SAMHSA
Substance Abuse and Mental Health Services Administration

2020 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

June 2022

2020 National Survey on Drug Use and Health (NSDUH) Methodological Resource Book, Section 12, Questionnaire Dwelling Unit-Level and Person Pair-Level Sampling Weight Calibration

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. Marlon Daniel 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

Center for Behavioral Health Statistics and Quality. (2022). *2020 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.

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

June 2022

Table of Contents

Chapter **Page**

Table of Contents (continued)

Chapter **Page**

Appendix

List of Tables

List of Tables (continued)

List of Exhibits

Exhibit **Page**

This page intentionally left blank

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 pair 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 2020 NSDUH editing and imputation report (Center for Behavioral Health Statistics and Quality, 2022a) 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 (2022a) for more detail.
nr	Nonresponse.
outwinsor	The proportion of weights trimmed after extreme value adjustment via winsorization.
pair age	Defined by ages of two people, not by a single person (e.g., "12-17, 18-25" is a pair age)
pair domain	A pair relationship in which 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 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 (through other evidence or through imputation).
ps	Poststratification.
QDU	Questionnaire dwelling unit: a dwelling unit in which at least one member responded to the questionnaire.
res.pr.ev	Respondent pair extreme value adjustment step. See Section 6.3.5 for more detail.
res.pr.nr	Respondent pair nonresponse adjustment step. See Section 6.3.3 for more detail.
res.pr.ps	Respondent pair poststratification adjustment step. See Section 6.3.4 for more detail.
res.qdu.ev	Respondent questionnaire dwelling unit extreme value adjustment step. See Section 6.2.5 for more detail.
res.qdu.nr	Respondent questionnaire dwelling unit nonresponse adjustment step. See Section 6.2.3 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 in which 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 in which the pair members are siblings (reported to be so or otherwise determined to be so).
spouse-spouse	A pair relationship in which the pair members are either married or living together as though married (reported to be so or otherwise determined to be so).
SSR	State sampling region.
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 weight adjustment that replaces extreme weights with the critical values used for defining low and high extreme weights.

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 2020. NSDUH is conducted under contract with RTI International.[1] The weighting team faced several challenges in this task and addressed them by resorting to innovative modifications of certain basic statistical ideas, which are listed as follows:

- 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 2020 NSDUH person-level sampling weight calibration report (Center for Behavioral Health Statistics and Quality, 2022b), 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, ratio adjustment (a form of poststratification) was performed 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[2] 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] RTI International is a trade name of Research Triangle Institute. RTI and the RTI logo are U.S. registered trademarks of Research Triangle Institute.

[2] The screening interview involves listing all eligible household members, along with their basic demographic information. Immediately after completion of the screening, 0, 1, or 2 people in the household are selected to complete the actual questionnaire interview. The first stage of data collection 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).[3] 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 et al., 2001). The experience of using the GEM with person weights is described in an earlier proceedings paper (Chen et al., 2000).

[3] 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 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. The National Survey on Drug Use and Health (NSDUH) allows for a richer analytic capability of a survey designed to ensure a nonzero pairwise probability of selection among all eligible household members in each sample household. Achieving nonzero probabilities for all pairs within sampled households permits unbiased estimation of the within-dwelling unit (DU) 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 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 2020 NSDUH person-level sampling weight calibration report (Center for Behavioral Health Statistics and Quality [CBHSQ], 2022b), 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 2020 NSDUHs. The coordinated design facilitated 50 percent overlap in third-stage units (area segments) within each successive 2-year period from 2014 through 2020. 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 half of the segments needed to be listed for the 2014 through 2020 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 2014-2020 survey sample was designed to yield

- 4,560 completed interviews in California;
- 3,300 completed interviews each in Florida, New York, and Texas;
- 2,400 completed interviews each in Illinois, Michigan, Ohio, and Pennsylvania;

- 1,500 completed interviews each in Georgia, New Jersey, North Carolina, and Virginia;
- 967 completed interviews in Hawaii; and
- 960 completed interviews in each of the remaining 37 states and the District of Columbia.

Under a stratified design with states serving as the primary strata and state sampling regions serving as the secondary strata, census tracts, census block groups within census tracts, segments within census block groups, and DUs within segments were each selected using probability proportional to size sampling. Also in the 2014-2020 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 in which 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 2014-2020 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 2014-2020 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 a 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 (2016-2020). 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 as follows, along with some adopted solutions.

Table 1.1 2016-2020 NSDUH Sample Sizes

Sample Unit		2016	2017	2018	2019	2020
SDU	Selected	173,149	184,266	193,456	210,244	536,203
	Completed	135,188	138,061	141,879	148,023	90,937
QDU	Selected	67,574	68,889	70,048	71,691	44,054
	Completed	50,095	50,328	50,373	50,488	29,200
Pair	Selected	28,033	28,778	29,063	29,818	18,461
	Completed	17,847	17,704	17,418	17,137	7,084
Person	Selected	95,607	97,667	99,111	101,509	62,515
	Completed	67,942	68,032	67,791	67,625	36,284

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016-2019 and Quarters 1 and 4, 2020.

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, 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 completed 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 (CBHSQ, 2022a). 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

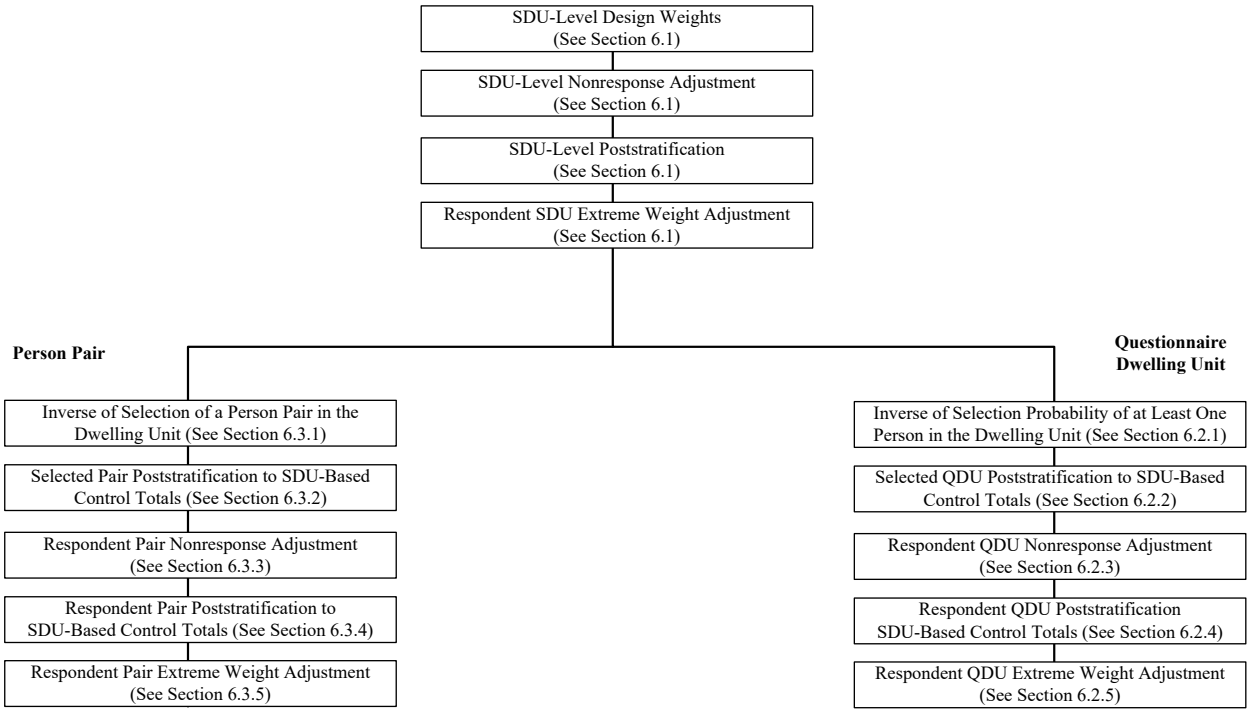
Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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 2020 NSDUH editing and imputation report (CBHSQ, 2022a). 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 QDU- and pair-level design weights through several steps of adjustment as shown in Exhibit 1.1. In the 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



Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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 (Hajek, 1971).[4] So, some treatment (such as winsorization[5]) may need to be performed 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 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

[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 ev adjustment that replaces evs with the critical values used for defining low and high evs.

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 the GEM method 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 2020 survey was the same as the one used in the 1999-2019 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 was categorically defined as zero, one, and two or more. For more details, see Chapter 5.

Before the 2020 NSDUH, quarterly samples were released, and in-person data collection was conducted in each quarter. However, for the 2020 NSDUH, the Substance Abuse and Mental Health Services Administration (SAMHSA) suspended in-person data collection on March 16, 2020, because of the coronavirus disease 2019 (COVID-19) pandemic. A small-scale data collection effort was conducted in July 2020 to test protocols to reduce the risk of COVID-19 infection through in-person data collection. Because of ongoing COVID-19 infection rates in the United States, it was nearly impossible to perform conventional in-person data collection, which could reduce the respondent sample size to an unacceptable level. Therefore, SAMHSA approved multimode data collection (in person and web based) for the 2020 NSDUH beginning in Quarter 4 (i.e., October to December 2020). In-person data collection resumed on October 1, 2020 (in locations where COVID-19 infection metrics were sufficiently low), and web-based data collection began on October 30, 2020. All sample DUs originally selected from the Quarters 2, 3, and 4 area segments for in-person data collection were released for web-based or in-person data collection in Quarter 4. Also, additional DUs were selected in some Quarter 2 and 3 segments and released to web-based data collection in Quarter 4 to partially compensate for the negative impact of the COVID-19 pandemic on data collection and response rates.

Because of the data collection interruption in 2020, and because the web-based data collection mode was introduced in Quarter 4, 2020 NSDUH person-level weighting was affected. Quarterly analysis weights for Quarters 1 and 4 were developed, then the annual analysis weights were calculated by combining data from Quarters 1 and 4 and dividing the quarterly analysis weights by a factor of 2 (CBHSQ, 2022b). The 2020 NSDUH person pair and QDU weighting were also affected. Unlike person-level weighting, only one set of person pair analysis weights and one set of QDU analysis weights were developed. The changes made to the QDU weighting and person pair weighting are discussed in Chapter 6.

[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 2002-2019 National Surveys on Drug Use and Health (NSDUHs),[7] the 2020 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 within census tracts, segments within census block groups, dwelling units (DUs) within segments, and people within DUs. 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 DU (Center for Behavioral Health Statistics and Quality, 2021).

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 2020, the pair-sampling algorithm was modified to increase the number of pairs selected in the sample (Chromy & Penne, 2002). 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 in which 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 completed 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 2020 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 2016-2020 NSDUHs

	2016			2017			2018			2019			2020		
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	173,149	135,188	78.08	184,266	138,061	74.92	193,456	141,879	73.34	210,244	148,023	70.41	536,203	90,937	16.96
QDUs															
Total QDUs	67,574	50,095	74.13	68,889	50,328	73.06	70,048	50,373	71.91	71,691	50,488	70.42	44,054	29,200	66.28
People															
Total People	95,607	67,942	71.06	97,667	68,032	69.66	99,111	67,791	68.40	101,509	67,625	66.62	62,515	36,284	58.04
12-17	22,323	17,081	76.52	22,750	17,026	74.84	22,962	16,820	73.25	23,630	16,858	71.34	14,313	6,330	44.23
18-25	22,836	16,435	71.97	23,707	16,469	69.47	24,363	16,561	67.98	25,169	16,516	65.62	15,223	8,915	58.56
26-34	15,022	10,528	70.08	15,140	10,416	68.80	15,241	10,224	67.08	15,438	10,207	66.12	9,684	6,241	64.45
35-49	19,988	13,572	67.90	20,280	13,639	67.25	20,958	13,889	66.27	21,050	13,535	64.30	12,896	8,049	62.41
50+	15,438	10,326	66.89	15,790	10,482	66.38	15,587	10,297	66.06	16,222	10,509	64.78	10,399	6,749	64.90
Non-Pairs[4]															
Total Non-Pairs	107,155	32,248	N/A	109,283	32,624	N/A	112,816	32,955	N/A	118,205	33,351	N/A	72,476	22,117	N/A
0,0	67,614	N/A	N/A	69,172	N/A	N/A	71,831	N/A	N/A	76,332	N/A	N/A	46,883	N/A	N/A
Total Singletons	39,541	32,248	81.56	40,111	32,624	81.33	40,985	32,955	80.41	41,873	33,351	79.65	25,593	22,117	86.42
0, 12-17	5,144	4,997	97.14	5,155	4,997	96.94	5,391	5,018	93.08	5,537	5,156	93.12	3,196	1,998	62.52
0, 18-25	7,647	6,895	90.17	7,858	7,079	90.09	7,895	7,019	88.90	8,172	7,258	88.82	4,697	4,870	103.7
0, 26-34	8,045	6,270	77.94	7,987	6,247	78.21	8,231	6,269	76.16	8,305	6,321	76.11	5,170	4,463	86.32
0, 35-49	8,442	6,596	78.13	8,601	6,679	77.65	8,878	6,923	77.98	8,914	6,791	76.18	5,552	5,262	94.78
0, 50+	10,263	7,490	72.98	10,510	7,622	72.52	10,590	7,726	72.96	10,945	7,825	71.49	6,978	5,524	79.16
Pairs															
Total Pairs[5]	28,033	17,847	63.66	28,778	17,704	61.52	29,063	17,418	59.93	29,818	17,137	57.47	18,461	7,084	38.37
12-17, 12-17	3,199	2,386	74.59	3,261	2,368	72.62	3,296	2,349	71.27	3,400	2,365	69.56	2,165	917	42.36
12-17, 18-25	2,548	1,774	69.62	2,679	1,758	65.62	2,680	1,764	65.82	2,779	1,706	61.39	1,743	599	34.37
12-17, 26-34	1,281	883	68.93	1,338	894	66.82	1,243	826	66.45	1,243	792	63.72	712	253	35.53
12-17, 35-49	5,829	3,930	67.42	5,845	3,870	66.21	5,939	3,836	64.59	6,101	3,753	61.51	3,602	1,377	38.23
12-17, 50+	1,123	725	64.56	1,211	771	63.67	1,117	678	60.70	1,170	721	61.62	730	269	36.85
18-25, 18-25	3,958	2,512	63.47	4,167	2,467	59.20	4,381	2,548	58.16	4,533	2,443	53.89	2,885	1,151	39.90
18-25, 26-34	1,429	886	62.00	1,443	839	58.14	1,557	851	54.66	1,623	872	53.73	960	363	37.81
18-25, 35-49	2,013	1,134	56.33	2,084	1,176	56.43	2,187	1,179	53.91	2,181	1,131	51.86	1,163	455	39.12
18-25, 50+	1,283	722	56.27	1,309	683	52.18	1,282	652	50.86	1,348	663	49.18	890	326	36.63
26-34, 26-34	1,518	905	59.62	1,551	871	56.16	1,495	827	55.32	1,515	798	52.67	1,010	436	43.17
26-34, 35-49	788	445	56.47	810	445	54.94	788	408	51.78	807	420	52.04	522	206	39.46
26-34, 50+	443	234	52.82	460	249	54.13	432	216	50.00	430	206	47.91	300	85	28.33
35-49, 35-49	1,213	627	51.69	1,233	628	50.93	1,338	661	49.40	1,273	606	47.60	879	324	36.86
35-49, 50+	490	213	43.47	474	213	44.94	490	221	45.10	501	228	45.51	299	101	33.78
50+, 50+	918	471	51.31	913	472	51.70	838	402	47.97	914	433	47.37	601	222	36.94

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.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016-2019 and Quarters 1 and 4, 2020.

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_h) 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, where $j \neq i$, 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, then used Brewer's method to select a pair, as in Case I. However, if the household had two or more people, a pair of real people was preferred in order 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 0.25.[8] 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),

$j = j^{th}$, where $j \neq i$, 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.

[8] Changes to the sample design in 2014 with respect to age group and state necessitated a new value of λ best suited for the 2014 through 2022 design. Using simulation analysis, λ was updated from the previously used 0.5 to 0.25 (CBHSQ, 2016).

The expression is now $\sum_{j \neq i} P'_{h(ij)} = P'_{h(i)}$. To maintain the original person selection probabilities despite the scale adjustment by F_s, Brewer's method was modified 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 DU 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 evs, 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 1, 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 evs.

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 earlier. 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 the 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 evs. 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 evs. 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 evs.

A large increase in the number of predictor variables in the GEM typically would result in a higher UWE, thus indicating a possible loss in precision. This was checked by comparing SUDAAN®-based (RTI International, 2013) 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 the GEM, several steps need to be followed: (1) define and create all the covariates; (2) define the extreme weights; (3) fit the GEM. The details of practical aspects of GEM implementation can be found in Chapters 4 and 5 of this report and Chapter 4 of the 2020 National Survey on Drug Use and Health person-level sampling weight calibration report (Center for Behavioral Health Statistics and Quality, 2022b).

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

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 2020 National Survey on Drug Use and Health person-level sampling weight calibration report (Center for Behavioral Health Statistics and Quality, 2022b).

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, whereas those used for SDU poststratification adjustment are counting variables. The variables of the first type (0/1 indicators) are population density,[9] 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 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 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.

[9] 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. 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 DU 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 earlier, 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 Quarters 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 within 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 Black or African American in Segments[b]
1: 50-100%, 2: 10-<50%, 3: 0-<10%[1]

Percentage of Hispanics or Latinos in Segments[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]

Interview Mode[2,b]
1: Web,[1] 2: Field

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

States[a,b,4]
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] Interview mode was included in only the res,qdu.nr step for Quarter 4 data.
[3] Segment-combined median rent and housing value (also known as the socioeconomic status indicator) is a composite measure based on rent, housing value, and percentage owner occupied.
[4] The states or district assigned to a particular model are based on census regions.
[a] Counting variable. A count of all people in the household.
[b] Binary variable.
Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

Exhibit 4.2 Definitions of Levels for Person 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]

Interview Mode[3]
1: Web,[1] 2: Field

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

(continued)

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[4]

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 (also known as the socioeconomic status indicator) is a composite measure based on rent, housing value, and percentage owner occupied.

[3] Interview mode was included in only the res,pr.nr step for Quarter 4 data.

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

* The pair member focused on.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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 $\pm 3 \times$ interquartile range (IQR). The interval is set for prespecified domains defined usually by design variables corresponding to deep stratification.[10] Similar to previous National Surveys on Drug Use and Health (NSDUHs), for the generalized exponential modeling used in the 2020 NSDUH, a more conservative (narrower) interval was defined, median $\pm 2.5 \times$ 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 evs by $b_{k(l)}$ and $b_{k(u)}$, respectively. For implementing ev control via the GEM, the variable m_k was defined as the minimum of $b_{k(u)} / w_k$ and 1 for high extreme weights, and the maximum of $b_{k(l)} / w_k$ and 1 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 1 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 the GEM. The 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

[10] Deep stratification refers to the stratification that was used in the sample design. In the case of the 2020 survey, deep stratification refers to the cross-classification of state sampling region by age group.

25-year-olds, and adult 26 years or older. Household type has seven categories as listed below:

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

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

The hierarchy is used to satisfy the minimum of 30 observations for defining the boundaries for evs. 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 2020 survey. This DU gave rise to a selected pair of respondents, one aged 49 and the other aged 51. The selection probability in this DU was 0.32400 for a respondent aged 26-49 and 0.12564 for a respondent aged 50 or older. Using the formula (2.6) in Chapter 2, the pair selection probability was computed to be 0.001043634. Therefore, the inverse of the selection probability, the pair-level design weight, was 958.190. Thus, a small pair selection probability can create a high initial weight, which is the product of the screener DU 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 1999-2019 surveys, for 2020, the extreme weight definition was based on the following hierarchy of domains:

1. Pair age group[11] (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 as follows]) 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 as follows]) by pair age group by number (0, 1, 2+) of people aged 12 to 25 in the household; and

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

4. State by pair age group by number (0, 1, 2+) of people aged 12 to 25 in the household.

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

The considerations behind the aforementioned definition for extreme weight domains are briefly introduced here. 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, because 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 earlier 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® (SAS Institute Inc., 2017), 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, some pair age categories had to be collapsed in forming domains of reasonable sample size to define extreme weights. More specifically, for the national level, all levels of number of people aged 12 to 25 for the pair age groups of 50+, 50+ and 26 to 49, 50+ were collapsed. 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 2020 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 2014-2020 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 within each state sampling region, (2) selection of census block groups within census tracts, (3) the selection of subareas or segments (comprising U.S. Census Bureau blocks) within census block groups, (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 2020 NSDUH sample design report (Center for Behavioral Health Statistics and Quality [CBHSQ], 2021).

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 SDU (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 (evs), nonresponse (nr), and coverage.

The final sample weights for screener dwelling units (SDUs) and QDU, person, and pair levels for the 2020 samples consisted of products of several factors, each representing a probability of selection at some particular stage or some form of ev, nr, or poststratification (ps) calibration adjustment. The following sections 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 2020 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 model (GEM) calibration 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, generalized exponential 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 manageable and desirable

because it assisted in bias reduction, ease of modeling, and workload reduction. Exhibit 6.3 provides more details of the data partition for generalized exponential modeling. The resulting sample sizes of selected and respondent units for the pair and QDU data partitions are shown for the 2015-2020 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 the 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 2002-2020 survey pair data, the built-in ev control feature was used for each adjustment step.

