340,469	2.97
26-34, 50+	216	909	20,903	44,040	77,631	417,067	2.12	640	19,096	39,014	81,549	425,403	2.21	627	19,006	38,853	81,276	421,644	2.21
35-49, 35-49	661	389	7,127	13,894	26,554	381,283	3.60	213	5,985	13,011	26,398	406,309	3.66	207	5,994	13,046	26,556	382,567	3.64
35-49, 50+	221	1,906	24,721	53,462	112,422	492,838	1.99	854	24,153	55,042	121,516	441,007	1.91	828	23,811	54,747	120,776	400,307	1.90
50+, 50+	402	2,159	42,164	92,140	140,072	558,734	1.73	1,463	39,842	91,034	141,356	479,818	1.69	1,428	39,555	90,409	141,589	458,425	1.68
Pair Race/Ethnicity																			
Hispanic or Latino	3,198	23	2,138	5,141	12,451	553,006	5.98	8	2,075	5,261	13,072	479,818	5.76	8	2,079	5,286	13,075	454,933	5.73
Black or African American	1,958	14	2,027	4,945	11,699	437,903	5.56	6	1,968	5,055	11,811	340,011	5.21	6	1,944	5,075	11,780	336,509	5.22
White	8,863	12	1,864	4,553	11,483	558,734	6.36	9	1,924	4,755	11,495	472,696	6.18	9	1,920	4,756	11,538	458,425	6.15
Other	1,208	33	1,134	4,039	13,052	555,914	7.46	25	1,021	4,046	14,265	465,363	7.01	25	1,028	4,085	14,371	455,146	6.86
White & Black or African American	168	274	2,669	5,471	15,290	177,596	4.28	77	2,061	5,441	16,778	205,952	4.52	77	2,099	5,409	16,846	206,876	4.54
White & Hispanic or Latino	728	126	2,176	5,280	14,303	260,592	4.64	36	2,117	5,395	14,681	310,507	4.75	36	2,096	5,472	14,576	312,217	4.73
White & Other	813	66	1,753	4,395	10,052	268,689	5.98	27	1,005	2,524	6,221	259,467	7.57	27	997	2,537	6,270	260,764	7.64
Black or African American & Hispanic or Latino	132	116	2,353	7,822	21,219	134,632	2.72	63	1,273	4,291	17,970	84,108	3.08	62	1,291	4,312	17,998	83,666	3.04
Black or African American & Other	191	264	1,846	4,309	10,068	154,950	4.75	59	730	2,229	5,181	114,017	6.62	57	738	2,195	5,106	114,462	6.73
Hispanic or Latino & Other	159	138	1,593	5,432	14,225	146,709	4.17	84	1,574	3,766	11,346	147,399	4.69	84	1,582	3,847	11,235	148,985	4.73

Table L.3 2018 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	3,554	12	1,810	4,507	11,384	417,067	5.33	11	1,751	4,446	11,469	425,403	5.56	11	1,744	4,441	11,538	421,644	5.57
Female, Female	4,061	14	1,783	4,367	9,974	301,088	5.13	9	1,750	4,287	10,099	316,953	5.33	9	1,745	4,278	10,136	316,574	5.33
Male, Female	9,803	14	1,971	4,993	13,055	558,734	6.32	6	1,867	5,013	12,740	479,818	6.14	6	1,860	5,004	12,791	458,425	6.09
Household Size																			
Two	4,169	28	1,411	3,626	11,870	294,827	5.51	18	1,368	3,701	11,635	332,476	5.65	17	1,354	3,671	11,580	315,942	5.66
Three	5,442	14	1,836	4,022	8,775	558,734	8.04	7	1,777	4,071	9,006	479,818	7.44	7	1,775	4,073	9,054	454,933	7.36
Four or More	7,807	12	2,395	6,127	14,183	555,914	5.57	6	2,237	5,972	14,112	472,696	5.60	6	2,225	5,996	14,091	458,425	5.56
Census Region																			
Northeast	3,186	23	1,353	4,032	10,990	558,734	7.42	8	1,302	3,943	11,000	447,023	7.06	8	1,300	3,966	10,970	431,045	6.95
South	5,886	12	2,382	5,522	13,029	496,482	5.22	6	2,222	5,427	13,037	445,589	5.31	6	2,208	5,431	13,046	413,394	5.31
Midwest	4,122	64	2,036	4,295	9,934	437,903	5.74	41	1,985	4,328	9,917	425,403	5.49	40	1,977	4,332	9,963	421,644	5.50
West	4,224	34	1,599	4,595	13,166	555,914	6.68	27	1,578	4,734	13,023	479,818	6.66	27	1,577	4,731	13,048	458,425	6.60
Quarter																			
Quarter1	4,233	14	2,195	5,244	12,836	555,914	5.39	7	2,075	5,316	13,005	465,363	5.22	7	2,067	5,318	13,003	445,785	5.21
Quarter2	4,563	15	1,750	4,293	11,073	553,006	6.53	8	1,621	4,165	10,806	456,673	6.76	7	1,607	4,162	10,796	455,146	6.76
Quarter3	4,272	12	1,925	4,797	12,214	558,734	6.53	8	1,853	4,834	12,167	479,818	6.28	8	1,844	4,863	12,197	454,933	6.20
Quarter4	4,350	15	1,810	4,602	11,549	492,838	6.19	6	1,705	4,555	11,606	472,696	6.14	6	1,701	4,562	11,709	458,425	6.08
% Hispanic or Latino in Segment																			
50-100%	1,633	39	2,698	6,464	15,749	553,006	5.68	18	2,537	6,435	15,821	479,818	5.66	18	2,526	6,477	15,884	455,146	5.61
10-<50%	4,456	12	2,397	5,930	15,136	496,482	5.39	8	2,276	5,889	14,831	445,589	5.39	8	2,274	5,890	14,831	421,644	5.35
<10%	11,329	14	1,664	4,148	10,367	558,734	6.55	6	1,562	4,117	10,209	472,696	6.42	6	1,557	4,110	10,218	458,425	6.39
% Black or African American in Segment																			
50-100%	1,320	12	2,032	5,083	11,950	376,847	5.07	6	1,883	4,943	12,175	344,467	5.23	6	1,874	4,938	12,232	344,688	5.25
10-<50%	3,542	14	2,176	5,315	13,211	545,341	5.77	9	2,059	5,292	13,459	447,023	5.55	9	2,036	5,308	13,430	431,045	5.51
<10%	12,556	15	1,806	4,535	11,483	558,734	6.38	8	1,738	4,510	11,313	479,818	6.34	8	1,739	4,510	11,353	458,425	6.30
% Owner-Occupied DUs[1] in Segment																			
50-100%	13,055	12	2,020	4,926	12,348	558,734	6.16	8	1,963	4,928	12,181	479,818	6.08	8	1,955	4,932	12,206	458,425	6.04
10-<50%	3,533	14	2,052	4,973	12,362	553,006	5.30	6	1,899	4,937	12,657	374,894	5.21	6	1,889	4,939	12,765	376,248	5.20
<10%	830	16	655	1,674	4,282	211,353	8.04	8	595	1,607	4,089	268,861	9.62	7	590	1,591	4,055	272,429	9.74
Combined Median Rent/Housing Value																			
1st Quintile	3,023	64	1,432	3,594	9,273	555,914	6.53	27	1,372	3,571	9,299	465,363	6.35	27	1,366	3,560	9,327	445,785	6.30
2nd Quintile	3,959	31	1,956	4,659	11,275	381,604	5.25	15	1,883	4,658	11,344	376,030	5.44	15	1,879	4,650	11,447	382,561	5.46
3rd Quintile	4,144	14	1,771	4,377	10,738	553,006	7.20	6	1,613	4,363	10,659	479,818	6.94	6	1,621	4,361	10,720	458,425	6.89
4th Quintile	3,615	12	2,027	5,153	13,135	558,734	6.38	11	1,961	5,082	13,087	441,007	6.19	11	1,961	5,116	13,129	421,644	6.13
5th Quintile	2,677	14	2,657	6,640	16,835	545,341	4.78	9	2,616	6,422	16,644	447,023	4.87	9	2,596	6,433	16,620	431,045	4.84