Exhibit 6.1 Summary of 2020 NSDUH QDU Sample Weight Components

SDU 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)***

QDU 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 2020 NSDUH person-level sampling weight calibration report (Center for Behavioral Health Statistics and Quality, 2022b). For computational feasibility, all weight adjustments were done using the four model groups based on census regions defined in Exhibit 6.3.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

Exhibit 6.2 Summary of 2020 NSDUH Person Pair Sample Weight Components

SDU 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)***

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 2020 NSDUH person-level sampling weight calibration report (Center for Behavioral Health Statistics and Quality, 2022b). For computational feasibility, all weight adjustments were done using the four model groups based on census regions defined in Exhibit 6.3.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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

QDU = questionnaire dwelling unit.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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

Model Group	2016		2017		2018		2019		2020	
	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs
QDU										
Northeast	13,414	9,552	14,037	9,915	13,414	9,552	14,037	9,915	9,219	5,760
South	22,287	16,810	22,628	16,901	22,287	16,810	22,628	16,901	13,057	8,862
Midwest	16,025	11,768	16,282	11,760	16,025	11,768	16,282	11,760	10,970	7,383
West	15,848	11,965	15,942	11,752	15,848	11,965	15,942	11,752	10,808	7,195
Total	67,574	50,095	68,889	50,328	67,574	50,095	68,889	50,328	44,054	29,200
	2016		**2017**		**2018**		**2019**		**2020**	
Model Group	**Selected Pairs**	**Completed Pairs**	**Selected Pairs**	**Completed Pairs**	**Selected Pairs**	**Completed Pairs**	**Selected Pairs**	**Completed Pairs**	**Selected Pairs**	**Completed Pairs**
Pair										
Northeast + South	14,543	9,182	15,072	9,284	14,543	9,182	15,072	9,284	9,185	3,561
Midwest + West	13,490	8,665	13,706	8,420	13,490	8,665	13,706	8,420	9,276	3,523
Total	28,033	17,847	28,778	17,704	28,033	17,847	28,778	17,704	18,461	7,084

QDU = questionnaire dwelling unit.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2016-2019 and Quarters 1 and 4, 2020.

Because of the data collection disruption in 2020, three changes were made to the 2020 NSDUH QDU weighting:

1. Selecting QDU poststratification to SDU-based control totals *(sel.qdu.ps, Weight component #13)* and respondent QDU nonresponse adjustment *(res.qdu.nr, Weight component #14)* were performed for Quarters 1 and 4 separately. The respondent QDU poststratification to SDU-based control totals *(res.qdu.ps, Weight component #15)* was conducted for the combined Quarter 1 and 4 data.
2. State/region three-way interactions were not included in models for all three adjustment steps because of the small sample size for the 2020 NSDUH.
3. The interview mode variable was added in the respondent QDU nonresponse adjustment models for Quarter 4.

Two changes were made to the 2020 NSDUH person pair weighting:

1. Selecting pair poststratification to SDU-based control totals *(sel.pr.ps, #13)* and respondent pair nonresponse adjustment *(res.pr.nr, #14)* were performed for Quarters 1 and 4 separately. The respondent pair poststratification to SDU-based control totals *(res.pr.ps, #15)* and respondent pair extreme value adjustment (*res.pr.ev, #16*) were conducted for the combined Quarter 1 and 4 data.
2. The interview mode variable was added in the respondent pair nonresponse adjustment models for Quarter 4.

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 for Quarters 1 and 4. 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 1 when compared with 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 census tract, (3) the probability of selecting the geographic segment within census block groups, (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 nr at the screening level, (10) DU ps 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 1 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 2020 NSDUH sample design report (CBHSQ, 2021), and for more detail on Weight Components #4 and #9 through #11, see the 2020 NSDUH person-level sampling weight calibration report (CBHSQ, 2022b).

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

Exhibit 4.1 lists the initially proposed predictor variables for the following adjustment steps via the GEM, and Exhibit 6.4 provides the initial covariates for the QDU-level adjustment modeling.

Exhibit 6.4 Covariates for 2020 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
Interview Mode[1]	Yes		2	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	
% Owner-Occupied × % Black or African American	Yes		3 × 3	4
% Owner-Occupied × % Hispanicity	Yes		3 × 3	4
% Owner-Occupied × Rent/Housing	Yes		3 × 5	8
Rent/Housing × % Black or African American	Yes		5 × 3	8
Rent/Housing × % Hispanicity	Yes		5 × 3	8
Three-Factor Effects[2]				
Race (3 Levels) × Age × Gender		Yes	8	8

[1] Interview mode was included in only the res.qdu.nr step for Quarter 4 data.

[2] State three-factor effects were used in previous years, but because of small sample sizes in the 2020 NSDUH, they were excluded.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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 DU 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 DU is selected (see Section 2.2 for details).

Table 6.2 presents the distributions and unequal weighting effect (UWE) for the design-based weight before and after the adjustment for selecting at one person in the DU.

Table 6.2 Distribution for Design-Based Weight and Unequal Weighting Effects before and after Selecting at Least One Person in the Dwelling Unit, 2020 NSDUH

Quarter	Weight	Min	25th Percentile	Median	75th Percentile	Max	Mean	Sample Size	UWE
1 and 4	Weight before QDU Selection	20	871	2,217	4,080	48,748	2,921	44,054	1.92
	QDU Selection Adjustment Factor	1.00	1.00	1.28	2.14	11.46	2.04	44,054	N/A
	Weight after QDU Selection	20	1,355	3,353	6,946	115,778	5,625	44,054	2.64

N/A = not applicable; QDU = questionnaire dwelling unit; UWE = unequal weighting effect.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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

This ps 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 nr adjustment (Weight Component #14). Exhibit 4.1 lists the initially proposed variables for generalized exponential modeling. The predictor variables are 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, so imputation was done to replace these missing data. For more details on the imputation of screener demographic information, see CBHSQ (2022b).

For the four model groups, one per census region, all main effects were kept except for some collapsing within the group quarters and household type, which both had some zero sample cells. Among two-factor effects, some segment-level variables were collapsed due to singularities, and some state × race variables were collapsed due to convergence problems.

Table 6.3 presents the distributions and UWE before and after the selected QDU-level postratification adjustment. The UWE increased slightly after adjustment for both quarters. There was little difference between the adjustments for each quarter. Data for this adjustment step were modeled separately for Quarters 1 and 4.

Table 6.3 Weight Distribution and Unequal Weighting Effects before and after Selected Questionnaire Dwelling Unit Poststratification, 2020 NSDUH

Quarter	Weight	Min	25th Percentile	Median	75th Percentile	Max	Mean	Sample Size	UWE
1	Weight before sel.qdu.ps	65	2,547	5,006	9,284	110,359	7,338	17,175	2.13
	sel.qdu.ps Adjustment Factor	0.16	0.85	0.98	1.13	4.98	1.01	17,175	N/A
	Weight after sel.qdu.ps	29	2,423	4,888	9,499	123,785	7,455	17,175	2.18
4	Weight before sel.qdu.ps	20	1,025	2,453	5,350	115,778	4,530	26,879	3.11
	sel.qdu.ps Adjustment Factor	0.20	0.88	1.00	1.14	4.96	1.02	26,879	N/A
	Weight after sel.qdu.ps	6	987	2,446	5,519	106,688	4,616	26,879	2.96
1 and 4	Weight before sel.qdu.ps	20	1,355	3,353	6,946	115,778	5,625	44,054	2.64
	sel.qdu.ps Adjustment Factor	0.16	0.87	0.99	1.13	4.98	1.02	44,054	N/A
	Weight after sel.qdu.ps	6	1,312	3,319	7,074	123,785	5,723	44,054	2.62

N/A = not applicable; sel.qdu.ps = selected questionnaire dwelling unit poststratification; UWE = unequal weighting effect.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

6.2.3 QDU Weight Component #14: Respondent QDU Nonresponse Adjustment

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

The modeling for this step followed a similar pattern of variables kept in the model as in the prior step, although more collapsing occurred among the segment-level variables and state × race interaction.

Table 6.4 presents the distributions and UWEs before and after the selected QDU-level postratification adjustment. The UWE increased slightly after adjustment for both quarters. There was little difference between the adjustments for each quarter. Data for this adjustment step were modeled separately for Quarters 1 and 4.

Table 6.4 Weight Distribution and Unequal Weighting Effects before and after Respondent Questionnaire Dwelling Unit Nonresponse Adjustment, 2020 NSDUH

Quarter	Weight	Min	25th Percentile	Median	75th Percentile	Max	Mean	Sample Size	UWE
1	Weight before res.qdu.nr	29	2,378	4,773	9,234	123,785	7,361	11,728	2.22
	res.qdu.nr Adjustment Factor	0.38	1.15	1.35	1.61	5.00	1.47	11,728	N/A
	Weight after res.qdu.nr	29	3,110	6,670	13,933	168,916	10,918	11,728	2.32
4	Weight before res.qdu.nr	18	1,066	2,644	5,951	106,688	4,941	17,472	2.90
	res.qdu.nr Adjustment Factor	0.60	1.15	1.33	1.67	5.00	1.55	17,472	N/A
	Weight after res.qdu.nr	18	1,557	3,863	8,788	138,996	7,102	17,472	2.72
1 and 4	Weight before res.qdu.nr	18	1,396	3,470	7,304	123,785	5,913	29,200	2.59
	res.qdu.nr Adjustment Factor	0.38	1.15	1.34	1.64	5.00	1.52	29,200	N/A
	Weight after res.qdu.nr	18	1,963	4,891	10,789	168,916	8,634	29,200	2.59

N/A = not applicable; res.qdu.nr = respondent questionnaire dwelling unit nonresponse; UWE = unequal weighting effect.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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

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

All proposed covariates were retained across all model groups for this step. The adjustment was performed on data for both quarters together, and reduced the median and maximum weights, as well as reduced the UWE. Data from Quarters 1 and 4 were modeled together for this adjustment step.

Table 6.5 presents the distributions and UWE before and after respondent QDU poststratification adjustment.

Table 6.5 Weight Distribution and Unequal Weighting Effects before and after Respondent Questionnaire Dwelling Unit Poststratification, 2020 NSDUH

Quarter	Weight	Min	25th Percentile	Median	75th Percentile	Max	Mean	Sample Size	UWE
1 and 4	Weight before res.qdu.ps	18	1,963	4,891	10,789	168,916	8,634	29,200	2.59
	res.qdu.ps Adjustment Factor	0.07	0.48	0.51	0.54	2.81	0.51	29,200	N/A
	Weight after res.qdu.ps	10	998	2,506	5,413	79,994	4,317	29,200	2.52

N/A = not applicable; res.qdu.ps = respondent questionnaire dwelling unit poststratification; UWE = unequal weighting effect.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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 ev adjustment (see Table 6.6). Weight Component #16 was set to 1 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, see Appendix D.

Table 6.6 Questionnaire Dwelling Unit Weight after res.qdu.ps Extreme Weight Proportions; by Model Group, 2020 NSDUH

Model Group	Extreme Weight Proportions		
	Unweighted	Weighted	Outwinsors
1. Northeast	0.71	2.14	0.40
2. Midwest	0.53	1.51	0.24
3. South	0.36	1.05	0.27
4. West	3.91	10.74	1.36

res.qdu.ps = respondent questionnaire dwelling unit poststratification; UWE = unequal weighting effect.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

6.2.6 Final QDU Analysis Weight

The final QDU analysis weight (distribution and UWE shown in Table 6.7) was constructed using the product of the final household weight, which is the product of SDU Weight Components #1 through #11 and QDU Weight Components #11 through #16. This adjusted weight is used for computing estimates of household characteristics.

Table 6.7 Final Weight Distribution of QDU Analysis Weight, 2020 NSDUH

Quarter	Weight	Min	25th Percentile	Median	75th Percentile	Max	Mean	Sample Size	UWE
1 and 4	**Final QDU Analysis Weight**	10	998	2,506	5,413	79,994	4,317	29,200	2.52

QDU = questionnaire dwelling unit; UWE = unequal weighting effect.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

6.3 Pair-Level Weight Components

Exhibit 4.2 lists the initially proposed predictor variables for the following adjustment steps via the GEM, and Exhibit 6.5 provides the initial covariates for the person pair-level adjustment modeling.

Exhibit 6.5 Covariates for 2020 NSDUH Person Pair-Level 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	3	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
Interview Mode[3]	2	1
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.

[3] Interview mode was included in only the res.pr.nr step for Quarter 4 data.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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 DU 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 earlier.

As in NSDUH years prior, Table 6.8 illustrates selecting person pairs within DUs produces high weights and large UWEs.

Table 6.8 Weight Distribution and Unequal Weighting Effects before and after Selection of a Person Pair in the Dwelling Unit, 2020 NSDUH

Quarter	Weight	Min	25th Percentile	Median	75th Percentile	Max	Mean	Sample Size	UWE
1 and 4	Weight before Pair Selection in QDU	20	899	2,251	4,174	48,748	3,034	18,461	1.96
	Pair Selection in QDU Adjustment Factor	1.02	2.48	3.98	8.47	989	8.55	18,461	N/A
	Weight after Pair Selection in QDU	33	3,440	9,318	24,001	3,625,562	25,289	18,461	9.99

N/A = not applicable; QDU = questionnaire dwelling unit; UWE = unequal weighting effect.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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 nr 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.

Modeling covariates were largely retained at proposed levels, although three factor effects were dropped for all models due to nonconvergence. Table 6.9 shows a reduced median weight as well as UWE, overall and for quarter-specific data after this adjustment. Data for this adjustment step were modeled separately for Quarters 1 and 4.

Table 6.9 Weight Distribution and Unequal Weighting Effects before and after Selected Pair Poststratification to SDU-Based Controls, 2020 NSDUH

Quarter	Weight	Min	25th Percentile	Median	75th Percentile	Max	Mean	Sample Size	UWE
1	Weight before sel.pr.ps	94	5,964	13,291	31,573	3,625,562	31,488	7,129	7.79
	sel.pr.ps Adjustment Factor	0.05	0.60	0.93	1.35	4.70	1.07	7,129	N/A
	Weight after sel.pr.ps	30	4,593	12,278	32,250	1,011,897	32,650	7,129	4.93
4	Weight before sel.pr.ps	33	2,551	7,123	19,103	3,602,542	21,389	11,332	12.14
	sel.pr.ps Adjustment Factor	0.04	0.58	0.92	1.39	4.84	1.08	11,332	N/A
	Weight after sel.pr.ps	18	2,075	6,396	19,358	952,486	21,888	11,332	6.68
1 and 4	Weight before sel.pr.ps	33	3,440	9,318	24,001	3,625,562	25,289	18,461	9.99
	sel.pr.ps Adjustment Factor	0.04	0.59	0.92	1.38	4.84	1.07	18,461	N/A
	Weight after sel.pr.ps	18	2,758	8,381	24,086	1,011,897	26,044	18,461	5.89

N/A = not applicable; sel.pr.ps = selected pair poststratification; SDU = screener dwelling unit; UWE = unequal weighting effect.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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.

Again, nearly all covariates were retained in the model with the exception of the three-factor interactions, which were all dropped due to convergence issues. Allowing for the lower response rates, the tolerance on the maximum adjustment factor was increased to 10 instead of 5 from prior years. The distribution in Table 6.10 reflects larger values after adjustment for both quarters and overall. Data for this adjustment step were modeled separately for Quarters 1 and 4.

Table 6.10 Weight Distribution and Unequal Weighting Effects before and after Respondent Pair Nonresponse Adjustment, 2020 NSDUH

Quarter	Weight	Min	25th Percentile	Median	75th Percentile	Max	Mean	Sample Size	UWE
1	Weight before res.pr.nr	34	4,371	11,514	29,383	1,011,897	29,018	3,900	4.59
	res.pr.nr Adjustment Factor	0.34	1.00	1.01	1.43	10.00	1.79	3,900	N/A
	Weight after res.pr.nr	34	5,035	14,567	43,787	3,060,480	59,683	3,900	7.80
4	Weight before res.pr.nr	22	2,069	6,291	19,087	891,523	21,912	3,177	6.53
	res.pr.nr Adjustment Factor	0.32	1.02	1.64	6.34	10.00	3.67	3,177	N/A
	Weight after res.pr.nr	37	3,772	14,676	63,228	3,433,813	78,006	3,177	7.37
1 and 4	Weight before res.pr.nr	22	3,072	9,074	24,743	1,011,897	25,812	7,084	5.31
	res.pr.nr Adjustment Factor	0.32	1.00	1.09	2.93	10.00	2.64	7,084	N/A
	Weight after res.pr.nr	34	4,554	14,598	50,518	3,433,813	67,871	7,084	7.68

N/A = not applicable; res.pr.nr = respondent pair nonresponse; UWE = unequal weighting effect.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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

This final ps 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 Chapter 1, 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 Chapter 1, 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 2020 NSDUH editing and imputation report (CBHSQ, 2022a) for more detail on the creation 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 ps.

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

This step retained the same covariates as in previous steps—no three-factor effects and all nearly all others. However, as stated earlier, the additional domains based on counts were included, and these had no convergence issues, so they were all kept in the models. The weights

were decreased after this adjustment, as shown in the distribution and UWE in Table 6.11. Data from Quarters 1 and 4 were modeled together for this adjustment step.

Table 6.11 Weight Distribution and Unequal Weighting Effects before and after Respondent Pair Poststratification to SDU-Based Controls, 2020 NSDUH

Quarter	Weight	Min	25th Percentile	Median	75th Percentile	Max	Mean	Sample Size	UWE
1 and 4	Weight before res.pr.ps	34	4,554	14,598	50,518	3,433,813	67,871	7,084	7.68
	res.pr.ps Adjustment Factor	0.02	0.36	0.51	0.72	2.20	0.57	7,084	N/A
	Weight after res.pr.ps	9	2,153	7,919	28,098	1,276,017	33,935	7,084	6.94

N/A = not applicable; res.pr.ps = respondent pair poststratification; SDU = screener dwelling unit; UWE = unequal weighting effect.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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

The extreme weight proportions for the weights up to Weight Component #15 were checked using the extreme weight domains (see Section 5.2). The built-in extreme weight control feature of the GEM implemented in previous adjustment steps successfully reduced the extreme weight proportions, although the values (see Table 6.12) were higher than the target values (see Section 7.2), so the extreme weight adjustment via the GEM was implemented using the same final set of predictor variables kept in the model for Weight Component #15. This step was somewhat successful in further reducing the extreme weight proportion in all model groups, per the distribution information shown in Table 6.13. Data from Quarters 1 and 4 were modeled together for this adjustment step.

Table 6.12 Person Pair-Level Weight after res.pr.ps Extreme Weight Proportions; by Model Group, 2020 NSDUH

Model Group	Weight	Extreme Weight Proportions		
		Unweighted	Weighted	Outwinsors
1. Northeast and South	Weight before res.pr.ps	4.38	13.11	3.79
	Weight after res.pr.ps	3.51	9.2	1.13
2. Midwest and West	Weight before res.pr.ps	4.80	17.03	3.94
	Weight after res.pr.ps	3.26	10.04	1.11

res.pr.ps = respondent pair poststratification.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

Table 6.13 Weight Distribution and Unequal Weighting Effects before and after Respondent Pair Extreme Weight Adjustment, 2020 NSDUH

Quarter	Weight	Min	25th Percentile	Median	75th Percentile	Max	Mean	Sample Size	UWE
1 and 4	Weight before res.pr.ev	9	2,153	7,919	28,098	1,276,017	33,935	7,084	6.94
	res.pr.ev Adjustment Factor	0.44	0.93	1.00	1.07	1.87	1.00	7,084	N/A
	Weight after res.pr.ev	9	2,134	7,989	28,755	1,142,106	33,935	7,084	6.83

N/A = not applicable; res.pr.ev = respondent pair extreme value; UWE = unequal weighting effect.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

6.3.6 Final Pair Analysis Weight

The final person-pair level analysis weight was constructed using the product of the final household weight, which is the product of Weight Components #1 through #11 and pair Weight Components #11 through #16. Distribution and UWE of the final pair analysis weight are shown in Table 6.14. This adjusted weight is used for computing estimates of characteristics of selected pairs within DUs.

Table 6.14 Final Weight Distribution Person Pair Analysis Weight, 2020 NSDUH

Quarter	Weight	Min	25th Percentile	Median	75th Percentile	Max	Mean	Sample Size	UWE
1 and 4	Final Pair Weight	9	2,134	7,989	28,755	1,142,106	33,935	7,084	6.83

UWE = unequal weighting effect.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

7. Evaluation of Calibration Weights

During the weight calibration process, several criteria for quality control were implemented to assess model adequacy. This chapter describes the individual procedures and summarizes their results. All tables referred to in this chapter can be found in Appendices C through F and G through J.

7.1 Response Rates

Table C.1 in Appendix C displays the final selected and responding questionnaire dwelling unit (QDU) sample sizes from the 2020 National Survey on Drug Use and Health for various national domains. This table also shows the weighted response rates.[12] Most domains reflect the overall 68.48 percent response rate, with most rates between 65 and 75 percent, although the highest response rate reached 100 percent (however, note the exceptionally small sample size of 6), from the age 12-17 category of the household type variable. The lowest response rate came from the age 12-17, 26+ category of the household type variable, with 58.24 percent.

Table G.1 in Appendix G displays the final selected and responding pair-level sample sizes and weighted response rates from the 2020 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 38.03 percent. The response rates range from a low of 27.12 percent in the pair race/ethnicity category White and Black or African American to a high of 79.39 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, whereas 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 of weighted, and 5 percent of outwinsor), a separate ev treatment step after ps could be performed. This year, because of small sample sizes caused in part by the challenges of the coronavirus disease 2019 (COVID-19) pandemic during 2020, the threshold was met and this step has been implemented for pair-level

[12] 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.

weighting to reduce final ev and outwinsor proportions. 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 D.1 and D.2 and Tables H.1 through H.3 present percentages of evs at the QDU and pair levels, 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, whereas 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, because this gave a measure of the impact of winsorization (or trimming) of ev weights (if this treatment was performed). See Sections 5.1 and 5.2 for the domains that were used to define evs.

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, 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, 2022b). Table E.1 displays QDU national domain-specific weight sums for before and after ps, as well as the desired totals to be met through ps. Table I.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 E.1, for the 2020 QDU domain household size of three, indicates a sample size of 6,446 with a total design-based weight of 19,191,247 and a census total of 19,344,396 with an initial slippage rate of -1.77 percent, which would imply a common weight adjustment approximately equal to 1.017988, if this were the only calibration control. Similarly, looking at pair data in Table I.1, the pair domain category of pair race Black or African American and Hispanic or Latino has a sample size of 63, a design-based weight of 2,245,576, and a census total of 1,820,037, showing an initial slippage of 23.38 percent. The resultant required adjustment would be approximately equal to 0.810499 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 F.1, F.2, and J.1 through J.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 J.2 and J.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 nr adjustment of respondent pairs, whereas in the ps 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 the 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, a sensitivity analysis for respondent QDU ps for the QDU weights, respondent pair ps, and extreme weight adjustment for the person pair weights was conducted. A simple baseline model was fitted with the same bounds and maximum number of iterations as were used for the chosen (more complex) final model. Then the possibility of substantial changes in point estimates and standard errors (SEs) was examined. 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 people aged 65 or older 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 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 GEMs were deemed reasonable.

Table 7.1 Estimates of Totals and SEs for Domains of Interest Based on QDU Sample, 2020

Domain	*n*	Baseline (B)[1]	Final (F)[2]	(B-F)/F% (Estimate)	(B-F)/F% (SE)
Households with Family Income					
$0-<$10,000	2,045	8,095,976 (341,731)	8,105,876 (341,103)	-0.12	0.18
$10,000-<$20,000	2,554	11,930,739 (426,677)	12,047,526 (432,109)	-0.97	-1.26
$20,000-<$30,000	2,376	10,992,537 (365,038)	11,033,380 (365,450)	-0.37	-0.11
$30,000-<$40,000	2,530	11,769,929 (393,966)	11,763,932 (396,521)	0.05	-0.64
$40,000-<$50,000	2,696	12,203,992 (389,752)	12,113,103 (384,878)	0.75	1.27
$50,000-<$75,000	4,441	19,380,212 (475,213)	19,294,690 (475,001)	0.44	0.04
$75,000-<$100,000	3,754	16,332,691 (435,204)	16,299,861 (441,735)	0.20	-1.48
$100,000+	8,804	35,354,868 (822,230)	35,402,576 (807,091)	-0.13	1.88
Households with Number of Youths (<18)					
0	16,247	86,860,425 (1,232,362)	86,871,188 (1,231,776)	-0.01	0.05
1	5,206	16,512,382 (377,028)	16,505,899 (372,360)	0.04	1.25
2	4,614	13,773,242 (337,259)	13,763,592 (337,108)	0.07	0.04
3	2,002	5,834,218 (192,296)	5,877,930 (195,304)	-0.74	-1.54
4+	1,131	3,080,678 (133,610)	3,042,335 (132,381)	1.26	0.93
Households with Insurance Coverage					
Yes	26,347	114,310,792 (1,334,651)	114,254,436 (1,334,482)	0.05	0.01
No	2,853	11,750,152 (375,276)	11,806,508 (377,163)	-0.48	-0.50
Households with Number of Older Adults (65+)					
0	24,106	88,533,814 (1,111,652)	88,749,708 (1,109,899)	-0.24	0.16
1	3,004	22,807,518 (669,345)	22,727,087 (662,524)	0.35	1.03
2	2,036	14,377,670 (498,513)	14,247,695 (497,041)	0.91	0.30
3+	54	341,942 (65,832)	336,454 (62,933)	1.63	4.61

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.
Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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, 2020

Drug	Mother User	*n*	Baseline[1]	Final[2]
Alcohol				
Lifetime	Yes	889	27.03 (3.19)	25.98 (3.14)
	No	159	22.82 (8.52)	21.11 (8.21)
	Overall	1,048	26.33 (2.88)	25.18 (2.94)
Past Year	Yes	739	25.90 (3.41)	24.86 (3.27)
	No	309	12.73 (3.84)	11.85 (3.20)
	Overall	1,048	22.05 (2.54)	20.89 (2.47)
Past Month	Yes	570	13.23 (2.89)	12.27 (2.42)
	No	478	3.52 (1.15)	2.85 (0.96)
	Overall	1,048	9.04 (1.66)	8.07 (1.44)
Cigarettes				
Lifetime	Yes	595	9.63 (1.96)	9.49 (1.88)
	No	453	4.23 (1.54)	2.66 (0.85)
	Overall	1,048	7.22 (1.22)	6.42 (1.09)
Past Year	Yes	200	15.76 (4.71)	12.62 (3.96)
	No	848	3.59 (1.03)	3.10 (0.89)
	Overall	1,048	5.35 (1.12)	4.54 (0.98)
Past Month	Yes	168	9.13 (4.80)	6.68 (3.64)
	No	880	1.65 (0.78)	1.61 (0.70)
	Overall	1,048	2.59 (0.96)	2.27 (0.79)

Note: Standard errors of prevalence estimates are provided in parentheses.
[1] Baseline refers to the weight obtained from using a main effects only model for the last two steps of calibration, res.pr.ps and res.pr.ev, and a full model for preceding steps.
[2] Final refers to the weight obtained using a full model throughout all steps of calibration.
Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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, 2020

Drug	Father User	*n*	Baseline[1]	Final[2]
Alcohol				
Lifetime	Yes	573	21.38 (3.22)	22.80 (3.54)
	No	79	18.75 (9.25)	20.23 (9.75)
	Overall	652	21.08 (3.22)	22.53 (3.47)
Past Year	Yes	481	22.35 (3.79)	24.02 (4.17)
	No	171	7.24 (2.18)	8.10 (2.40)
	Overall	652	18.04 (2.81)	19.43 (3.09)
Past Month	Yes	407	10.26 (2.89)	12.03 (3.61)
	No	245	2.26 (0.98)	2.35 (1.06)
	Overall	652	7.09 (1.89)	8.06 (2.30)
Cigarettes				
Lifetime	Yes	399	3.79 (1.58)	4.07 (1.64)
	No	253	5.56 (4.73)	5.26 (4.21)
	Overall	652	4.56 (2.01)	4.58 (1.65)
Past Year	Yes	131	4.57 (2.73)	5.28 (3.18)
	No	521	0.96 (0.40)	1.19 (0.57)
	Overall	652	1.50 (0.55)	1.80 (0.70)
Past Month	Yes	113	3.59 (2.96)	4.28 (3.44)
	No	539	0.40 (0.31)	0.56 (0.48)
	Overall	652	0.82 (0.48)	1.06 (0.63)

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.
Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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, 2020

Drug	Mother User	*n*	Baseline[1]	Final[2]
Any Illicit				
Lifetime	Yes	557	31.15 (5.22)	29.79 (5.00)
	No	491	12.02 (2.45)	11.89 (2.92)
	Overall	1,048	21.73 (3.04)	20.82 (3.03)
Past Year	Yes	173	28.04 (7.32)	27.95 (7.87)
	No	875	13.65 (2.66)	11.59 (2.17)
	Overall	1,048	15.85 (2.66)	14.25 (2.39)
Past Month	Yes	108	14.30 (5.28)	12.29 (4.86)
	No	940	7.46 (1.73)	6.41 (1.55)
	Overall	1,048	8.01 (1.63)	6.96 (1.46)
Marijuana				
Lifetime	Yes	522	19.38 (4.02)	16.96 (3.26)
	No	526	8.15 (2.15)	7.51 (2.50)
	Overall	1,048	13.47 (2.38)	11.92 (2.12)
Past Year	Yes	137	17.34 (5.71)	17.34 (6.02)
	No	911	11.67 (2.64)	9.58 (2.11)
	Overall	1,048	12.30 (2.41)	10.48 (1.94)
Past Month	Yes	92	13.94 (5.69)	12.23 (5.29)
	No	956	6.49 (1.64)	5.65 (1.47)
	Overall	1,048	7.01 (1.56)	6.19 (1.40)

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.
Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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, 2020

Drug	Father User	*n*	Baseline[1]	Final[2]
Any Illicit				
Lifetime	Yes	368	23.81 (4.14)	24.64 (4.47)
	No	284	13.78 (5.09)	13.07 (4.57)
	Overall	652	19.11 (3.39)	19.36 (3.19)
Past Year	Yes	127	28.47 (10.20)	32.59 (11.26)
	No	525	9.52 (2.13)	8.87 (1.90)
	Overall	652	12.48 (2.54)	12.75 (2.73)
Past Month	Yes	74	28.98 (15.85)	31.36 (16.87)
	No	578	5.61 (1.61)	5.59 (1.53)
	Overall	652	7.74 (2.42)	8.23 (2.67)
Marijuana				
Lifetime	Yes	327	17.70 (4.41)	18.49 (4.83)
	No	325	3.31 (1.02)	3.85 (1.29)
	Overall	652	10.00 (2.33)	10.76 (2.61)
Past Year	Yes	101	24.41 (12.69)	28.41 (14.19)
	No	551	6.79 (1.45)	6.61 (1.35)
	Overall	652	8.96 (2.30)	9.39 (2.56)
Past Month	Yes	67	32.34 (17.43)	34.87 (18.48)
	No	585	2.92 (0.66)	2.96 (0.76)
	Overall	652	5.23 (2.07)	5.80 (2.44)

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.
Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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, 2020

Drug	Parent Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Alcohol				
Lifetime	Yes	1,026	22.29 (2.57)	22.04 (2.54)
	No	632	26.25 (3.59)	26.40 (3.79)
	Overall	1,658	23.77 (1.93)	23.67 (2.12)
Past Year	Yes	1,026	18.94 (2.49)	18.68 (2.48)
	No	632	22.08 (3.05)	22.19 (3.20)
	Overall	1,658	20.11 (1.70)	20.00 (1.80)
Past Month	Yes	1,026	8.68 (1.93)	7.97 (1.83)
	No	632	8.46 (2.01)	9.21 (2.24)
	Overall	1,658	8.60 (1.34)	8.44 (1.37)
Cigarettes				
Lifetime	Yes	1,026	7.08 (1.70)	6.55 (1.60)
	No	632	5.61 (1.28)	5.44 (1.36)
	Overall	1,658	6.53 (1.11)	6.14 (1.01)
Past Year	Yes	1,026	4.39 (1.00)	3.85 (0.93)
	No	632	3.54 (1.26)	3.64 (1.39)
	Overall	1,658	4.07 (0.77)	3.77 (0.78)
Past Month	Yes	1,026	2.05 (0.84)	1.76 (0.77)
	No	632	2.32 (1.23)	2.68 (1.37)
	Overall	1,658	2.15 (0.70)	2.10 (0.71)

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.
Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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, 2020

Drug	Parent Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Any Illicit				
Lifetime	Yes	1,026	20.03 (2.73)	19.46 (2.48)
	No	632	21.17 (2.95)	21.10 (3.16)
	Overall	1,658	20.45 (1.95)	20.07 (1.85)
Past Year	Yes	1,026	14.87 (2.17)	13.72 (1.96)
	No	632	13.48 (2.50)	13.42 (2.65)
	Overall	1,658	14.35 (1.65)	13.61 (1.57)
Past Month	Yes	1,026	8.53 (1.50)	8.11 (1.65)
	No	632	6.39 (1.63)	6.28 (1.70)
	Overall	1,658	7.73 (1.17)	7.43 (1.25)
Marijuana				
Lifetime	Yes	1,026	13.21 (2.24)	12.35 (2.07)
	No	632	10.60 (2.03)	10.61 (2.25)
	Overall	1,658	12.23 (1.58)	11.70 (1.54)
Past Year	Yes	1,026	12.47 (2.25)	11.42 (2.05)
	No	632	9.05 (1.86)	8.67 (1.79)
	Overall	1,658	11.20 (1.59)	10.39 (1.46)
Past Month	Yes	1,026	7.35 (1.66)	7.10 (1.84)
	No	632	5.35 (1.52)	5.31 (1.64)
	Overall	1,658	6.60 (1.24)	6.43 (1.34)

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.
Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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, 2020

Drug	Mother Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Alcohol				
Lifetime	0 times	105	19.59 (10.42)	20.21 (10.22)
	1-2 times	182	20.95 (5.49)	20.47 (5.24)
	A few times	276	24.55 (5.14)	25.81 (5.73)
	Many times	439	30.02 (4.49)	26.67 (3.98)
	Overall	1,002	25.73 (2.86)	24.57 (2.91)
Past Year	0 times	105	5.95 (3.63)	7.49 (4.34)
	1-2 times	182	20.18 (5.46)	19.54 (5.20)
	A few times	276	22.92 (5.18)	23.68 (5.75)
	Many times	439	26.01 (4.52)	22.93 (3.88)
	Overall	1,002	21.73 (2.59)	20.58 (2.53)
Past Month	0 times	105	3.96 (3.30)	4.38 (3.74)
	1-2 times	182	8.80 (4.06)	9.13 (4.10)
	A few times	276	6.51 (2.97)	6.18 (2.90)
	Many times	439	11.81 (2.77)	9.68 (2.15)
	Overall	1,002	8.94 (1.68)	7.95 (1.45)
Cigarettes				
Lifetime	0 times	105	0.41 (0.27)	0.25 (0.19)
	1-2 times	182	4.21 (1.84)	3.65 (1.40)
	A few times	276	10.90 (3.55)	9.67 (3.24)
	Many times	439	8.42 (1.96)	7.49 (1.64)
	Overall	1,002	7.35 (1.26)	6.51 (1.14)
Past Year	0 times	105	0.35 (0.26)	0.22 (0.19)
	1-2 times	182	2.78 (1.64)	2.22 (1.18)
	A few times	276	9.84 (3.53)	8.54 (3.18)
	Many times	439	5.42 (1.65)	4.52 (1.24)
	Overall	1,002	5.49 (1.16)	4.67 (1.02)
Past Month	0 times	105	0.00 (0.00)	0.00 (0.00)
	1-2 times	182	1.09 (0.68)	1.18 (0.78)
	A few times	276	6.61 (3.06)	5.72 (2.72)
	Many times	439	1.65 (1.17)	1.32 (0.73)
	Overall	1,002	2.64 (0.99)	2.32 (0.82)

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.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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, 2020

Drug	Father Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Alcohol				
Lifetime	0 times	90	6.02 (2.85)	5.97 (2.77)
	1-2 times	142	11.08 (3.98)	11.60 (4.02)
	A few times	186	25.17 (7.31)	26.31 (8.17)
	Many times	188	29.41 (6.07)	30.41 (5.90)
	Overall	606	19.73 (3.27)	20.66 (3.43)
Past Year	0 times	90	3.52 (1.74)	3.33 (1.42)
	1-2 times	142	10.19 (3.94)	10.41 (3.93)
	A few times	186	20.07 (7.07)	21.31 (8.07)
	Many times	188	26.02 (5.81)	27.05 (5.63)
	Overall	606	16.63 (2.84)	17.47 (3.06)
Past Month	0 times	90	2.19 (1.58)	1.58 (1.01)
	1-2 times	142	3.24 (1.57)	3.21 (1.72)
	A few times	186	5.60 (2.25)	5.09 (2.08)
	Many times	188	8.83 (3.66)	10.75 (4.49)
	Overall	606	5.35 (1.39)	5.69 (1.61)
Cigarettes				
Lifetime	0 times	90	1.00 (0.72)	1.09 (0.80)
	1-2 times	142	8.58 (7.66)	8.07 (6.87)
	A few times	186	0.45 (0.21)	0.43 (0.22)
	Many times	188	7.95 (3.45)	8.66 (3.62)
	Overall	606	4.83 (2.16)	4.90 (1.79)
Past Year	0 times	90	1.00 (0.72)	1.09 (0.80)
	1-2 times	142	0.54 (0.38)	0.59 (0.42)
	A few times	186	0.19 (0.11)	0.17 (0.11)
	Many times	188	4.57 (2.12)	5.60 (2.66)
	Overall	606	1.61 (0.59)	1.95 (0.77)
Past Month	0 times	90	0.47 (0.44)	0.46 (0.45)
	1-2 times	142	0.00 (0.00)	0.00 (0.00)
	A few times	186	0.00 (0.00)	0.00 (0.00)
	Many times	188	2.97 (1.93)	3.84 (2.48)
	Overall	606	0.88 (0.52)	1.15 (0.69)

Note: Standard errors of prevalence estimates are provided in parentheses.
[1] Baseline refers to the weight obtained from using a main effects only model for the last two steps of calibration, res.pr.ps and res.pr.ev, and a full model for preceding steps.
[2] Final refers to the weight obtained using a full model throughout all steps of calibration.
Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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, 2020

Drug	Mother Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Any Illicit				
Lifetime	0 times	105	8.27 (3.63)	8.15 (3.96)
	1-2 times	182	14.34 (4.10)	14.35 (3.98)
	A few times	276	22.55 (6.25)	23.09 (6.32)
	Many times	439	27.04 (5.22)	24.57 (5.12)
	Overall	1,002	21.35 (3.09)	20.34 (3.09)
Past Year	0 times	105	4.95 (3.38)	5.23 (3.78)
	1-2 times	182	11.24 (3.78)	10.76 (3.29)
	A few times	276	17.61 (6.37)	17.21 (6.33)
	Many times	439	19.43 (4.28)	16.07 (3.37)
	Overall	1,002	15.75 (2.71)	14.10 (2.44)
Past Month	0 times	105	0.35 (0.26)	0.22 (0.19)
	1-2 times	182	5.21 (2.34)	5.76 (2.62)
	A few times	276	10.88 (5.53)	9.42 (4.61)
	Many times	439	8.95 (2.27)	7.21 (1.95)
	Overall	1,002	7.75 (1.64)	6.67 (1.47)
Marijuana				
Lifetime	0 times	105	0.98 (0.65)	0.76 (0.56)
	1-2 times	182	10.70 (3.74)	10.24 (3.25)
	A few times	276	12.30 (5.49)	10.78 (4.60)
	Many times	439	19.41 (4.12)	17.18 (3.80)
	Overall	1,002	13.78 (2.46)	12.16 (2.21)
Past Year	0 times	105	0.98 (0.65)	0.76 (0.56)
	1-2 times	182	10.21 (3.71)	9.62 (3.22)
	A few times	276	11.99 (5.50)	10.57 (4.61)
	Many times	439	17.06 (4.17)	14.10 (3.20)
	Overall	1,002	12.57 (2.49)	10.67 (2.02)
Past Month	0 times	105	0.35 (0.26)	0.22 (0.19)
	1-2 times	182	5.07 (2.35)	5.52 (2.63)
	A few times	276	10.63 (5.53)	9.08 (4.62)
	Many times	439	7.78 (2.16)	6.62 (1.90)
	Overall	1,002	7.14 (1.61)	6.28 (1.45)

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.
Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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, 2020

Drug	Father Talked about Dangers with Child	*n*	Baseline[1]	Final[2]
Any Illicit				
Lifetime	0 times	90	5.77 (3.16)	5.87 (2.81)
	1-2 times	142	24.43 (8.93)	24.07 (7.57)
	A few times	186	16.43 (6.90)	18.04 (7.94)
	Many times	188	25.79 (5.39)	25.40 (5.07)
	Overall	606	19.69 (3.62)	20.02 (3.42)
Past Year	0 times	90	5.36 (3.13)	5.23 (2.72)
	1-2 times	142	7.70 (3.90)	7.52 (3.51)
	A few times	186	13.69 (6.82)	14.50 (7.89)
	Many times	188	20.16 (5.10)	20.03 (4.79)
	Overall	606	12.63 (2.68)	12.88 (2.90)
Past Month	0 times	90	3.77 (2.98)	3.36 (2.45)
	1-2 times	142	1.66 (1.04)	1.68 (0.99)
	A few times	186	9.76 (6.76)	10.90 (7.98)
	Many times	188	14.53 (5.37)	15.08 (5.21)
	Overall	606	7.99 (2.58)	8.55 (2.88)
Marijuana				
Lifetime	0 times	90	4.56 (3.04)	4.18 (2.55)
	1-2 times	142	6.29 (3.78)	6.15 (3.38)
	A few times	186	10.87 (6.64)	13.19 (7.88)
	Many times	188	16.54 (3.83)	16.90 (3.73)
	Overall	606	10.27 (2.45)	11.08 (2.76)
Past Year	0 times	90	4.56 (3.04)	4.18 (2.55)
	1-2 times	142	6.29 (3.78)	6.15 (3.38)
	A few times	186	10.10 (6.67)	11.49 (7.91)
	Many times	188	13.46 (3.68)	13.59 (3.51)
	Overall	606	9.18 (2.43)	9.62 (2.72)
Past Month	0 times	90	3.34 (2.93)	2.72 (2.34)
	1-2 times	142	0.37 (0.34)	0.43 (0.40)
	A few times	186	8.63 (6.74)	10.04 (8.00)
	Many times	188	8.64 (3.54)	9.56 (3.77)
	Overall	606	5.60 (2.22)	6.30 (2.65)

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.
Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

References

American Association for Public Opinion Research. (2016). *Standard definitions: Final dispositions of case codes and outcome rates for surveys (9th edition)*. https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf

Center for Behavioral Health Statistics and Quality. (2014). *2013 National Survey on Drug Use and Health: Methodological resource book (Section 2, Sample design report)*. Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (2016). *2015 National Survey on Drug Use and Health (NSDUH) methodological resource book, Section 2: Sample design report*. 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.* Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (2021). *2020 National Survey on Drug Use and Health (NSDUH) methodological resource book, Section 2: Sample design report*. Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (2022a). *2020 National Survey on Drug Use and Health (NSDUH) methodological resource book, Section 10: Editing and imputation report*. Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (2022b). *2020 National Survey on Drug Use and Health (NSDUH) methodological resource book, Section 11: Person-level sampling weight calibration*. Substance Abuse and Mental Health Services Administration.

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). American Statistical Association. http://www.asasrms.org/Proceedings/papers/2000_100.pdf

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*. American Statistical Association. http://www.asasrms.org/Proceedings/y2001/Proceed/00222.pdf

Chromy, J. R., & Penne, M. A. (2002). Pair sampling in household surveys. In *Proceedings of the 2002 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Atlanta, GA*. American Statistical Association. http://www.asasrms.org/Proceedings/y2002/Files/JSM2002-001055.pdf

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). American Statistical Association. http://www.asasrms.org/Proceedings/papers/2000_099.pdf

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). 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). 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*. American Statistical Association. http://www.asasrms.org/Proceedings/y2001/Proceed/00447.pdf

RTI International. (2013). *SUDAAN® language manual, release 11.0.1*.

SAS Institute Inc. (2017). *SAS/STAT software: Release 14.1*.

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. Work by RTI was performed under Contract No. HHSS283201700002C. Marlon Daniel served as government project officer and as the contracting officer representative, and David Hunter served as the RTI project director.