Table L.3 2018 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]	7,401	12	3,060	7,209	16,687	558,734	5.22	6	2,964	7,130	16,759	479,818	5.08	6	2,963	7,164	16,820	458,425	5.03
Medium to Small MSA[1]	8,671	23	1,434	3,559	8,897	555,914	6.79	11	1,320	3,540	8,721	465,363	6.92	11	1,310	3,532	8,746	445,785	6.92
Non-MSA,[1] Urban	424	60	958	3,111	7,428	230,317	6.14	27	1,016	3,044	8,022	205,415	5.72	27	1,020	3,045	7,984	207,552	5.76
Non-MSA,[1] Rural	922	34	1,243	3,309	8,062	276,299	6.48	27	1,271	3,405	8,075	289,386	6.53	27	1,276	3,403	8,106	289,207	6.50
Group Quarters																			
Group	195	68	733	1,678	4,248	30,000	2.39	41	696	1,709	4,010	31,011	2.56	40	681	1,679	3,977	31,487	2.60
Non-Group	17,223	12	1,928	4,764	12,026	558,734	6.12	6	1,835	4,741	11,993	479,818	6.05	6	1,832	4,746	12,032	458,425	6.01
Pair Relationship Domain[4]																			
Parent-Child (12-14)	2,668	54	1,967	4,066	8,533	153,151	3.28	64	2,139	4,503	9,351	153,730	3.11	61	2,143	4,519	9,373	154,373	3.09
Parent-Child (12-17)	4,872	54	2,157	4,551	9,570	153,151	3.07	64	2,254	4,793	10,232	183,985	3.06	61	2,249	4,791	10,166	185,582	3.05
Parent-Child (12-20)	5,677	54	2,348	5,060	11,325	157,521	3.01	64	2,427	5,319	11,798	183,985	3.01	61	2,410	5,316	11,800	185,582	3.00
Sibling (12-14)-Sibling (15-17)	1,361	14	655	1,784	4,322	27,589	2.36	6	564	1,795	4,423	30,636	2.50	6	565	1,799	4,426	28,462	2.49
Sibling (12-17)-Sibling (18-25)	1,588	28	1,683	3,204	5,992	60,861	2.07	18	1,706	3,399	6,264	33,276	1.99	17	1,725	3,390	6,216	33,018	1.99
Spouse-Spouse/ Partner-Partner	3,271	84	2,002	6,462	17,692	558,734	5.61	27	1,809	5,927	18,067	479,818	5.50	27	1,796	5,927	18,049	455,146	5.45
Spouse-Spouse/ Partner-Partner with Children (Younger Than 18)	1,601	84	2,212	5,999	14,958	506,047	6.96	78	2,605	6,980	17,542	479,818	5.86	79	2,632	7,108	17,536	454,933	5.72

[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.

This page intentionally left blank

SAMHSA
Substance Abuse and Mental Health Services Administration

SAMHSA's mission is to reduce the impact of substance abuse and mental illness on America's communities.

1-877-SAMHSA-7 (1-877-726-4727) | 1-800-487-4889 (TDD) | www.samhsa.gov