Contributors to this report at SAMHSA were Rong Cai and Jingsheng Yan.

Contributors to this report at RTI included Patrick Chen, Harper Gordek, and Matthew Westlake. Also at RTI, report and web production staff included Margaret Johnson and Teresa Mink.

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 (e.g., selected dwelling units or selected persons). The distance minimized under the generalized exponential model (GEM), subject to calibration constraints, is given by the distance function below:

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

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, where $\ell_k < c_k < u_k$. 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 model parameters λ from GEM which is a *p*-dimensional vector with *p* as the number of predictor variables; 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 ≤ *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 $k = 1,\ldots,n$,

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

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 $\ell_k = L,\ u_k = U$, and $c_k = 1$ for all *k*. The new GEM was introduced by Folsom and Singh (2000).

A.2 Generalized Exponential Model 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 $b_k = \text{median}\{d_k\} \pm 3 * \text{IQR}$ 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, let L_1 and U_1 be user-specified lower and upper bounds for high extreme weights, L_2 and U_2 the lower and upper bounds for nonextreme weights, and L_3 and U_3 the lower and upper bounds for low extreme weights. The sample is first divided into two parts: the nonextreme weight subsample and the extreme weight subsample. For nonextreme weights, the following are set: $L_2 = 1,\ C_2 = \rho^{-1},$ and $U_2 = U > \rho^{-1}$, where ρ is the overall response propensity. For extreme weights with high weights, $\ell_k = L_1 m_k,\ c_k = \rho^{-1} m_k,$ and $u_k = U_1 m_k$, where $m_k = b_k / d_k$ and $1 \le L_1 < \rho^{-1} = C_1 < U_1$ are prescribed numbers. Similarly, for extreme weights with low weights, $\ell_k = L_3 m_k,\ c_k = \rho^{-1} m_k,\ u_k = U_3 m_k$, and $1 \le L_3 < \rho^{-1} = C_3 < U_3$.

For the poststratification adjustment, the following weights are set: for nonextreme weights, $\ell_k = L_2,\ c_k = C_2 = 1,$ and $u_k = U_2$; for high extreme weights, $\ell_k = L_1 m_k,\ c_k = m_k,$ and $u_k = U_1 m_k$; and similarly, for low extreme weights, $\ell_k = L_3 m_k,\ c_k = m_k,$ 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, $\ell_k = L$, $u_k = U$, and $c_k = C = 1$, such that $L < 1 < U$. 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 1.

[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

Unlike for 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 who were members of the 10 selected pair domains given as follows:

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 aged 17 or younger living in the household.

B.1 Derivation of Questionnaire Dwelling Unit-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 desired control totals. 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 people aged 12 or older were provided by the Population Estimates Program of the U.S. Census Bureau. One state-level file contained estimates of the population broken down by levels of month (12), Hispanicity (2), race (6), gender (2), and age into single-year estimates from people aged 0 to 84, and a combined age group estimate for people aged 85 or older.

The breakdown received from the census did not match the levels of the domains for which control totals were desired. To account for this, levels were collapsed. From these altered data, datasets were created with model group-specific control totals. Observations in these datasets corresponded to a breakdown by quarter (4), Hispanicity (2), race (5), gender (2), age (6), 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 Screener Dwelling Unit 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 people aged 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 people aged 12 or older and could be utilized as a basis for constructing controls at the QDU and person pair levels.

B.1.2 Questionnaire Dwelling Unit Level

B.1.2.1 Deriving Questionnaire Dwelling Unit-Level Poststratification Control Totals from Adjusted Screener Dwelling Unit Weights

Because there were no controls for QDU-level poststratification available directly, the adjusted SDU weights were used. 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

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

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.

B.1.2.2 Adjusting Questionnaire Dwelling Unit-Level Data to the Control Totals

As was done for the SDU data, the QDU-level data were adjusted via calibration in the GEM of sample weights to multiple control totals. Each QDU received an adjustment factor, similar to that described for the SDU weight in Section B.1.1.2. The controls utilized in this calibration were based on the SDU weight as described in Section B.1.2.1. The adjusted weight was representative of the civilian, noninstitutionalized population estimates for people aged 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 Screener Dwelling Unit 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. 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 2020 NSDUH editing and imputation report (Center for Behavioral Health Statistics and Quality, 2022a) 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 were adjusted via the 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 people aged 12 or older, and the sum of all final weights in a domain corresponded to the control total for that domain.

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

This page intentionally left blank

Table C.1 2020 NSDUH QDU-Level Response Rates

Domain	Selected QDU	Respondent QDU	QDU-Level Response Rate[1]
Total	44,054	29,200	68.48
Census Region			
Northeast	9,219	5,760	64.41
South	13,057	8,862	69.71
Midwest	10,970	7,383	69.30
West	10,808	7,195	68.70
Quarter			
Quarter 1	17,175	11,728	67.42
Quarter 4	26,879	17,472	69.57
Household Type			
12-17, 18-25, 26+	3,202	1,952	64.12
12-17, 18-25	25	13	65.34
12-17, 26+	9,571	5,309	58.24
18-25, 26+	7,313	4,888	66.87
12-17	6	6	100.00
18-25	3,290	2,565	76.45
26+	20,647	14,467	70.03
Race/Ethnicity of Householder			
Hispanic or Latino White	5,231	3,399	65.98
Hispanic or Latino Black or African American	217	136	65.25
Hispanic or Latino Other	481	311	69.17
Non-Hispanic or Latino White	30,091	20,007	69.10
Non-Hispanic or Latino Black or African American	3,948	2,614	68.51
Non-Hispanic or Latino Other	4,086	2,733	67.30
% Hispanic or Latino in Segment			
50-100%	2,996	1,940	67.00
10-<50%	10,498	7,001	68.12
<10%	30,560	20,259	68.82
% Black or African American in Segment			
50-100%	2,352	1,554	67.19
10-<50%	7,740	5,296	70.38
<10%	33,962	22,350	68.08
% Owner-Occupied DUs in Segment			
50-100%	32,916	21,413	67.92
10-<50%	8,804	6,097	69.76
<10%	2,334	1,690	70.76
Combined Median Rent/Housing Value			
1st Quintile	6,382	4,260	69.90
2nd Quintile	9,169	6,193	69.57
3rd Quintile	10,265	6,805	68.15
4th Quintile	9,734	6,401	67.04
5th Quintile	8,504	5,541	68.26
Population Density			
Large MSA	18,808	12,319	67.47
Medium to Small MSA	22,398	14,975	69.58
Non-MSA, Urban	1,229	811	68.95
Non-MSA, Rural	1,619	1,095	70.39
Group Quarters			
Group	136	128	96.14
Non-Group	43,918	29,072	68.41
Household Size			
One	5,358	4,141	74.91
Two	18,854	13,060	68.66
Three	10,647	6,446	62.15
Four or More	9,195	5,553	62.44

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 the selected QDU poststratification adjustment. This weight is the product of WT1*...*WT11*DUWT12*DUWT13.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

This page intentionally left blank

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

This page intentionally left blank

Table D.1 2020 NSDUH Selected QDU-Level Proportions of Extreme Values and Outwinsors

Domain	*n*	SDU-Level Weights (SDUWT: WT1*...*WT11)			Before sel.qdu.ps (SDUWT*DUWT12)			After sel.qdu.ps (SDUWT*DUWT12*DUWT13)		
		% Unweighted	% Weighted[1]	% Outwinsor[2]	% Unweighted	% Weighted[1]	% Outwinsor[2]	% Unweighted	% Weighted[1]	% Outwinsor[2]
Total	44,054	1.89	6.32	1.56	2.03	7.00	1.74	2.07	5.98	1.22
Census Region										
Northeast	9,219	1.88	6.85	1.89	2.27	7.75	1.96	2.84	8.77	2.13
South	13,057	1.91	6.25	1.48	1.72	6.17	1.21	1.55	5.04	0.88
Midwest	10,970	1.45	5.22	1.19	1.70	6.07	1.41	1.77	5.67	1.26
West	10,808	2.30	7.02	1.77	2.53	8.74	2.79	2.36	5.76	1.08
Quarter										
Quarter 1	17,175	2.61	6.64	1.78	2.81	6.69	1.65	3.43	7.74	1.86
Quarter 4	26,879	1.42	5.98	1.33	1.53	7.31	1.83	1.21	4.16	0.56
Household Type										
12-17, 18-25, 26+	3,202	1.75	6.84	1.56	1.72	6.77	1.54	2.53	10.29	2.45
12-17, 18-25	25	4.00	18.01	6.10	0.00	0.00	0.00	4.00	24.62	6.65
12-17, 26+	9,571	1.69	6.12	1.60	1.69	6.09	1.59	1.92	6.73	1.56
18-25, 26+	7,313	2.23	7.32	1.79	2.30	7.48	1.70	2.24	7.59	1.55
12-17	6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18-25	3,290	2.71	6.81	1.66	2.74	7.01	1.44	3.19	8.48	2.40
26+	20,647	1.74	5.83	1.42	2.02	7.08	1.79	1.83	5.22	0.98
Race/Ethnicity of Householder										
Hispanic or Latino White	5,231	2.03	6.63	1.61	1.97	7.22	1.60	2.20	6.24	1.07
Hispanic or Latino Black or African American	217	12.44	36.69	10.88	11.06	33.64	10.21	10.60	28.15	8.50
Hispanic or Latino Other	481	6.24	25.38	9.69	4.99	25.84	9.09	6.44	21.70	5.40
Non-Hispanic or Latino White	30,091	1.31	3.98	0.94	1.66	5.75	1.37	1.53	4.30	0.83
Non-Hispanic or Latino Black or African American	3,948	4.46	11.84	2.92	3.85	10.77	2.89	4.36	10.86	2.16
Non-Hispanic or Latino Other	4,086	2.37	8.30	1.72	2.23	6.63	1.53	2.77	8.35	2.18

(continued)

Table D.1 2020 NSDUH Selected QDU-Level Proportions of Extreme Values and Outwinsors (continued)

Domain	*n*	SDU-Level Weights (SDUWT: WT1*...*WT11)			Before sel.qdu.ps (SDUWT*DUWT12)			After sel.qdu.ps (SDUWT*DUWT12*DUWT13)		
		% Unweighted	% Weighted[1]	% Outwinsor[2]	% Unweighted	% Weighted[1]	% Outwinsor[2]	% Unweighted	% Weighted[1]	% Outwinsor[2]
% Hispanic or Latino in Segment										
50-100%	2,996	1.94	6.97	1.67	1.77	5.56	1.60	2.14	5.70	1.18
10-<50%	10,498	1.68	5.48	1.21	1.49	4.61	0.89	1.90	5.24	0.98
<10%	30,560	1.95	6.60	1.70	2.24	8.17	2.10	2.13	6.32	1.33
% Black or African American in Segment										
50-100%	2,352	4.68	12.48	3.37	4.34	13.32	3.67	4.00	11.35	2.00
10-<50%	7,740	1.99	6.77	1.61	1.60	4.67	0.96	2.04	5.68	1.25
<10%	33,962	1.67	5.54	1.35	1.96	6.99	1.75	1.95	5.53	1.14
% Owner-Occupied DUs in Segment										
50-100%	32,916	1.49	5.22	1.24	1.80	6.53	1.51	1.68	4.92	0.96
10-<50%	8,804	2.64	7.72	1.87	2.36	7.14	1.89	3.02	8.47	1.87
<10%	2,334	4.71	13.11	3.74	3.98	11.95	3.88	3.98	10.03	2.15
Combined Median Rent/Housing Value										
1st Quintile	6,382	2.63	7.61	1.80	2.63	8.66	2.42	2.66	5.98	1.20
2nd Quintile	9,169	1.97	5.04	1.11	2.16	6.13	1.24	1.92	4.29	0.93
3rd Quintile	10,265	1.88	6.33	1.52	1.89	6.41	1.59	1.92	5.57	1.07
4th Quintile	9,734	1.62	5.98	1.56	1.97	6.68	1.66	2.03	5.99	1.20
5th Quintile	8,504	1.54	7.06	1.89	1.66	7.70	2.01	2.02	8.25	1.75
Population Density										
Large MSA	18,808	2.08	7.14	1.78	1.98	7.25	1.86	2.24	7.06	1.44
Medium to Small MSA	22,398	1.49	4.93	1.16	1.76	5.63	1.30	1.61	4.02	0.86
Non-MSA, Urban	1,229	2.44	5.05	0.93	2.12	4.41	0.71	1.79	3.15	0.69
Non-MSA, Rural	1,619	4.63	10.36	3.07	6.18	19.07	5.16	6.67	13.07	2.32
Group Quarters										
Group	136	3.68	7.51	1.55	5.15	16.63	3.60	11.03	28.10	7.18
Non-Group	43,918	1.88	6.31	1.56	2.02	6.97	1.73	2.04	5.93	1.21
Household Size										
One	5,358	1.70	5.32	1.33	1.89	6.62	1.51	1.38	3.65	0.58
Two	18,854	1.74	5.50	1.33	2.27	7.46	1.96	2.23	6.25	1.30
Three	10,647	2.02	7.36	1.89	1.82	6.56	1.59	2.12	7.48	1.60
Four or More	9,195	2.13	7.20	1.72	1.85	6.51	1.51	2.10	7.88	1.77

DU = dwelling unit; MSA = metropolitan statistical area; ps = poststratification adjustment; QDU = questionnaire dwelling unit; SDU = screener dwelling unit; sel = selected.

[1] Weighted extreme value percentage: $100*\sum_k w_{ek}/\sum_k w_k$, where w_{ek} denotes the weight for extreme weights and w_k denotes the weight for extreme weights and nonextreme weights.

[2] Outwinsor weight percentage: $100*\sum_k |w_{ek} - b_k|/\sum_k w_k$, where b_k denotes the cut-off point for defining the extreme weight.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

Table D.2 2020 NSDUH Respondent QDU-Level Proportions of Extreme Values and Outwinsors

Domain	*n*	Before res.qdu.nr (SDUWT*DUWT12*DUWT13)			After res.qdu.nr (SDUWT*DUWT12*...*DUWT14)			Final Weight after res.qdu.ps (SDUWT*DUWT12*...*DUWT15)		
		% Unweighted	% Weighted[1]	% Outwinsor[2]	% Unweighted	% Weighted[1]	% Outwinsor[2]	% Unweighted	% Weighted[1]	% Outwinsor[2]
Total	29,200	1.86	5.62	1.17	2.27	6.97	1.56	1.29	3.44	0.48
Census Region										
Northeast	5,760	2.80	8.79	2.48	3.32	11.57	2.92	0.66	2.01	0.33
South	8,862	1.41	4.86	0.87	1.46	4.10	0.88	0.30	0.92	0.17
Midwest	7,383	1.31	4.65	0.96	2.07	8.33	1.84	0.47	1.42	0.23
West	7,195	2.24	5.58	0.96	2.63	7.03	1.39	3.85	10.69	1.36
Quarter										
Quarter 1	11,728	3.04	7.21	1.80	3.74	9.96	2.33	1.78	3.37	0.44
Quarter 4	17,472	1.08	4.02	0.53	1.28	3.88	0.76	0.96	3.50	0.52
Household Type										
12-17, 18-25, 26+	1,952	2.05	8.66	2.18	2.56	10.76	2.79	1.18	4.01	0.62
12-17, 18-25	13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12-17, 26+	5,309	2.00	7.20	1.82	2.96	11.28	2.80	1.62	5.24	0.74
18-25, 26+	4,888	2.31	8.41	1.78	2.43	9.24	2.21	1.74	6.47	1.09
12-17	6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18-25	2,565	3.08	8.72	2.31	3.43	12.65	3.90	1.87	4.72	0.72
26+	14,467	1.42	4.66	0.87	1.71	5.42	1.06	0.93	2.60	0.33
Race/Ethnicity of Householder										
Hispanic or Latino White	3,399	1.97	5.90	1.09	2.68	8.51	1.72	2.21	6.53	0.84
Hispanic or Latino Black or African American	136	11.03	35.31	12.70	11.03	31.98	7.15	8.82	18.73	2.30
Hispanic or Latino Other	311	5.79	18.16	3.83	5.14	18.31	2.73	3.86	9.27	1.84
Non-Hispanic or Latino White	20,007	1.38	4.01	0.75	1.75	5.17	1.13	0.84	2.18	0.28
Non-Hispanic or Latino Black or African American	2,614	3.98	10.35	2.11	4.06	10.48	2.51	1.57	3.57	0.67
Non-Hispanic or Latino Other	2,733	2.34	7.58	2.10	3.07	10.76	2.81	2.49	6.90	0.99
% Hispanic or Latino in Segment										
50-100%	1,940	2.16	5.52	1.20	2.47	6.87	1.47	2.27	5.46	0.83
10-<50%	7,001	1.67	4.77	0.89	2.07	6.19	1.56	1.40	4.14	0.57
<10%	20,259	1.90	5.97	1.28	2.32	7.30	1.57	1.16	2.88	0.40

(continued)

Table D.2 2020 NSDUH Respondent QDU-Level Proportions of Extreme Values and Outwinsors (continued)

Domain	*n*	Before res.qdu.nr (SDUWT*DUWT12*DUWT13)			After res.qdu.nr (SDUWT*DUWT12*...*DUWT14)			Final Weight after res.qdu.ps (SDUWT*DUWT12*...*DUWT15)		
		% Unweighted	% Weighted[1]	% Outwinsor[2]	% Unweighted	% Weighted[1]	% Outwinsor[2]	% Unweighted	% Weighted[1]	% Outwinsor[2]
% Black or African American in Segment										
50-100%	1,554	3.86	11.21	2.07	3.99	11.07	2.84	1.48	3.85	0.89
10-<50%	5,296	1.83	4.88	0.97	2.13	5.52	1.37	0.87	2.19	0.35
<10%	22,350	1.73	5.28	1.14	2.18	6.96	1.48	1.38	3.74	0.48
% Owner-Occupied DUs in Segment										
50-100%	21,413	1.48	4.44	0.89	1.93	5.75	1.16	1.15	2.87	0.39
10-<50%	6,097	2.79	8.40	1.79	3.03	9.30	2.35	1.36	3.84	0.52
<10%	1,690	3.37	9.66	2.28	3.79	13.45	3.54	2.78	8.75	1.42
Combined Median Rent/Housing Value										
1st Quintile	4,260	2.54	5.68	1.22	2.98	7.83	1.77	1.64	4.44	0.77
2nd Quintile	6,193	1.68	3.18	0.66	2.15	5.46	1.24	1.50	2.66	0.40
3rd Quintile	6,805	1.54	4.77	1.05	2.00	5.71	1.38	1.09	2.91	0.40
4th Quintile	6,401	1.87	6.47	1.19	2.48	8.11	1.72	1.45	4.64	0.57
5th Quintile	5,541	1.93	8.23	1.80	1.93	8.10	1.75	0.85	2.73	0.35
Population Density										
Large MSA	12,319	2.13	6.77	1.42	2.59	8.40	1.90	1.33	4.11	0.60
Medium to Small MSA	14,975	1.37	3.75	0.78	1.70	4.91	1.06	0.99	2.16	0.29
Non-MSA, Urban	811	2.34	3.99	0.96	1.97	3.77	0.74	1.36	1.70	0.14
Non-MSA, Rural	1,095	5.30	10.51	2.00	6.58	10.77	2.50	4.84	8.45	1.04
Group Quarters										
Group	128	10.94	27.73	5.22	10.16	16.53	2.06	4.69	2.68	0.19
Non-Group	29,072	1.82	5.54	1.16	2.23	6.94	1.56	1.28	3.44	0.48
Household Size										
One	4,141	1.33	3.94	0.63	1.62	4.22	0.75	0.70	1.36	0.18
Two	13,060	1.91	5.62	1.17	2.26	6.93	1.50	1.32	3.55	0.49
Three	6,446	2.06	7.45	1.68	2.34	8.84	2.20	1.30	4.76	0.66
Four or More	5,553	1.91	7.46	1.84	2.68	10.52	2.66	1.66	5.71	0.86

DU = dwelling unit; MSA = metropolitan statistical area; nr = nonresponse adjustment; ps = poststratification adjustment; QDU = questionnaire dwelling unit; res = respondent; SDU = screener dwelling unit.

[1] Weighted extreme value percentage: $100*\sum_k w_{ek}/\sum_k w_k$, where w_{ek} denotes the weight for extreme weights and w_k denotes the weight for extreme weights and nonextreme weights.

[2] Outwinsor weight percentage: $100*\sum_k |w_{ek} - b_k|/\sum_k w_k$, where b_k denotes the cut-off point for defining the extreme weight.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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

This page intentionally left blank

Table E.1 2020 NSDUH Respondent 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	29,200	126,060,944	126,060,944	126,060,944	0.00	0.00
Census Region						
Northeast	5,760	21,700,642	21,700,642	21,700,642	0.00	0.00
South	8,862	48,426,656	48,426,656	48,426,656	0.00	-0.00
Midwest	7,383	27,305,445	27,305,445	27,305,445	0.00	0.00
West	7,195	28,628,201	28,628,201	28,628,201	0.00	0.00
Quarter						
Quarter 1	11,728	64,020,354	64,020,354	64,020,354	0.00	0.00
Quarter 4	17,472	62,040,590	62,040,590	62,040,590	0.00	0.00
Household Type						
12-17, 18-25, 26+	1,952	5,048,192	5,042,838	5,042,838	0.11	0.00
12-17, 18-25	13	28,547	28,583	33,900	-15.79	-15.68
12-17, 26+	5,309	13,431,427	13,436,744	13,431,427	0.00	0.04
18-25, 26+	4,888	13,326,546	13,326,546	13,326,546	0.00	0.00
12-17	6	6,918	6,918	6,918	0.00	0.00
18-25	2,565	5,361,520	5,361,520	5,361,520	0.00	0.00
26+	14,467	88,857,795	88,857,795	88,857,795	0.00	0.00
Race/Ethnicity of Householder						
Hispanic or Latino White	3,399	15,555,346	15,555,346	15,555,346	0.00	0.00
Hispanic or Latino Black or African American	136	880,074	889,099	889,099	-1.02	0.00
Hispanic or Latino Other	311	1,264,296	1,255,270	1,255,270	0.72	0.00
Non-Hispanic or Latino White	20,007	83,491,543	83,491,543	83,491,543	0.00	0.00
Non-Hispanic or Latino Black or African American	2,614	15,146,610	15,146,610	15,146,610	0.00	-0.00
Non-Hispanic or Latino Other	2,733	9,723,075	9,723,075	9,723,075	-0.00	0.00
% Hispanic or Latino in Segment						
50-100%	1,940	10,680,533	10,680,533	10,680,533	0.00	0.00
10-<50%	7,001	33,718,556	33,718,556	33,718,556	0.00	0.00
<10%	20,259	81,661,856	81,661,856	81,661,856	0.00	0.00
% Black or African American in Segment						
50-100%	1,554	9,112,173	9,112,173	9,112,173	-0.00	-0.00
10-<50%	5,296	25,252,217	25,252,217	25,252,217	0.00	0.00
<10%	22,350	91,696,554	91,696,554	91,696,554	0.00	0.00
% Owner-Occupied DUs in Segment						
50-100%	21,413	92,013,748	92,013,749	92,013,748	0.00	0.00
10-<50%	6,097	26,174,988	26,174,988	26,174,988	0.00	0.00
<10%	1,690	7,872,208	7,872,208	7,872,208	0.00	-0.00

(continued)

Table E.1 2020 NSDUH Respondent 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	4,260	18,778,022	18,778,022	18,778,022	0.00	-0.00
2nd Quintile	6,193	26,041,717	26,041,717	26,041,717	0.00	0.00
3rd Quintile	6,805	28,839,108	28,839,108	28,839,108	0.00	0.00
4th Quintile	6,401	27,882,537	27,882,537	27,882,537	0.00	0.00
5th Quintile	5,541	24,519,560	24,519,560	24,519,560	0.00	0.00
Population Density						
Large MSA	12,319	66,671,376	66,671,376	66,671,376	0.00	0.00
Medium to Small MSA	14,975	51,289,886	51,289,886	51,289,886	0.00	0.00
Non-MSA, Urban	811	2,951,604	2,951,604	2,951,604	0.00	0.00
Non-MSA, Rural	1,095	5,148,078	5,148,078	5,148,078	0.00	0.00
Group Quarters						
Group	128	309,032	309,032	309,032	0.00	0.00
Non-Group	29,072	125,751,912	125,751,912	125,751,912	0.00	0.00
Household Size						
One	4,141	32,199,299	32,478,228	32,256,700	-0.18	0.69
Two	13,060	59,292,781	58,833,959	59,050,742	0.41	-0.37
Three	6,446	19,191,247	19,344,396	19,536,453	-1.77	-0.98
Four or More	5,553	15,377,617	15,404,362	15,217,050	1.06	1.23

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).

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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

This page intentionally left blank

Table F.1 2020 NSDUH Selected QDU-Level Weight Summary Statistics

Domain	*n*	SDU-Level Weights (SDUWT: WT1*...*WT11)						Before sel.qdu.ps (SDUWT*DUWT12)						After sel.qdu.ps (SDUWT*DUWT12*DUWT13)					
		Min	Q1[1]	Med	Q3[1]	Max	UWE[2]	Min	Q1[1]	Med	Q3[1]	Max	UWE[2]	Min	Q1[1]	Med	Q3[1]	Max	UWE[2]
Total	44,054	20	871	2,217	4,080	48,748	1.92	20	1,355	3,353	6,946	115,778	2.64	6	1,312	3,319	7,074	123,785	2.62
Census Region																			
Northeast	9,219	20	698	1,726	3,526	29,416	1.97	20	1,034	2,691	5,635	115,778	2.82	17	985	2,574	5,576	72,584	2.95
South	13,057	22	1,443	3,140	5,192	48,748	1.73	31	2,229	4,801	9,340	104,364	2.26	6	2,170	4,802	9,581	123,785	2.25
Midwest	10,970	24	940	2,064	3,388	24,480	1.66	24	1,428	3,104	5,829	111,049	2.56	18	1,380	3,069	5,942	68,880	2.50
West	10,808	33	609	1,701	4,231	27,837	2.14	35	937	2,665	6,499	110,359	3.03	22	910	2,684	6,589	94,047	2.89
Quarter																			
Quarter 1	17,175	65	1,830	3,440	5,049	48,748	1.57	65	2,547	5,006	9,284	110,359	2.13	29	2,423	4,888	9,499	123,785	2.18
Quarter 4	26,879	20	681	1,547	3,087	32,886	2.21	20	1,025	2,453	5,350	115,778	3.11	6	987	2,446	5,519	106,688	2.96
Household Type																			
12-17, 18-25, 26+	3,202	35	995	2,372	4,334	25,736	1.91	35	995	2,372	4,334	25,736	1.91	17	923	2,243	4,240	31,638	2.07
12-17, 18-25	25	73	484	2,415	3,708	12,210	1.88	73	484	2,415	3,708	12,210	1.88	73	478	2,216	3,361	16,692	2.47
12-17, 26+	9,571	24	799	2,096	3,909	32,886	1.93	24	803	2,106	3,936	32,895	1.92	18	777	2,032	3,937	31,842	1.97
18-25, 26+	7,313	20	1,003	2,460	4,423	32,582	1.88	20	1,137	2,763	5,030	32,720	1.88	6	1,082	2,639	4,998	35,015	1.95
12-17	6	377	1,407	2,392	3,375	3,894	1.26	380	1,421	2,405	3,409	3,933	1.26	229	1,710	2,464	3,110	3,858	1.26
18-25	3,290	28	707	1,912	4,028	35,147	2.10	28	799	2,134	4,652	41,026	2.14	23	717	2,069	4,561	59,501	2.35
26+	20,647	21	888	2,201	4,027	48,748	1.89	29	2,280	5,586	11,305	115,778	2.21	25	2,322	5,618	11,701	123,785	2.15
Race/Ethnicity of Householder																			
Hispanic or Latino White	5,231	46	1,640	3,139	5,269	32,886	1.69	65	1,975	3,931	7,122	107,785	2.28	29	1,951	4,068	7,365	85,233	2.24
Hispanic or Latino Black or African American	217	61	1,136	3,407	6,831	29,416	2.14	63	1,311	4,655	10,381	91,042	2.72	19	1,178	4,427	11,199	72,540	2.69
Hispanic or Latino Other	481	31	631	1,632	3,994	48,748	3.09	31	805	2,293	5,662	95,516	4.40	6	690	2,068	6,298	77,986	3.49
Non-Hispanic or Latino White	30,091	21	778	1,957	3,702	30,736	1.89	24	1,259	3,171	6,704	110,359	2.69	18	1,214	3,113	6,739	123,785	2.68
Non-Hispanic or Latino Black or African American	3,948	51	1,738	3,542	5,636	29,868	1.62	51	2,401	5,002	9,363	115,778	2.32	22	2,327	5,033	9,762	86,564	2.22
Non-Hispanic or Latino Other	4,086	20	693	1,724	3,832	35,147	2.18	20	918	2,508	5,735	68,223	2.78	17	859	2,450	5,978	94,047	3.00
% Hispanic or Latino in Segment																			
50-100%	2,996	67	2,157	3,903	5,935	35,147	1.55	68	2,701	5,176	8,408	80,939	2.00	19	2,694	5,075	8,828	59,501	1.94
10-<50%	10,498	22	1,253	2,722	4,946	32,582	1.77	31	1,820	3,994	8,127	94,248	2.31	6	1,717	3,951	8,261	94,047	2.37
<10%	30,560	20	729	1,882	3,639	48,748	1.99	20	1,146	2,990	6,299	115,778	2.87	17	1,108	2,952	6,428	123,785	2.81

(continued)

Table F.1 2020 NSDUH Selected QDU-Level Weight Summary Statistics (continued)

Domain	*n*	SDU-Level Weights (SDUWT: WT1*...*WT11)						Before sel.qdu.ps (SDUWT*DUWT12)						After sel.qdu.ps (SDUWT*DUWT12*DUWT13)					
		Min	Q1[1]	Med	Q3[1]	Max	UWE[2]	Min	Q1[1]	Med	Q3[1]	Max	UWE[2]	Min	Q1[1]	Med	Q3[1]	Max	UWE[2]
% Black or African American in Segment																			
50-100%	2,352	35	1,642	3,403	5,638	29,868	1.66	35	2,304	4,967	9,422	115,778	2.47	48	2,113	4,974	9,821	86,564	2.30
10-<50%	7,740	22	1,363	2,890	4,921	29,807	1.72	31	2,007	4,303	8,115	68,823	2.16	6	1,914	4,232	8,401	77,986	2.22
<10%	33,962	20	772	1,976	3,791	48,748	1.98	20	1,205	3,058	6,478	110,359	2.78	17	1,161	3,027	6,555	123,785	2.75
% Owner-Occupied DUs in Segment																			
50-100%	32,916	24	858	2,135	3,895	48,748	1.89	24	1,348	3,285	6,783	115,778	2.62	17	1,306	3,231	6,871	123,785	2.60
10-<50%	8,804	20	843	2,381	4,527	29,416	1.90	20	1,314	3,446	7,124	104,364	2.65	19	1,250	3,443	7,410	106,688	2.72
<10%	2,334	30	1,168	2,932	5,434	35,147	1.95	31	1,676	4,115	8,507	111,049	2.69	6	1,688	4,143	8,716	70,962	2.37
Combined Median Rent/Housing Value																			
1st Quintile	6,382	30	958	2,379	4,214	25,362	1.82	35	1,436	3,482	7,193	115,778	2.70	17	1,438	3,560	7,335	94,047	2.53
2nd Quintile	9,169	31	785	2,244	4,096	32,886	1.88	31	1,229	3,363	7,044	104,364	2.65	6	1,155	3,251	6,999	86,564	2.71
3rd Quintile	10,265	38	847	2,187	4,017	32,582	1.93	47	1,323	3,334	6,871	104,667	2.58	22	1,262	3,346	7,096	77,899	2.54
4th Quintile	9,734	20	919	2,214	4,164	35,147	1.94	20	1,423	3,370	6,976	94,248	2.55	18	1,374	3,316	7,164	78,778	2.56
5th Quintile	8,504	28	890	2,084	3,950	48,748	2.00	28	1,390	3,242	6,789	110,359	2.77	18	1,350	3,187	6,865	123,785	2.72
Population Density																			
Large MSA	18,808	22	1,563	3,044	4,910	48,748	1.69	35	2,283	4,538	8,477	115,778	2.31	19	2,190	4,492	8,823	123,785	2.29
Medium to Small MSA	22,398	20	582	1,518	3,306	28,202	2.07	20	872	2,438	5,493	110,359	2.89	6	845	2,415	5,567	106,688	2.88
Non-MSA, Urban	1,229	35	621	1,519	3,464	17,935	2.01	35	984	2,673	5,885	45,937	2.61	22	901	2,704	5,886	46,843	2.60
Non-MSA, Rural	1,619	53	717	1,969	3,782	32,886	2.07	53	1,220	3,367	7,475	104,667	3.10	65	1,369	3,610	7,605	94,047	2.77
Group Quarters																			
Group	136	122	842	1,991	4,295	10,204	1.78	150	1,146	2,619	5,093	46,515	2.93	83	940	2,183	5,535	64,943	3.31
Non-Group	43,918	20	871	2,217	4,079	48,748	1.92	20	1,356	3,354	6,952	115,778	2.64	6	1,313	3,321	7,078	123,785	2.62
Household Size																			
One	5,358	26	840	2,228	4,007	28,561	1.81	28	2,620	7,183	16,103	115,778	2.23	23	2,646	7,345	16,858	123,785	2.16
Two	18,854	20	843	2,147	3,979	48,748	1.92	20	1,652	4,080	8,330	88,156	2.20	23	1,645	4,100	8,586	77,986	2.16
Three	10,647	22	889	2,241	4,101	32,886	1.93	24	1,041	2,563	4,910	45,141	2.07	18	991	2,465	4,896	48,973	2.16
Four or More	9,195	28	944	2,362	4,324	32,582	1.94	28	982	2,432	4,504	37,482	1.97	6	906	2,310	4,339	35,015	2.12

DU = dwelling unit; MSA = metropolitan statistical area; ps = poststratification adjustment; QDU = questionnaire dwelling unit; SDU = screener dwelling unit; sel = selected; UWE = unequal weighting effect.

[1] Q1 and Q3 refer to the first and third quartiles of the weight distribution.

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

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

Table F.2 2020 NSDUH Respondent QDU-Level Weight Summary Statistics

Domain	*n*	Before res.qdu.nr (SDUWT*DUWT12*DUWT13)						After res.qdu.nr (SDUWT*DUWT12*...*DUWT14)						Final Weight after res.qdu.ps (SDUWT*DUWT12*...*DUWT15)					
		Min	Q1[1]	Med	Q3[1]	Max	UWE[2]	Min	Q1[1]	Med	Q3[1]	Max	UWE[2]	Min	Q1[1]	Med	Q3[1]	Max	UWE[2]
Total	29,200	18	1,396	3,470	7,304	123,785	2.59	18	1,963	4,891	10,789	168,916	2.59	10	998	2,506	5,413	79,994	2.52
Census Region																			
Northeast	5,760	19	1,007	2,645	5,740	72,584	2.98	19	1,521	3,867	8,786	109,331	2.95	10	791	1,969	4,524	53,765	2.80
South	8,862	18	2,336	5,006	9,789	123,785	2.23	18	3,197	7,151	14,325	168,916	2.21	14	1,633	3,660	7,182	54,651	2.14
Midwest	7,383	18	1,433	3,213	6,184	61,192	2.43	36	1,956	4,452	8,953	90,573	2.52	13	999	2,297	4,588	40,210	2.41
West	7,195	27	946	2,807	6,706	94,047	2.89	29	1,381	4,038	9,801	138,996	2.90	17	692	2,016	4,894	79,994	2.92
Quarter																			
Quarter 1	11,728	29	2,378	4,773	9,234	123,785	2.22	29	3,110	6,670	13,933	168,916	2.32	13	1,615	3,486	7,023	54,651	2.20
Quarter 4	17,472	18	1,066	2,644	5,951	106,688	2.90	18	1,557	3,863	8,788	138,996	2.72	10	791	1,942	4,367	79,994	2.73
Household Type																			
12-17, 18-25, 26+	1,952	19	952	2,443	4,420	31,638	2.05	19	1,538	3,586	6,718	46,762	2.14	10	777	1,854	3,455	17,884	1.98
12-17, 18-25	13	224	1,715	2,255	4,004	16,692	2.42	224	1,851	2,255	4,225	18,605	2.24	99	859	993	2,892	8,568	2.13
12-17, 26+	5,309	18	826	2,225	4,152	31,842	1.91	31	1,505	3,742	6,815	51,165	2.00	12	783	1,926	3,525	19,582	1.86
18-25, 26+	4,888	18	1,117	2,674	4,980	35,015	1.95	18	1,497	3,708	7,348	81,094	2.11	13	772	1,901	3,821	40,851	2.04
12-17	6	229	1,710	2,464	3,110	3,858	1.26	229	1,710	2,464	3,110	3,858	1.26	100	806	1,230	1,607	1,944	1.27
18-25	2,565	23	702	1,977	4,516	59,501	2.34	26	892	2,400	5,507	63,915	2.61	17	457	1,264	2,914	25,177	2.32
26+	14,467	31	2,316	5,581	11,688	123,785	2.17	39	3,067	7,741	16,555	168,916	2.21	21	1,530	3,880	8,471	79,994	2.17
Race/Ethnicity of Householder																			
Hispanic or Latino White	3,399	29	2,047	4,170	7,385	71,652	2.22	81	2,752	6,124	11,437	143,596	2.34	32	1,415	3,122	5,708	51,273	2.27
Hispanic or Latino Black or African American	136	19	1,473	4,642	11,529	72,540	2.68	19	1,784	7,556	18,769	109,331	2.48	10	1,059	4,476	9,944	39,010	2.16
Hispanic or Latino Other	311	23	690	2,061	7,323	52,407	3.26	32	1,302	3,503	10,088	75,299	3.03	12	595	1,641	5,183	32,491	2.90
Non-Hispanic or Latino White	20,007	18	1,306	3,280	7,026	123,785	2.65	24	1,859	4,647	10,283	168,916	2.61	13	943	2,359	5,157	50,639	2.56
Non-Hispanic or Latino Black or African American	2,614	22	2,489	5,271	9,898	86,564	2.23	41	3,465	7,385	14,907	124,222	2.23	27	1,815	3,885	7,466	53,765	2.13
Non-Hispanic or Latino Other	2,733	18	897	2,564	6,106	94,047	2.92	18	1,268	3,504	8,707	139,426	3.15	13	660	1,791	4,495	79,994	2.98
% Hispanic or Latino in Segment																			
50-100%	1,940	19	2,808	5,164	9,082	59,501	1.93	19	3,889	7,452	14,526	84,771	1.95	10	1,947	3,764	7,239	47,380	1.94
10-<50%	7,001	22	1,791	4,063	8,403	94,047	2.37	41	2,424	5,742	12,374	143,596	2.44	17	1,235	2,950	6,179	79,994	2.40
<10%	20,259	18	1,173	3,115	6,703	123,785	2.77	18	1,704	4,416	9,834	168,916	2.73	12	867	2,261	4,979	53,765	2.64

(continued)

Table F.2 2020 NSDUH Respondent QDU-Level Weight Summary Statistics (continued)

Domain	*n*	Before res.qdu.nr (SDUWT*DUWT12*DUWT13)						After res.qdu.nr (SDUWT*DUWT12*...*DUWT14)						Final Weight: After res.qdu.ps (SDUWT*DUWT12*...*DUWT15)					
		Min	Q1[1]	Med	Q3[1]	Max	UWE[2]	Min	Q1[1]	Med	Q3[1]	Max	UWE[2]	Min	Q1[1]	Med	Q3[1]	Max	UWE[2]
% Black or African American in Segment																			
50-100%	1,554	59	2,070	5,063	10,127	86,564	2.29	59	2,910	7,264	15,163	124,222	2.29	32	1,589	3,896	7,594	53,765	2.15
10-<50%	5,296	18	2,076	4,478	8,645	71,652	2.18	18	2,820	6,211	12,465	95,327	2.18	10	1,427	3,138	6,189	50,582	2.17
<10%	22,350	18	1,233	3,160	6,756	123,785	2.73	24	1,763	4,510	10,035	168,916	2.72	13	896	2,299	5,076	79,994	2.65
% Owner-Occupied DUs in Segment																			
50-100%	21,413	18	1,410	3,438	7,191	123,785	2.56	18	2,040	4,934	10,765	168,916	2.53	14	1,031	2,525	5,377	54,651	2.47
10-<50%	6,097	19	1,280	3,447	7,357	106,688	2.77	19	1,684	4,600	10,704	138,996	2.78	10	874	2,389	5,377	79,994	2.72
<10%	1,690	59	1,646	4,194	8,458	70,962	2.37	59	1,925	5,369	11,709	143,596	2.60	32	988	2,811	6,002	50,327	2.47
Combined Median Rent/Housing Value																			
1st Quintile	4,260	19	1,592	3,731	7,722	94,047	2.50	19	2,235	5,149	11,277	138,996	2.49	10	1,144	2,650	5,514	79,994	2.49
2nd Quintile	6,193	27	1,231	3,418	7,243	86,564	2.67	59	1,729	4,819	10,468	99,213	2.65	18	885	2,472	5,270	51,273	2.60
3rd Quintile	6,805	22	1,338	3,465	7,267	77,899	2.53	29	1,856	4,816	10,752	143,596	2.57	17	940	2,455	5,344	50,639	2.53
4th Quintile	6,401	18	1,447	3,393	7,212	78,778	2.56	24	2,042	4,795	10,811	101,432	2.61	12	1,043	2,457	5,419	54,651	2.52
5th Quintile	5,541	18	1,435	3,376	7,264	123,785	2.68	18	2,043	4,978	10,870	168,916	2.61	14	1,056	2,569	5,556	50,327	2.45
Population Density																			
Large MSA	12,319	19	2,297	4,634	9,076	123,785	2.29	19	3,290	6,750	13,678	168,916	2.29	10	1,671	3,485	6,837	54,651	2.21
Medium to Small MSA	14,975	18	902	2,574	5,761	106,688	2.83	18	1,296	3,586	8,318	109,331	2.81	12	658	1,820	4,234	47,916	2.75
Non-MSA, Urban	811	33	898	2,829	6,163	46,843	2.63	33	1,390	4,032	8,993	60,118	2.55	13	711	1,939	4,491	29,052	2.56
Non-MSA, Rural	1,095	78	1,433	3,980	8,228	94,047	2.60	79	2,183	5,345	12,235	138,996	2.57	34	1,161	2,786	5,899	79,994	2.61
Group Quarters																			
Group	128	83	940	2,213	5,535	64,943	3.32	211	1,063	2,226	5,800	67,834	3.27	97	506	1,160	3,021	36,015	3.40
Non-Group	29,072	18	1,398	3,473	7,311	123,785	2.59	18	1,971	4,905	10,821	168,916	2.59	10	1,004	2,510	5,425	79,994	2.52
Household Size																			
One	4,141	23	2,491	6,957	16,002	123,785	2.22	26	3,063	9,058	21,425	168,916	2.29	17	1,564	4,569	10,836	79,994	2.25
Two	13,060	24	1,649	4,077	8,451	72,584	2.15	30	2,171	5,455	12,302	94,268	2.27	17	1,096	2,797	6,152	40,851	2.19
Three	6,446	18	1,047	2,600	5,015	48,973	2.13	24	1,653	3,967	7,681	92,010	2.26	13	847	2,032	3,972	45,466	2.19
Four or More	5,553	18	944	2,482	4,461	35,015	2.08	18	1,510	3,722	7,198	81,094	2.23	10	784	1,921	3,671	28,544	2.12

DU = dwelling unit; MSA = metropolitan statistical area; nr = nonresponse adjustment; ps = poststratification adjustment; QDU = questionnaire dwelling unit; res = respondent; SDU = screener dwelling unit; UWE = unequal weighting effect.

[1] Q1 and Q3 refer to the first and third quartiles of the weight distribution.

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

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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

This page intentionally left blank

Table G.1 2020 NSDUH Person Pair-Level Response Rates

Domain	Selected Pairs	Respondent Pairs	Person Pair-Level Response Rate[1]
Total	18,461	7,084	38.03
Pair Age Group			
12-17, 12-17	2,165	913	47.66
12-17, 18-25	1,743	609	40.40
12-17, 26-34	712	253	39.31
12-17, 35-49	3,602	1,375	42.30
12-17, 50+	730	274	40.28
18-25, 18-25	2,885	1,162	42.67
18-25, 26-34	960	359	36.58
18-25, 35-49	1,163	457	45.69
18-25, 50+	890	324	33.80
26-34, 26-34	1,010	421	35.39
26-34, 35-49	522	208	37.73
26-34, 50+	300	86	29.16
35-49, 35-49	879	324	36.62
35-49, 50+	299	98	33.65
50+, 50+	601	221	36.57
Pair Race/Ethnicity			
Hispanic or Latino	2,990	1,088	34.18
Black or African American	1,507	629	37.05
White	10,222	3,911	40.17
Other	1,627	632	34.14
White & Black or African American	222	71	27.12
White & Hispanic or Latino	801	309	36.40
White & Other	735	292	44.09
Black or African American & Hispanic or Latino	110	52	36.64
Black or African American & Other	111	46	41.96
Hispanic or Latino & Other	136	54	54.57
Pair Gender			
Male, Male	3,741	1,349	34.47
Female, Female	4,037	1,607	40.66
Male, Female	10,683	4,128	38.27
Household Size			
Two	4,630	1,923	42.66
Three	5,540	2,031	35.76
Four or More	8,291	3,130	36.89

(continued)

Table G.1 2020 NSDUH Person Pair-Level Response Rates (continued)

Domain	Selected Pairs	Respondent Pairs	Person Pair-Level Response Rate[1]
Census Region			
Northeast	3,857	1,431	37.21
South	5,328	2,130	39.79
Midwest	4,499	1,720	39.15
West	4,777	1,803	35.23
Quarter			
Quarter 1	7,129	3,900	48.62
Quarter 4	11,332	3,184	28.09
% Hispanic or Latino in Segment			
50-100%	1,622	585	31.17
10-<50%	4,561	1,710	34.77
<10%	12,278	4,789	41.24
% Black or African American in Segment			
50-100%	962	387	34.96
10-<50%	3,275	1,245	38.02
<10%	14,224	5,452	38.33
% Owner-Occupied DUs in Segment			
50-100%	13,846	5,322	39.11
10-<50%	3,642	1,373	33.25
<10%	973	389	39.45
Combined Median Rent/Housing Value			
1st Quintile	2,759	1,120	38.87
2nd Quintile	3,872	1,547	35.38
3rd Quintile	4,298	1,640	36.01
4th Quintile	4,113	1,518	39.03
5th Quintile	3,419	1,259	41.71
Population Density			
Large MSA	7,836	2,939	38.47
Medium to Small MSA	9,451	3,641	37.35
Non-MSA, Urban	511	208	37.07
Non-MSA, Rural	663	296	39.21
Group Quarters			
Group	64	46	79.39
Non-Group	18,397	7,038	37.99

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 the selected pair poststratification adjustment. This weight is the product of WT1*…*WT11*PRWT12*PRWT13.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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

This page intentionally left blank

Table H.1 2020 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	18,461	1.85	7.32	1.84	4.06	22.31	10.06	3.18	12.35	2.98
Pair Age Group										
12-17, 12-17	2,165	1.11	4.45	1.44	2.77	14.39	4.90	1.76	10.40	3.43
12-17, 18-25	1,743	1.32	5.15	1.28	6.71	31.69	13.08	2.75	11.67	1.92
12-17, 26-34	712	1.69	6.12	1.52	1.54	8.11	2.87	1.97	7.42	2.79
12-17, 35-49	3,602	1.42	5.93	1.55	2.19	11.27	2.92	2.19	7.49	1.76
12-17, 50+	730	2.05	7.09	1.37	2.19	9.75	2.75	1.92	7.84	1.38
18-25, 18-25	2,885	1.77	6.38	1.54	6.07	26.81	11.69	4.26	16.64	3.84
18-25, 26-34	960	4.48	15.03	2.76	5.10	24.96	10.49	4.17	13.67	3.21
18-25, 35-49	1,163	3.27	11.46	2.83	6.28	24.15	7.76	6.71	19.32	4.59
18-25, 50+	890	1.57	5.98	1.26	4.61	19.44	6.18	2.25	9.50	2.93
26-34, 26-34	1,010	1.98	9.73	3.11	2.08	11.61	4.04	3.07	16.54	3.80
26-34, 35-49	522	2.87	13.34	3.65	4.21	23.56	8.63	3.83	11.64	2.47
26-34, 50+	300	3.33	10.73	2.20	6.00	24.11	8.32	3.33	13.57	2.00
35-49, 35-49	879	1.02	4.31	1.00	2.84	22.69	11.56	4.78	12.41	2.73
35-49, 50+	299	2.34	8.47	3.05	5.69	39.63	22.47	5.02	20.78	5.19
50+, 50+	601	1.66	8.55	2.40	4.33	28.87	17.58	2.50	10.54	3.00
Pair Race/Ethnicity										
Hispanic or Latino	2,990	2.61	9.56	2.10	4.78	23.71	10.52	3.48	14.71	3.57
Black or African American	1,507	4.58	12.05	3.25	6.77	39.35	21.12	8.10	20.33	5.37
White	10,222	1.00	3.85	0.92	3.38	18.47	7.95	2.04	9.30	2.15
Other	1,627	2.27	8.85	2.24	4.43	16.17	6.36	4.00	12.69	3.40
White & Black or African American	222	3.15	11.83	3.53	5.86	23.08	10.38	7.66	20.36	4.82
White & Hispanic or Latino	801	2.50	11.98	3.60	3.12	12.04	4.52	3.50	11.59	2.21
White & Other	735	1.77	9.77	2.76	3.13	17.79	6.20	3.40	13.48	2.57
Black or African American & Hispanic or Latino	110	4.55	15.62	4.94	9.09	49.10	13.41	12.73	24.64	7.18
Black or African American & Other	111	6.31	30.11	8.10	5.41	27.33	9.04	0.90	4.54	1.79
Hispanic or Latino & Other	136	2.94	8.00	1.93	8.09	49.53	27.23	1.47	11.45	3.38
Pair Gender										
Male, Male	3,741	2.22	8.51	2.15	5.32	18.99	6.07	4.09	12.33	2.86
Female, Female	4,037	1.78	7.22	1.94	3.89	21.36	10.12	3.20	11.17	2.29
Male, Female	10,683	1.75	6.92	1.68	3.69	23.42	11.07	2.86	12.68	3.21
Household Size										
Two	4,630	1.66	7.08	1.71	1.19	6.63	1.69	1.68	4.88	1.11
Three	5,540	2.04	8.02	2.34	2.56	24.06	11.90	2.76	12.28	2.90
Four or More	8,291	1.83	6.99	1.57	6.67	29.13	13.28	4.29	15.97	3.92

(continued)

Table H.1 2020 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	3,857	1.92	8.39	2.62	4.20	26.01	13.11	4.72	15.81	3.99
South	5,328	1.60	6.17	1.38	4.13	23.23	11.20	2.91	12.20	2.71
Midwest	4,499	1.67	6.73	1.65	4.05	18.66	7.64	2.82	12.48	3.84
West	4,777	2.26	8.87	2.16	3.89	21.00	8.01	2.57	10.17	2.06
Quarter										
Quarter 1	7,129	2.75	7.85	2.17	5.09	22.21	10.22	5.04	14.54	3.70
Quarter 4	11,332	1.29	6.81	1.50	3.42	22.39	9.92	2.01	10.30	2.30
% Hispanic or Latino in Segment										
50-100%	1,622	2.10	6.72	1.31	4.99	30.25	17.25	4.07	14.18	3.80
10-<50%	4,561	1.95	6.81	1.54	3.84	22.34	10.01	3.20	11.52	2.26
<10%	12,278	1.78	7.71	2.10	4.02	20.28	8.27	3.05	12.31	3.13
% Black or African American in Segment										
50-100%	962	4.26	13.59	3.95	7.17	26.60	11.84	8.42	20.70	5.92
10-<50%	3,275	2.11	7.68	1.99	4.49	29.10	15.73	3.60	12.13	2.73
<10%	14,224	1.63	6.55	1.56	3.75	19.87	8.20	2.73	11.60	2.76
% Owner-Occupied DUs[1] in Segment										
50-100%	13,846	1.58	6.38	1.54	3.87	21.75	10.04	2.90	11.73	2.95
10-<50%	3,642	2.64	9.21	2.59	4.59	26.15	11.04	4.86	15.73	3.34
<10%	973	2.77	11.44	2.61	4.83	16.11	6.48	0.92	5.48	0.96
Combined Median Rent/Housing Value										
1st Quintile	2,759	2.28	7.69	1.95	5.18	23.76	12.25	3.59	10.50	2.12
2nd Quintile	3,872	1.70	5.88	1.35	4.39	26.85	12.40	3.64	15.81	3.76
3rd Quintile	4,298	1.79	7.71	1.77	3.86	19.20	7.08	2.70	11.99	3.31
4th Quintile	4,113	1.97	7.87	2.23	3.99	21.68	9.74	3.57	13.23	3.18
5th Quintile	3,419	1.61	7.47	1.89	3.13	20.31	9.64	2.46	9.35	2.14
Population Density										
Large MSA[1]	7,836	2.25	8.43	2.13	4.38	25.91	12.73	3.52	12.44	2.89
Medium to Small MSA[1]	9,451	1.33	5.45	1.28	3.39	15.62	5.42	2.74	11.18	2.69
Non-MSA,[1] Urban	511	2.54	5.38	0.90	5.87	26.42	14.79	6.07	17.22	3.82
Non-MSA,[1] Rural	663	4.07	13.02	4.34	8.60	32.89	14.21	3.17	20.63	6.97
Group Quarters										
Group	64	3.13	7.76	1.25	14.06	37.36	17.82	7.81	33.69	10.29
Non-Group	18,397	1.85	7.32	1.84	4.03	22.29	10.06	3.16	12.33	2.97

DU = dwelling unit; MSA = metropolitan statistical area; pr = pair; ps = poststratification adjustment; SDU = screener dwelling unit; sel = selected.

[1] This step used demographic variables from screener data for all selected person pairs.

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

[3] Outwinsor weight percentage: $100*\sum_k |w_{ek} - b_k|/\sum_k w_k$, where b_k denotes the cut-off point for defining the extreme weight.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

Table H.2 2020 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	7,084	3.39	12.42	3.04	7.37	29.46	11.26
Pair Age Group							
12-17, 12-17	913	2.08	10.82	2.93	4.82	25.34	9.49
12-17, 18-25	609	2.96	12.68	2.19	8.70	42.71	18.62
12-17, 26-34	253	2.37	11.97	6.10	5.93	33.61	11.92
12-17, 35-49	1,375	2.25	6.86	1.74	6.98	34.91	13.02
12-17, 50+	274	1.82	7.49	0.86	2.55	8.27	3.24
18-25, 18-25	1,162	4.48	16.40	3.61	8.26	41.22	16.50
18-25, 26-34	359	3.62	14.31	2.75	10.31	48.23	23.98
18-25, 35-49	457	6.56	18.51	3.91	9.85	38.42	17.89
18-25, 50+	324	3.70	13.40	2.83	8.64	28.11	9.94
26-34, 26-34	421	3.56	19.39	5.59	7.13	42.94	19.20
26-34, 35-49	208	5.29	14.71	2.75	7.69	27.83	6.57
26-34, 50+	86	2.33	9.37	1.45	3.49	14.79	3.05
35-49, 35-49	324	5.25	14.80	4.83	9.57	30.75	11.27
35-49, 50+	98	5.10	22.88	6.21	4.08	15.79	4.32
50+, 50+	221	1.81	7.49	2.01	7.69	27.65	10.39
Pair Race/Ethnicity							
Hispanic or Latino	1,088	4.32	16.83	3.93	10.85	39.77	16.98
Black or African American	629	8.43	19.29	5.22	11.61	34.05	14.08
White	3,911	1.97	8.35	2.08	5.45	23.86	8.03
Other	632	3.48	18.11	4.47	6.80	34.01	12.55
White & Black or African American	71	11.27	24.60	9.94	22.54	51.99	25.27
White & Hispanic or Latino	309	3.56	11.71	1.75	9.71	26.40	11.11
White & Other	292	4.11	10.59	2.10	4.11	23.48	9.10
Black or African American & Hispanic or Latino	52	13.46	31.37	10.30	23.08	46.83	19.34
Black or African American & Other	46	2.17	10.82	3.58	6.52	18.58	3.96
Hispanic or Latino & Other	54	3.70	20.98	1.48	3.70	9.49	2.04
Pair Gender							
Male, Male	1,349	4.60	14.30	3.08	10.90	32.14	14.74
Female, Female	1,607	3.61	10.97	2.11	6.78	26.85	10.70
Male, Female	4,128	2.91	12.39	3.30	6.44	29.46	10.46
Household Size							
Two	1,923	2.13	6.40	1.67	3.38	18.37	7.50
Three	2,031	3.50	15.11	3.75	7.63	35.58	13.84
Four or More	3,130	4.09	14.52	3.47	9.65	31.87	11.83

(continued)

Table H.2 2020 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	1,431	4.33	13.50	3.62	7.69	23.71	9.83
South	2,130	3.19	10.52	2.44	7.98	28.83	11.09
Midwest	1,720	3.20	10.32	3.20	6.69	27.91	9.85
West	1,803	3.05	16.47	3.47	7.04	35.31	13.50
Quarter							
Quarter 1	3,900	4.64	14.41	3.69	5.56	25.01	9.53
Quarter 4	3,184	1.85	9.19	1.97	9.58	33.63	12.88
% Hispanic or Latino in Segment							
50-100%	585	4.62	15.00	2.87	12.48	38.02	13.75
10-<50%	1,710	3.45	10.12	2.28	7.43	27.24	10.63
<10%	4,789	3.22	12.89	3.38	6.72	28.49	10.97
% Black or African American in Segment							
50-100%	387	10.08	22.03	6.38	13.70	39.04	16.81
10-<50%	1,245	4.58	12.36	2.53	9.16	30.49	13.11
<10%	5,452	2.64	11.59	2.88	6.51	28.24	10.21
% Owner-Occupied DUs[1] in Segment							
50-100%	5,322	3.03	11.55	2.89	6.91	28.93	11.16
10-<50%	1,373	5.39	17.45	3.98	10.78	34.13	12.56
<10%	389	1.29	7.57	1.43	1.54	7.83	3.70
Combined Median Rent/Housing Value							
1st Quintile	1,120	3.84	7.76	1.57	6.79	24.59	7.92
2nd Quintile	1,547	3.10	13.34	3.28	10.08	34.91	14.24
3rd Quintile	1,640	2.99	12.28	3.10	6.89	32.30	12.70
4th Quintile	1,518	3.95	10.75	2.53	6.26	24.93	8.96
5th Quintile	1,259	3.18	16.96	4.38	6.51	28.93	11.45
Population Density							
Large MSA[1]	2,939	4.32	13.97	3.47	7.93	29.10	11.16
Medium to Small MSA[1]	3,641	2.58	10.09	2.17	7.03	28.97	10.75
Non-MSA,[1] Urban	208	6.73	7.55	2.67	7.69	27.56	14.16
Non-MSA,[1] Rural	296	1.69	16.03	5.68	5.74	41.21	16.58
Group Quarters							
Group	46	8.70	40.63	14.61	8.70	50.29	12.04
Non-Group	7,038	3.35	12.37	3.01	7.36	29.44	11.26

DU = dwelling unit; MSA = metropolitan statistical area; nr = nonresponse adjustment; pr = pair; res = respondent; SDU = screener dwelling unit.

[1] This step used demographic variables from screener data for all responding person pairs.

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

[3] Outwinsor weight percentage: $100*\sum_k |w_{ek} - b_k|/\sum_k w_k$, where b_k denotes the cut-off point for defining the extreme weight.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

Table H.3 2020 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	7,084	7.13	27.94	10.26	4.59	14.89	3.86	3.39	9.58	1.12
Pair Age Group										
12-17, 12-17	917	4.91	25.30	9.10	1.85	13.71	2.28	0.98	7.01	0.87
12-17, 18-25	599	8.51	42.28	18.24	6.18	24.95	6.25	3.67	14.83	1.73
12-17, 26-34	253	5.93	29.34	11.04	3.95	15.51	2.50	1.98	9.63	1.38
12-17, 35-49	1,377	6.61	33.61	12.16	3.34	14.53	4.06	2.61	8.68	0.94
12-17, 50+	269	1.49	6.28	1.40	1.12	4.30	0.56	0.74	2.11	0.19
18-25, 18-25	1,151	8.25	39.53	15.95	6.17	24.35	7.14	5.13	15.95	1.75
18-25, 26-34	363	9.64	46.33	22.83	7.44	21.66	6.09	4.96	15.99	2.00
18-25, 35-49	455	10.77	39.93	17.09	8.13	27.69	6.50	6.15	18.97	2.60
18-25, 50+	326	7.36	24.70	8.40	3.37	12.22	2.12	2.15	6.06	0.66
26-34, 26-34	436	6.88	39.18	17.92	5.73	26.05	9.20	5.50	22.28	2.22
26-34, 35-49	206	9.71	36.43	10.63	4.85	24.03	6.03	3.88	11.11	1.86
26-34, 50+	85	1.18	6.19	1.47	1.18	5.97	1.82	1.18	4.93	0.60
35-49, 35-49	324	8.64	23.46	10.71	6.79	13.99	3.96	4.63	8.75	0.94
35-49, 50+	101	1.98	9.80	2.92	0.99	3.41	0.11	0.00	0.00	0.00
50+, 50+	222	6.76	30.77	8.58	3.15	12.57	3.51	2.70	9.33	1.05
Pair Race/Ethnicity										
Hispanic or Latino	1,089	9.37	35.74	16.25	5.88	24.16	6.17	4.59	15.88	1.92
Black or African American	611	11.46	34.84	12.48	8.84	20.97	5.91	9.17	19.58	2.47
White	3,819	5.39	22.35	7.02	3.30	9.40	2.49	1.96	5.50	0.54
Other	600	6.83	31.74	10.13	4.00	11.14	2.64	2.67	7.75	0.83
White & Black or African American	84	14.29	48.21	26.38	19.05	26.93	10.28	19.05	18.33	2.42
White & Hispanic or Latino	330	8.48	21.01	8.50	6.67	21.88	4.43	3.94	6.64	1.00
White & Other	348	6.61	33.57	11.72	2.30	14.10	1.99	2.30	5.01	0.31
Black or African American & Hispanic or Latino	63	17.46	34.86	16.32	9.52	35.97	11.61	6.35	24.09	5.74
Black or African American & Other	75	14.67	44.38	15.00	4.00	10.47	2.61	1.33	0.74	0.14
Hispanic or Latino & Other	65	1.54	8.16	1.79	3.08	3.34	1.87	1.54	0.20	0.01
Pair Gender										
Male, Male	1,351	10.36	30.02	13.24	7.55	17.68	5.47	6.81	15.10	1.66
Female, Female	1,616	6.50	24.15	9.53	4.15	11.08	3.13	3.28	6.45	0.66
Male, Female	4,117	6.32	28.46	9.67	3.79	15.19	3.62	2.31	8.94	1.10
Household Size										
Two	1,923	3.33	20.89	7.83	2.44	11.25	3.03	2.08	7.29	1.00
Three	2,031	7.39	31.65	12.50	4.87	19.54	5.58	4.09	15.76	1.75
Four or More	3,130	9.30	29.55	10.36	5.72	14.42	3.44	3.74	7.73	0.88

(continued)

Table H.3 2020 NSDUH Respondent Pair-Level Proportions of Extreme Values and Outwinsors (continued)

Domain	*n*	Before res.pr.ps[1] (SDUWT*PRWT12*...*PRWT14) % Unweighted	Before res.pr.ps[1] (SDUWT*PRWT12*...*PRWT14) % Weighted[2]	Before res.pr.ps[1] (SDUWT*PRWT12*...*PRWT14) % Outwinsor[3]	After res.pr.ps[1] (SDUWT*PRWT12*...*PRWT15) % Unweighted	After res.pr.ps[1] (SDUWT*PRWT12*...*PRWT15) % Weighted[2]	After res.pr.ps[1] (SDUWT*PRWT12*...*PRWT15) % Outwinsor[3]	Final Weight: After res.pr.ev[1] (SDUWT*PRWT12*...*PRWT16) % Unweighted	Final Weight: After res.pr.ev[1] (SDUWT*PRWT12*...*PRWT16) % Weighted[2]	Final Weight: After res.pr.ev[1] (SDUWT*PRWT12*...*PRWT16) % Outwinsor[3]
Census Region										
Northeast	1,431	7.83	25.61	9.93	4.47	11.50	3.39	4.05	8.51	1.19
South	2,130	7.09	25.47	8.52	4.32	13.86	3.97	3.15	9.51	1.10
Midwest	1,720	7.21	28.42	10.40	4.83	12.62	2.86	3.08	6.75	0.85
West	1,803	6.54	32.65	12.87	4.77	20.29	4.74	3.44	12.47	1.30
Quarter										
Quarter 1	3,900	4.90	23.51	8.48	2.74	10.42	2.57	2.15	6.89	0.82
Quarter 4	3,184	9.86	32.09	11.93	6.85	19.08	5.07	4.90	12.10	1.41
% Hispanic or Latino in Segment										
50-100%	585	10.60	34.11	12.61	6.84	22.13	4.93	4.27	11.75	1.41
10-<50%	1,710	7.49	27.39	9.83	5.03	15.65	4.18	4.15	10.68	1.21
<10%	4,789	6.58	26.73	9.91	4.16	12.80	3.44	3.01	8.53	1.01
% Black or African American in Segment										
50-100%	387	13.18	44.14	15.93	12.40	28.27	8.06	11.63	23.85	3.13
10-<50%	1,245	8.92	29.03	11.44	5.14	15.01	3.53	4.02	9.79	1.10
<10%	5,452	6.29	26.05	9.38	3.91	13.55	3.54	2.66	8.13	0.93
% Owner-Occupied DUs[1] in Segment										
50-100%	5,322	6.69	27.88	10.09	4.17	13.74	3.59	2.72	8.01	0.89
10-<50%	1,373	10.49	30.57	11.76	6.55	20.39	5.15	5.97	16.56	2.06
<10%	389	1.29	6.27	3.46	3.34	9.68	2.49	3.34	7.20	1.42
Combined Median Rent/Housing Value										
1st Quintile	1,120	6.34	21.93	7.60	4.20	12.39	4.33	3.39	8.10	1.04
2nd Quintile	1,547	9.95	32.05	13.26	6.33	19.94	5.23	5.37	15.45	1.80
3rd Quintile	1,640	6.59	29.60	11.53	5.18	17.62	4.09	3.23	8.77	1.19
4th Quintile	1,518	6.39	27.49	8.03	3.36	11.83	2.84	2.70	8.16	0.79
5th Quintile	1,259	5.96	26.50	10.02	3.49	11.36	2.85	1.99	6.84	0.72
Population Density										
Large MSA[1]	2,939	7.93	28.78	10.41	5.75	16.04	3.74	4.42	11.15	1.23
Medium to Small MSA[1]	3,641	6.51	25.80	9.38	3.65	11.56	3.24	2.50	6.50	0.82
Non-MSA,[1] Urban	208	7.69	30.51	15.34	4.33	24.39	8.12	4.33	20.25	1.98
Non-MSA,[1] Rural	296	6.42	36.56	14.49	4.73	27.74	9.85	3.38	12.75	2.20
Group Quarters										
Group	46	6.52	26.02	7.90	2.17	19.60	3.38	0.00	0.00	0.00
Non-Group	7,038	7.13	27.94	10.26	4.60	14.89	3.86	3.41	9.59	1.12

(continued)

Table H.3 2020 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)	980	5.92	26.76	8.23	3.27	14.35	3.96	2.35	8.70	0.99
Parent-Child (12-17)	1,728	5.73	24.84	8.32	3.18	12.72	3.13	2.26	7.47	0.85
Parent-Child (12-20)	2,096	6.39	27.31	9.08	3.34	12.48	2.89	2.43	7.86	0.92
Sibling (12-14)-Sibling (15-17)	567	4.94	26.03	10.62	1.23	9.39	1.96	0.71	5.21	0.61
Sibling (12-17)-Sibling (18-25)	563	8.35	42.38	18.45	5.68	24.81	6.30	3.20	14.65	1.68
Spouse-Spouse/Partner-Partner	1,646	5.53	27.02	9.19	3.22	12.12	2.90	2.19	7.33	0.87
Spouse-Spouse/Partner-Partner with Children (Younger than 18)	711	6.61	24.52	9.08	5.63	15.72	4.35	3.94	10.82	1.29

DU = dwelling unit; ev = extreme value adjustment; MSA = metropolitan statistical area; pr = pair; ps = poststratification adjustment; res = respondent; SDU = screener dwelling unit.

[1] This step used demographic variables from questionnaire data for all responding person pairs.

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

[3] Outwinsor weight percentage: $100*\sum_k |w_{ek} - b_k|/\sum_k w_k$, where b_k denotes the cut-off point for defining the extreme weight.

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

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

This page intentionally left blank

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

This page intentionally left blank

Table I.1 2020 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	7,084	240,397,692	240,397,692	240,397,692	-0.00	-0.00
Pair Age Group						
12-17, 12-17	917	7,719,564	7,688,445	7,688,445	0.40	0.00
12-17, 18-25	599	8,181,049	8,282,841	8,282,841	-1.23	0.00
12-17, 26-34	253	4,875,548	4,826,331	4,826,331	1.02	-0.00
12-17, 35-49	1,377	29,502,701	29,372,080	29,372,080	0.44	-0.00
12-17, 50+	269	13,280,887	13,961,817	13,961,817	-4.88	0.00
18-25, 18-25	1,151	12,752,905	12,631,480	12,631,480	0.96	-0.00
18-25, 26-34	363	7,348,227	7,471,650	7,471,650	-1.65	-0.00
18-25, 35-49	455	16,072,806	15,926,764	15,926,765	0.92	-0.00
18-25, 50+	326	21,330,525	21,107,295	21,107,295	1.06	0.00
26-34, 26-34	436	12,394,077	12,390,970	12,390,970	0.03	-0.00
26-34, 35-49	206	9,384,627	9,333,675	9,333,675	0.55	0.00
26-34, 50+	85	14,388,881	14,451,082	14,451,082	-0.43	0.00
35-49, 35-49	324	19,173,706	19,445,946	19,445,946	-1.40	0.00
35-49, 50+	101	18,694,089	18,138,936	18,138,936	3.06	0.00
50+, 50+	222	45,298,100	45,368,379	45,368,379	-0.15	0.00
Pair Race/Ethnicity						
Hispanic or Latino	1,089	48,030,925	47,996,518	47,996,518	0.07	-0.00
Black or African American	611	26,037,331	26,427,678	26,427,678	-1.48	-0.00
White	3,819	119,635,997	120,443,269	120,443,269	-0.67	0.00
Other	600	18,105,946	18,289,699	18,289,699	-1.00	0.00
White & Black or African American	84	2,613,809	3,305,022	3,305,022	-20.91	-0.00
White & Hispanic or Latino	330	10,460,493	11,239,022	11,239,022	-6.93	-0.00
White & Other	348	9,237,844	7,724,641	7,724,641	19.59	-0.00
Black or African American & Hispanic or Latino	63	2,245,576	1,820,037	1,820,037	23.38	-0.00
Black or African American & Other	75	2,110,397	1,410,105	1,410,105	49.66	-0.00
Hispanic or Latino & Other	65	1,919,373	1,741,702	1,741,702	10.20	0.00
Pair Gender						
Male, Male	1,351	41,901,818	42,585,031	42,585,031	-1.60	-0.00
Female, Female	1,616	44,236,353	43,467,263	43,467,263	1.77	-0.00
Male, Female	4,117	154,259,521	154,345,398	154,345,398	-0.06	-0.00
Pair Relationship Domain[3,4,5]						
Parent-Child (12-14)*	980	12,001,635	12,179,472	12,179,472	-1.46	-0.00
Parent-Child (12-17)*	1,728	23,162,817	23,701,383	23,701,383	-2.27	-0.00
Parent-Child (15-17)*	748	11,161,183	11,521,911	11,521,911	-3.13	0.00
Parent-Child (12-20)*	2,096	32,931,434	32,364,655	32,364,655	1.75	-0.00
Parent*-Child (12-14)	980	19,226,314	18,780,525	18,780,525	2.37	-0.00
Parent*-Child (12-17)	1,728	30,510,682	30,516,893	30,516,893	-0.02	-0.00
Parent*-Child (15-17)	748	18,123,549	18,451,591	18,109,735	0.08	1.89
Parent*-Child (12-20)	2,096	39,027,787	38,003,040	38,003,040	2.70	-0.00
Sibling (12-14)-Sibling (15-17)*	567	4,199,153	4,035,857	4,035,857	4.05	0.00
Sibling (12-17)-Sibling (18-25)*	563	6,276,076	6,008,791	6,008,791	4.45	0.00
Spouse-Spouse/Partner-Partner	1,646	77,254,316	73,978,957	73,978,957	4.43	-0.00
Spouse-Spouse/Partner-Partner with Children (Younger Than 18)	711	24,577,758	27,722,610	27,722,610	-11.34	-0.00

(continued)

Table I.1 2020 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	1,923	58,989,240	58,989,240	58,989,240	0.00	0.00
Three	2,031	58,591,079	58,591,079	58,591,079	-0.00	-0.00
Four or More	3,130	122,817,373	122,817,373	122,817,373	-0.00	-0.00
Census Region						
Northeast	1,431	41,494,881	41,494,881	41,494,881	0.00	-0.00
South	2,130	89,865,421	89,865,421	89,865,421	-0.00	-0.00
Midwest	1,720	46,319,089	46,319,089	46,319,089	0.00	-0.00
West	1,803	62,718,301	62,718,301	62,718,301	-0.00	0.00
Quarter						
Quarter 1	3,900	116,381,915	116,381,915	116,381,915	0.00	0.00
Quarter 4	3,184	124,015,777	124,015,777	124,015,777	-0.00	-0.00
% Hispanic or Latino in Segment						
50-100%	585	33,311,179	33,311,179	33,311,179	-0.00	-0.00
10-<50%	1,710	67,484,470	67,484,470	67,484,470	0.00	-0.00
<10%	4,789	139,602,043	139,602,043	139,602,043	-0.00	-0.00
% Black or African American in Segment						
50-100%	387	16,999,536	16,999,536	16,999,536	-0.00	-0.00
10-<50%	1,245	48,823,527	48,823,527	48,823,527	0.00	0.00
<10%	5,452	174,574,629	174,574,629	174,574,629	0.00	0.00
% Owner-Occupied DUs in Segment						
50-100%	5,322	190,751,347	190,751,347	190,751,347	-0.00	-0.00
10-<50%	1,373	44,672,435	44,672,435	44,672,435	0.00	-0.00
<10%	389	4,973,911	4,973,911	4,973,911	0.00	0.00
Combined Median Rent/Housing Value						
1st Quintile	1,120	35,589,312	35,589,312	35,589,312	0.00	0.00
2nd Quintile	1,547	50,575,186	50,575,186	50,575,186	0.00	-0.00
3rd Quintile	1,640	56,448,397	56,448,397	56,448,397	-0.00	0.00
4th Quintile	1,518	52,751,267	52,751,267	52,751,267	-0.00	0.00
5th Quintile	1,259	45,033,531	45,033,531	45,033,531	-0.00	0.00
Population Density						
Large MSA	2,939	132,339,355	132,339,355	132,339,355	-0.00	-0.00
Medium to Small MSA	3,641	94,160,816	94,160,815	94,160,815	0.00	-0.00
Non-MSA, Urban	208	5,125,616	5,125,616	5,125,616	0.00	0.00
Non-MSA, Rural	296	8,771,906	8,771,906	8,771,906	0.00	0.00
Group Quarters						
Group	46	224,380	224,758	224,758	-0.17	-0.00
Non-Group	7,038	240,173,313	240,172,934	240,172,934	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 model and thus have higher slippage rates than the other domains listed. However, because 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 because no household counts for this domain were calculated. Household counts are required to construct the appropriate control totals.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 2, 2020.

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

This page intentionally left blank

Table J.1 2020 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	18,461	20	899	2,251	4,174	48,748	1.96	33	3,440	9,318	24,001	3,625,562	9.99	18	2,758	8,381	24,086	1,011,897	5.89
Pair Age Group																			
12-17, 12-17	2,165	35	777	1,977	3,733	29,807	1.98	49	1,623	3,965	8,505	134,524	2.91	18	1,095	3,221	8,617	117,312	3.37
12-17, 18-25	1,743	35	962	2,257	4,266	23,211	1.90	83	2,080	5,113	10,915	319,977	3.87	36	1,897	4,995	12,029	106,457	2.73
12-17, 26-34	712	24	937	2,420	4,708	26,301	1.84	71	3,459	8,279	15,997	253,203	2.93	30	2,376	6,198	15,862	170,268	3.16
12-17, 35-49	3,602	37	827	2,115	3,911	32,886	1.96	122	3,728	9,220	18,606	313,620	3.02	55	3,073	7,831	19,554	255,501	3.04
12-17, 50+	730	49	1,073	2,602	4,453	20,376	1.77	493	11,008	25,040	42,315	423,621	2.19	411	10,542	23,657	50,493	299,828	2.26
18-25, 18-25	2,885	32	824	2,098	4,256	35,147	2.04	33	1,698	4,580	10,073	236,389	3.60	22	1,382	4,092	10,764	120,699	3.06
18-25, 26-34	960	20	1,018	2,454	4,508	24,588	1.99	55	3,989	9,003	17,992	694,558	5.63	28	2,584	7,110	17,704	251,324	3.60
18-25, 35-49	1,163	68	1,125	2,581	4,533	23,816	1.89	307	6,112	13,205	29,168	408,603	3.10	154	5,091	14,357	32,783	296,563	2.83
18-25, 50+	890	31	1,120	2,754	4,558	24,411	1.79	540	12,705	28,151	54,174	694,825	2.87	285	11,431	28,326	60,094	937,930	2.70
26-34, 26-34	1,010	21	866	2,085	4,175	26,097	2.11	150	6,498	14,572	29,831	690,147	3.75	97	3,859	10,003	23,811	755,848	5.69
26-34, 35-49	522	43	890	2,510	4,722	48,748	2.25	302	8,604	20,817	44,903	965,598	4.55	253	5,952	15,667	41,480	588,204	4.01
26-34, 50+	300	94	1,085	2,849	5,056	21,312	1.71	2,509	30,211	62,248	115,488	1,100,048	2.57	2,006	23,323	53,065	132,522	604,629	2.23
35-49, 35-49	879	68	907	2,106	4,034	22,033	1.88	530	9,291	22,489	45,211	2,758,107	8.15	347	6,678	17,074	45,513	906,880	4.47
35-49, 50+	299	78	1,166	2,544	4,237	22,469	1.86	1,558	31,379	60,585	104,978	3,410,245	6.28	1,106	30,840	72,535	153,304	1,011,897	2.39
50+, 50+	601	38	943	2,183	3,911	21,419	1.85	2,128	42,218	93,937	154,535	3,625,562	4.44	2,559	48,188	100,525	202,834	1,008,312	1.99
Pair Race/Ethnicity																			
Hispanic or Latino	2,990	37	1,618	3,073	5,396	32,886	1.75	59	5,289	12,686	31,027	3,625,562	9.06	18	3,763	11,658	32,104	937,930	5.38
Black or African American	1,507	51	1,892	3,717	6,017	24,480	1.63	103	6,027	14,033	32,154	3,410,245	14.10	56	4,435	14,489	37,590	1,008,312	4.49
White	10,222	21	799	1,995	3,718	32,582	1.90	33	3,070	8,369	21,896	3,602,542	9.16	20	2,635	7,599	20,776	1,011,897	6.46
Other	1,627	31	614	1,629	3,664	23,994	2.21	53	2,326	6,659	17,048	1,371,088	7.55	28	1,994	6,108	20,315	906,880	6.24
White & Black or African American	222	46	1,313	2,690	4,789	23,196	1.92	178	5,399	14,755	27,306	507,687	4.29	68	3,477	11,339	28,577	455,136	4.26
White & Hispanic or Latino	801	24	824	1,976	3,757	48,748	2.42	49	3,020	8,590	23,556	408,184	3.89	30	2,697	8,174	24,451	588,204	5.08
White & Other	735	20	660	1,520	3,261	35,147	2.49	55	3,017	7,705	19,829	687,431	5.16	42	2,049	6,198	18,133	439,371	5.24
Black or African American & Hispanic or Latino	110	117	1,217	3,143	6,420	17,658	1.85	143	4,114	12,689	27,906	687,618	6.92	57	3,831	11,901	44,120	315,065	3.34
Black or African American & Other	111	80	622	1,664	3,832	29,807	2.83	326	2,761	8,089	21,616	424,268	6.13	70	2,723	9,646	26,325	277,939	4.23
Hispanic or Latino & Other	136	35	643	1,508	3,710	11,259	1.89	64	2,337	7,810	22,814	1,137,084	12.26	89	1,344	3,514	19,101	394,370	6.71

(continued)

Table J.1 2020 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	3,741	20	939	2,385	4,307	32,886	1.99	49	3,263	8,528	21,982	596,056	4.58	20	2,766	8,604	23,377	755,848	4.78
Female, Female	4,037	24	861	2,195	4,050	29,807	1.94	33	3,206	8,423	20,050	3,410,245	11.66	32	2,488	7,327	21,740	675,945	5.17
Male, Female	10,683	21	896	2,235	4,172	48,748	1.95	40	3,606	10,076	26,537	3,625,562	10.37	18	2,868	8,730	25,508	1,011,897	6.14
Household Size																			
Two	4,630	20	838	2,113	4,049	48,748	2.01	33	3,427	10,038	26,687	687,431	4.24	22	2,180	6,744	20,509	669,056	5.53
Three	5,540	22	903	2,271	4,147	32,886	1.97	59	3,045	8,155	18,065	2,758,107	13.13	18	2,469	7,151	19,191	1,008,312	7.12
Four or More	8,291	28	937	2,333	4,283	32,582	1.93	49	3,709	10,119	27,342	3,625,562	11.06	28	3,537	10,673	29,616	1,011,897	5.42
Census Region																			
Northeast	3,857	20	743	1,781	3,472	26,301	2.01	33	2,610	7,235	18,425	3,625,562	16.37	20	1,831	6,057	17,935	906,880	7.17
South	5,328	31	1,517	3,227	5,601	48,748	1.76	88	5,716	13,726	33,319	3,602,542	9.80	18	4,633	12,519	32,438	1,011,897	5.06
Midwest	4,499	24	931	2,063	3,427	24,480	1.72	59	3,221	7,927	18,834	1,586,399	6.29	30	2,646	7,352	19,294	984,054	5.77
West	4,777	35	651	1,798	4,292	27,837	2.13	49	2,737	8,451	23,825	1,371,088	6.66	37	2,298	7,638	24,703	723,767	5.84
Quarter																			
Quarter 1	7,129	65	1,777	3,366	5,224	48,748	1.62	94	5,964	13,291	31,573	3,625,562	7.79	30	4,593	12,278	32,250	1,011,897	4.93
Quarter 4	11,332	20	706	1,632	3,262	32,886	2.24	33	2,551	7,123	19,103	3,602,542	12.14	18	2,075	6,396	19,358	952,486	6.68
% Hispanic or Latino in Segment																			
50-100%	1,622	68	2,169	3,841	6,102	35,147	1.53	265	7,780	17,453	38,965	3,625,562	11.98	58	5,897	16,298	42,571	937,930	4.50
10-<50%	4,561	31	1,300	2,725	4,972	32,582	1.83	88	4,617	11,903	27,974	3,410,245	9.01	18	3,609	10,336	28,644	906,880	5.08
<10%	12,278	20	731	1,864	3,665	48,748	2.06	33	2,823	7,818	20,230	3,602,542	8.55	20	2,296	7,049	20,261	1,011,897	6.48
% Black or African American in Segment																			
50-100%	962	35	1,855	3,559	5,885	24,728	1.68	94	6,210	14,067	34,265	1,100,048	5.17	90	4,401	14,494	37,480	1,008,312	4.41
10-<50%	3,275	31	1,422	2,939	5,164	29,807	1.77	88	5,087	12,081	28,195	3,625,562	14.70	18	3,861	10,952	28,737	937,930	4.97
<10%	14,224	20	796	2,008	3,841	48,748	2.01	33	3,024	8,451	22,334	3,602,542	8.50	20	2,523	7,662	21,967	1,011,897	6.30
% Owner-Occupied DUs[1] in Segment																			
50-100%	13,846	24	888	2,172	4,027	48,748	1.93	53	3,541	9,494	24,560	3,625,562	10.40	20	2,983	9,032	25,290	1,011,897	5.81
10-<50%	3,642	20	887	2,408	4,614	28,202	1.94	33	3,115	8,730	21,747	2,758,107	9.48	22	2,498	8,024	23,771	723,767	5.57
<10%	973	31	1,081	2,875	5,510	35,147	2.06	88	3,309	9,892	23,336	319,977	3.61	18	1,310	3,998	10,702	224,398	4.30
Combined Median Rent/Housing Value																			
1st Quintile	2,759	35	976	2,327	4,366	24,588	1.88	49	3,541	9,461	24,399	3,625,562	12.59	40	2,766	8,345	24,311	755,848	5.68
2nd Quintile	3,872	31	827	2,317	4,266	32,886	1.93	53	3,234	9,309	23,846	3,410,245	11.35	18	2,599	8,332	23,832	1,008,312	6.29
3rd Quintile	4,298	47	888	2,284	4,187	32,582	1.96	90	3,400	9,432	23,917	1,371,088	6.21	27	2,472	7,717	24,312	984,054	6.02
4th Quintile	4,113	20	899	2,244	4,234	35,147	1.99	33	3,634	9,542	24,692	3,602,542	10.00	20	2,890	8,556	24,011	952,486	5.74
5th Quintile	3,419	28	926	2,079	3,884	48,748	2.03	59	3,491	8,927	22,849	2,758,107	11.03	36	3,127	9,003	23,862	1,011,897	5.60

(continued)

Table J.1 2020 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]	7,836	35	1,645	3,115	5,069	48,748	1.72	94	6,075	13,489	32,149	3,625,562	10.42	28	5,065	12,993	33,596	1,008,312	4.69
Medium to Small MSA[1]	9,451	20	599	1,569	3,324	28,202	2.12	33	2,263	6,632	17,707	936,327	6.01	18	1,733	5,683	17,351	1,011,897	7.01
Non-MSA,[1] Urban	511	35	658	1,556	3,603	15,906	2.03	49	2,359	6,492	17,311	1,351,344	12.82	34	2,392	6,726	19,155	697,924	6.06
Non-MSA,[1] Rural	663	53	740	1,937	4,048	32,886	2.23	137	2,844	8,715	23,272	1,586,399	9.34	60	1,810	6,291	20,819	984,054	9.09
Group Quarters																			
Group	64	284	796	2,523	4,201	9,966	1.69	531	1,840	5,777	8,318	57,972	2.55	138	1,054	2,739	6,898	74,638	4.02
Non-Group	18,397	20	900	2,251	4,174	48,748	1.96	33	3,446	9,356	24,048	3,625,562	9.98	18	2,775	8,421	24,156	1,011,897	5.88

DU = dwelling unit; MSA = metropolitan statistical area; pr = pair; ps = poststratification; SDU = screener dwelling unit; sel = selected; UWE = unequal weighting effect.

[1] This step used demographic variables from screener data for all selected person pairs.

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

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

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

Table J.2 2020 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	7,084	22	3,072	9,074	24,743	1,011,897	5.31	34	4,554	14,598	50,518	3,433,813	7.68
Pair Age Group													
12-17, 12-17	913	28	1,288	3,959	9,574	109,122	3.23	34	1,962	6,122	16,284	416,815	4.36
12-17, 18-25	609	55	2,261	5,721	13,899	106,457	2.71	55	3,440	10,162	29,393	399,243	4.14
12-17, 26-34	253	99	2,967	7,202	17,029	170,268	3.36	236	4,661	13,100	32,289	805,638	5.70
12-17, 35-49	1,375	55	3,712	8,792	21,816	224,337	2.95	55	5,088	14,362	38,617	773,224	4.90
12-17, 50+	274	411	13,334	27,244	52,962	291,076	2.15	411	18,069	47,469	145,082	941,717	2.70
18-25, 18-25	1,162	22	1,467	4,317	11,287	104,341	3.07	37	1,920	6,180	19,791	413,736	4.93
18-25, 26-34	359	44	2,461	6,808	16,861	251,324	3.86	63	3,955	9,816	32,135	891,463	5.84
18-25, 35-49	457	154	6,758	17,100	38,460	296,563	2.63	172	9,051	25,784	83,032	1,389,494	4.16
18-25, 50+	324	285	13,081	27,316	52,458	364,424	2.38	780	19,665	46,935	178,200	1,096,496	3.05
26-34, 26-34	421	97	3,984	9,852	18,777	472,413	4.72	365	5,751	13,841	38,954	1,658,411	7.68
26-34, 35-49	208	253	5,827	16,748	46,564	413,619	2.92	465	9,264	23,969	108,665	1,109,520	3.96
26-34, 50+	86	2,601	23,873	51,499	135,447	511,238	2.18	6,658	72,414	234,390	470,698	1,782,715	2.11
35-49, 35-49	324	347	7,573	18,081	51,489	639,325	3.56	347	12,380	36,814	103,421	2,137,366	4.73
35-49, 50+	98	2,658	29,080	73,909	154,039	1,011,897	2.63	2,843	76,109	192,435	506,427	1,987,353	2.35
50+, 50+	221	3,819	51,864	108,381	196,251	772,087	1.85	11,224	86,851	210,778	487,388	3,433,813	2.72
Pair Race/Ethnicity													
Hispanic or Latino	1,088	34	4,212	12,420	32,533	891,523	4.71	34	6,083	21,330	76,674	3,433,813	6.73
Black or African American	629	62	4,820	14,057	35,977	669,056	3.78	90	7,118	22,401	77,535	2,172,504	5.95
White	3,911	22	2,944	8,427	21,965	1,011,897	5.91	42	4,221	13,204	43,397	3,060,480	8.81
Other	632	37	2,108	6,049	19,336	488,589	5.20	37	2,887	9,941	40,381	1,881,091	8.64
White & Black or African American	71	132	3,141	9,249	29,177	229,024	3.60	297	5,915	26,919	91,127	649,565	3.78
White & Hispanic or Latino	309	63	2,964	8,293	23,677	412,433	4.70	63	4,895	16,959	80,082	1,109,520	4.71
White & Other	292	54	2,537	6,233	19,731	439,371	5.01	55	3,712	10,419	40,434	1,183,755	6.34
Black or African American & Hispanic or Latino	52	696	2,854	7,927	40,329	171,787	2.86	1,075	7,610	36,616	81,675	692,266	3.58
Black or African American & Other	46	587	4,617	9,666	19,914	217,529	4.24	587	5,616	10,420	32,630	693,065	6.26
Hispanic or Latino & Other	54	157	1,387	3,105	19,802	394,370	6.53	157	1,477	6,251	34,891	608,962	5.08
Pair Gender													
Male, Male	1,349	28	2,926	8,528	22,571	399,493	4.25	34	4,773	15,361	49,639	1,987,353	7.11
Female, Female	1,607	32	2,975	8,849	23,918	675,945	4.59	55	4,294	13,161	43,963	1,593,117	6.42
Male, Female	4,128	22	3,165	9,412	26,322	1,011,897	5.55	37	4,643	14,977	52,387	3,433,813	7.90
Household Size													
Two	1,923	22	2,283	7,171	20,015	669,056	5.69	37	2,848	9,837	37,759	3,433,813	9.03
Three	2,031	34	3,094	8,173	20,467	658,691	5.53	34	4,611	13,222	40,820	3,060,480	8.93
Four or More	3,130	28	3,778	11,360	30,562	1,011,897	4.87	50	5,924	20,420	65,910	2,604,750	6.48

(continued)

Table J.2 2020 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.nr) (continued)

Domain	*n*	Before res.pr.nr[1] (SDUWT*PRWT12*PRWT13) Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	After res.pr.nr[1] (SDUWT*PRWT12*...*PRWT14) Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
Census Region													
Northeast	1,431	22	2,006	6,694	19,759	891,523	6.55	34	3,142	10,753	40,096	2,172,504	8.21
South	2,130	50	5,070	13,158	33,789	1,011,897	4.40	50	7,044	20,626	71,229	2,604,750	6.49
Midwest	1,720	64	2,987	8,206	19,581	772,087	4.77	89	4,487	13,313	41,331	2,540,730	7.26
West	1,803	37	2,317	7,649	24,203	658,691	5.97	37	3,373	13,526	50,096	3,433,813	8.78
Quarter													
Quarter 1	3,900	34	4,371	11,514	29,383	1,011,897	4.59	34	5,035	14,567	43,787	3,060,480	7.80
Quarter 4	3,184	22	2,067	6,291	19,055	891,523	6.53	37	3,771	14,664	63,107	3,433,813	7.37
% Hispanic or Latino in Segment													
50-100%	585	182	5,513	14,959	39,196	675,945	3.98	423	9,122	32,877	111,096	1,881,091	4.64
10-<50%	1,710	55	3,869	10,947	29,131	445,423	3.93	55	5,687	18,347	68,241	2,604,750	6.36
<10%	4,789	22	2,645	7,852	21,006	1,011,897	6.15	34	3,788	12,760	41,399	3,433,813	9.08
% Black or African American in Segment													
50-100%	387	90	4,039	13,822	35,901	669,056	3.77	90	7,258	24,954	90,096	2,172,504	5.24
10-<50%	1,245	62	4,391	12,065	29,151	458,040	4.01	68	6,605	20,538	74,084	1,987,353	5.32
<10%	5,452	22	2,818	8,221	22,856	1,011,897	5.85	34	3,931	13,051	44,869	3,433,813	8.69
% Owner-Occupied DUs[1] in Segment													
50-100%	5,322	28	3,448	9,852	26,869	1,011,897	5.30	34	4,818	15,314	51,385	3,433,813	7.84
10-<50%	1,373	22	2,675	8,497	22,573	511,238	4.24	68	4,724	15,408	56,668	1,782,715	5.98
<10%	389	41	1,155	3,850	10,702	136,613	4.02	42	1,891	5,945	21,334	545,406	5.74
Combined Median Rent/Housing Value													
1st Quintile	1,120	55	2,815	9,033	24,086	637,660	4.82	64	3,501	12,930	42,149	2,052,532	7.60
2nd Quintile	1,547	28	2,958	8,937	23,455	772,087	4.93	34	4,548	15,575	56,867	2,540,730	6.46
3rd Quintile	1,640	64	2,657	7,817	23,799	511,238	5.19	89	4,102	13,282	45,529	3,433,813	8.59
4th Quintile	1,518	22	3,373	9,464	26,798	675,945	5.11	37	4,854	16,024	55,234	2,604,750	7.80
5th Quintile	1,259	50	3,813	10,377	27,151	1,011,897	6.00	50	5,602	15,304	52,949	3,060,480	7.71
Population Density													
Large MSA[1]	2,939	50	5,636	13,930	34,843	891,523	4.43	50	8,083	23,262	75,766	3,060,480	6.32
Medium to Small MSA[1]	3,641	22	1,923	6,185	18,526	1,011,897	5.53	37	3,002	10,714	37,472	3,433,813	8.15
Non-MSA,[1] Urban	208	34	2,491	6,160	17,365	231,204	4.46	34	2,627	7,260	27,331	879,723	7.36
Non-MSA,[1] Rural	296	60	1,588	4,736	17,807	772,087	9.91	64	1,776	7,052	28,214	2,540,730	16.34
Group Quarters													
Group	46	138	799	2,059	9,176	74,638	4.22	138	799	2,059	10,802	108,913	5.03
Non-Group	7,038	22	3,094	9,119	24,875	1,011,897	5.29	34	4,598	14,727	51,019	3,433,813	7.65

DU = dwelling unit; MSA = metropolitan statistical area; nr = nonresponse adjustment; pr = pair; res = respondent; SDU = screener dwelling unit; UWE = unequal weighting effect.

[1] This step used demographic variables from screener data for all selected person pairs.

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

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

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

Table J.3 2020 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	7,084	34	4,554	14,598	50,518	3,433,813	7.68	9	2,153	7,919	28,098	1,276,017	6.94	9	2,134	7,989	28,755	1,142,106	6.83
Pair Age Group																			
12-17, 12-17	917	34	1,976	6,149	16,284	416,815	4.34	12	956	3,011	9,581	129,816	3.91	11	920	3,016	9,530	133,699	3.78
12-17, 18-25	599	55	3,382	10,157	29,909	399,243	4.14	12	1,606	6,005	17,376	161,012	3.31	10	1,640	6,250	17,993	130,004	3.04
12-17, 26-34	253	236	4,797	13,381	33,605	805,638	5.60	58	2,451	7,458	22,445	299,298	3.99	52	2,363	7,727	22,154	316,288	4.06
12-17, 35-49	1,377	55	5,039	14,362	38,617	773,224	4.88	13	2,345	7,592	21,641	473,236	4.22	11	2,323	7,886	22,050	335,296	3.95
12-17, 50+	269	411	18,057	46,391	144,654	821,002	2.56	145	8,451	25,512	74,491	464,281	2.58	97	8,586	25,697	73,971	461,876	2.59
18-25, 18-25	1,151	37	1,899	6,058	19,876	413,736	5.01	9	848	3,246	12,719	150,033	3.84	9	907	3,374	13,267	145,213	3.57
18-25, 26-34	363	63	4,034	10,368	31,893	891,463	5.92	24	2,090	6,679	19,605	303,956	4.00	20	2,070	6,626	21,732	338,394	4.22
18-25, 35-49	455	172	8,959	25,382	83,853	1,389,494	4.15	84	3,854	14,431	43,117	434,718	3.43	79	3,893	14,555	44,649	384,103	3.30
18-25, 50+	326	780	19,963	46,760	175,019	1,096,496	3.13	239	10,446	31,921	79,025	671,142	2.79	240	10,497	32,325	77,190	620,667	2.70
26-34, 26-34	436	365	5,594	13,795	39,125	1,658,411	7.76	142	2,873	6,370	23,659	811,580	7.05	105	2,931	6,466	24,578	841,267	6.84
26-34, 35-49	206	465	9,095	25,333	101,576	1,109,520	4.02	112	3,453	11,148	42,785	703,674	5.26	88	2,894	11,315	43,081	761,209	5.34
26-34, 50+	85	6,658	77,410	235,323	472,220	1,782,715	2.10	1,589	39,544	112,435	223,085	862,406	2.07	1,463	38,948	103,964	229,054	731,703	2.07
35-49, 35-49	324	347	12,380	36,141	102,765	2,137,366	4.75	73	4,803	17,299	54,364	1,163,616	5.52	55	4,629	16,977	54,677	1,142,106	5.69
35-49, 50+	101	2,843	82,501	194,210	489,167	1,987,353	2.33	936	31,491	88,927	274,192	1,102,470	2.32	831	27,793	88,595	274,657	1,129,581	2.36
50+, 50+	222	9,795	86,107	210,135	484,655	3,433,813	2.74	4,865	47,459	124,386	269,332	1,276,017	2.21	3,945	48,544	128,817	264,741	993,675	2.12
Pair Race/Ethnicity																			
Hispanic or Latino	1,089	34	6,165	21,055	76,643	3,433,813	6.73	12	2,899	11,814	40,136	1,102,470	6.45	10	2,972	11,968	41,351	1,129,581	6.19
Black or African American	611	90	7,389	22,657	78,749	2,172,504	5.92	19	3,364	13,943	45,459	1,163,616	4.99	13	3,304	13,866	47,263	1,142,106	4.99
White	3,819	42	4,215	13,180	43,494	3,060,480	8.75	20	2,125	7,058	24,293	1,276,017	7.61	16	2,130	7,116	24,439	993,675	7.46
Other	600	37	2,936	10,947	43,230	1,881,091	8.10	9	1,390	6,404	23,961	816,135	6.65	9	1,362	6,876	25,009	857,075	6.85
White & Black or African American	84	297	5,400	17,577	48,170	649,565	5.18	75	4,051	14,489	48,186	341,231	3.34	86	4,924	15,955	47,373	373,636	3.49
White & Hispanic or Latino	330	63	4,585	14,924	63,472	1,109,520	4.79	24	2,028	7,711	31,176	703,674	5.95	20	2,027	8,403	32,991	761,209	6.07
White & Other	348	55	3,665	10,567	41,353	1,183,755	6.42	12	1,248	4,551	17,254	591,245	7.72	10	1,185	4,831	17,985	549,717	7.76
Black or African American & Hispanic or Latino	63	542	6,037	27,544	85,477	692,266	3.45	155	2,046	6,858	35,419	287,576	4.18	204	1,682	10,144	31,884	273,965	3.94
Black or African American & Other	75	68	4,411	13,418	41,240	693,065	5.16	14	1,673	6,015	13,941	311,650	6.76	11	1,567	6,064	14,010	323,047	6.89
Hispanic or Latino & Other	65	55	1,682	5,749	20,631	1,058,604	8.02	13	749	3,015	16,577	240,325	5.28	11	624	3,147	18,185	245,315	5.38

(continued)

Table J.3 2020 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	1,351	34	4,684	15,142	49,474	1,987,353	7.22	12	2,550	9,266	30,439	936,514	5.64	11	2,632	9,288	31,451	993,675	5.59
Female, Female	1,616	55	4,306	13,295	44,538	1,593,117	6.37	12	2,000	7,009	24,295	958,240	6.52	10	1,977	7,241	25,996	945,440	6.47
Male, Female	4,117	37	4,650	14,959	52,393	3,433,813	7.89	9	2,109	7,778	28,108	1,276,017	7.16	9	2,074	7,784	28,914	1,142,106	7.02
Household Size																			
Two	1,923	37	2,848	9,837	37,759	3,433,813	9.03	9	1,286	4,775	19,498	949,599	7.71	9	1,231	4,676	20,204	856,741	7.73
Three	2,031	34	4,611	13,222	40,820	3,060,480	8.93	12	2,211	7,140	22,771	1,276,017	8.17	10	2,212	7,319	23,692	993,675	7.67
Four or More	3,130	50	5,924	20,420	65,910	2,604,750	6.48	20	3,003	11,656	36,098	1,163,616	5.99	21	3,016	11,903	36,472	1,142,106	5.96
Census Region																			
Northeast	1,431	34	3,142	10,753	40,096	2,172,504	8.21	12	1,464	5,261	22,106	1,102,470	8.15	11	1,421	5,240	22,507	1,129,581	8.21
South	2,130	50	7,044	20,626	71,229	2,604,750	6.49	19	3,477	11,278	39,398	1,276,017	6.08	13	3,456	11,738	40,358	1,142,106	5.96
Midwest	1,720	89	4,487	13,313	41,331	2,540,730	7.26	20	2,162	7,378	21,883	924,006	6.69	16	2,171	7,510	23,150	923,562	6.79
West	1,803	37	3,373	13,526	50,096	3,433,813	8.78	9	1,512	6,952	28,751	949,599	7.07	9	1,485	7,119	29,018	857,075	6.76
Quarter																			
Quarter 1	3,900	34	5,035	14,567	43,787	3,060,480	7.80	12	2,241	7,715	24,339	936,514	6.74	11	2,156	7,603	24,980	993,675	6.70
Quarter 4	3,184	37	3,771	14,664	63,107	3,433,813	7.37	9	2,061	8,470	33,580	1,276,017	6.88	9	2,107	8,599	33,800	1,142,106	6.71
% Hispanic or Latino in Segment																			
50-100%	585	423	9,122	32,877	111,096	1,881,091	4.64	96	4,350	19,068	57,653	958,240	4.91	89	4,452	19,767	57,818	945,440	4.90
10-<50%	1,710	55	5,687	18,347	68,241	2,604,750	6.36	12	2,865	10,207	36,661	869,585	5.47	10	2,871	10,449	37,590	838,240	5.33
<10%	4,789	34	3,788	12,760	41,399	3,433,813	9.08	9	1,848	6,611	22,106	1,276,017	8.05	9	1,826	6,707	22,721	1,142,106	7.91
% Black or African American in Segment																			
50-100%	387	90	7,258	24,954	90,096	2,172,504	5.24	19	3,617	14,994	54,271	692,130	4.15	13	3,754	15,111	51,106	725,848	4.20
10-<50%	1,245	68	6,605	20,538	74,084	1,987,353	5.32	14	3,391	11,416	41,176	727,271	4.70	11	3,263	11,960	39,972	755,071	4.75
<10%	5,452	34	3,931	13,051	44,869	3,433,813	8.69	9	1,872	6,809	24,451	1,276,017	7.97	9	1,878	6,893	25,152	1,142,106	7.78
% Owner-Occupied DUs[1] in Segment																			
50-100%	5,322	34	4,818	15,314	51,385	3,433,813	7.84	9	2,396	8,577	29,412	1,276,017	6.79	9	2,414	8,648	29,680	1,142,106	6.66
10-<50%	1,373	68	4,724	15,408	56,668	1,782,715	5.98	14	1,924	7,738	28,676	924,006	6.84	11	1,842	7,585	29,423	923,562	6.76
<10%	389	42	1,891	5,945	21,334	545,406	5.74	12	809	3,140	12,644	188,847	4.67	10	782	3,144	13,136	212,007	4.74
Combined Median Rent/Housing Value																			
1st Quintile	1,120	64	3,501	12,930	42,149	2,052,532	7.60	14	1,643	6,161	22,936	1,276,017	8.26	11	1,539	6,168	23,977	993,675	7.83
2nd Quintile	1,547	34	4,548	15,575	56,867	2,540,730	6.46	12	2,258	8,713	30,497	873,475	6.09	10	2,206	8,747	31,531	916,870	6.04
3rd Quintile	1,640	89	4,102	13,282	45,529	3,433,813	8.59	19	1,905	7,155	24,476	949,599	7.08	13	1,796	7,141	26,011	857,075	7.03
4th Quintile	1,518	37	4,854	16,024	55,234	2,604,750	7.80	9	2,400	8,509	29,707	1,163,616	6.82	9	2,472	8,414	30,532	1,142,106	6.69
5th Quintile	1,259	50	5,602	15,304	52,949	3,060,480	7.71	45	2,938	8,675	30,862	1,102,470	6.81	38	2,870	8,997	31,526	1,129,581	6.79

(continued)

Table J.3 2020 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]	2,939	50	8,083	23,262	75,766	3,060,480	6.32	14	4,550	13,411	41,774	1,163,616	5.66	11	4,557	13,421	42,416	1,142,106	5.68
Medium to Small MSA[1]	3,641	37	3,002	10,714	37,472	3,433,813	8.15	9	1,411	5,226	19,945	949,599	7.40	9	1,375	5,203	20,477	993,675	7.26
Non-MSA,[1] Urban	208	34	2,627	7,260	27,331	879,723	7.36	12	1,245	3,567	14,290	464,151	8.09	11	1,167	3,622	15,345	492,102	7.73
Non-MSA,[1] Rural	296	64	1,776	7,052	28,214	2,540,730	16.34	37	718	3,727	16,028	1,276,017	14.79	40	738	3,695	18,120	916,870	12.23
Group Quarters																			
Group	46	138	799	2,059	10,802	108,913	5.03	31	473	1,278	3,678	52,104	5.80	26	362	1,290	4,485	54,910	6.07
Non-Group	7,038	34	4,598	14,727	51,019	3,433,813	7.65	9	2,196	8,011	28,331	1,276,017	6.91	9	2,171	8,096	28,950	1,142,106	6.79
Pair Relationship Domain[4]																			
Parent-Child (12-14)	980	118	4,734	14,242	42,878	773,224	4.50	50	2,239	7,351	22,745	473,236	4.34	46	2,247	7,478	22,728	335,296	3.98
Parent-Child (12-17)	1,728	55	5,512	16,208	45,576	821,002	4.41	13	2,581	8,817	25,224	473,236	4.15	11	2,562	8,943	25,876	461,876	3.98
Parent-Child (12-20)	2,096	55	5,961	18,307	52,241	1,389,494	4.76	13	2,867	9,811	28,687	473,236	4.17	11	2,831	10,010	29,075	461,876	4.02
Sibling (12-14)-Sibling (15-17)	567	34	1,751	5,902	16,712	416,815	4.70	12	750	2,922	8,908	102,028	3.78	11	753	2,867	8,981	103,202	3.70
Sibling (12-17)-Sibling (18-25)	563	55	3,398	10,176	30,588	399,243	4.14	12	1,574	5,984	17,482	161,012	3.33	10	1,569	6,129	18,023	130,004	3.06
Spouse-Spouse/ Partner-Partner	1,646	37	3,791	13,462	66,240	3,433,813	8.12	9	1,817	6,326	31,259	1,163,616	7.19	9	1,791	6,340	30,734	1,142,106	7.24
Spouse-Spouse/ Partner-Partner with Children (Younger than 18)	711	61	4,175	13,252	50,096	2,137,366	7.65	18	2,297	7,643	29,618	1,163,616	8.21	15	2,250	7,494	28,920	1,142,106	8.43

DU = dwelling unit; ev = extreme value adjustment; MSA = metropolitan statistical area; pr = pair; ps = poststratification adjustment; res = respondent; SDU = screener dwelling unit; UWE = unequal weighting effect.

[1] This step used demographic variables from questionnaire data for all selected person pairs.

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

[3] 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 because extreme values were not controlled within this domain.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

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