2011 NATIONAL SURVEY ON DRUG USE AND HEALTH

QUESTIONNAIRE DWELLING UNIT-LEVEL AND PERSON PAIR-LEVEL SAMPLING WEIGHT CALIBRATION

Prepared for the 2011 Methodological Resource Book

RTI Project No. 0211838.207.005
Contract No. HHSS283200800004C
Deliverable No. 39

Authors:

Matthew Westlake
Patrick Chen
Harper Gordek

Project Director: Thomas G. Virag

Prepared for:

Substance Abuse and Mental Health Services Administration
Rockville, Maryland 20857

Prepared by:

RTI International
Research Triangle Park, North Carolina 27709

January 2013

2011 NATIONAL SURVEY ON DRUG USE AND HEALTH

QUESTIONNAIRE DWELLING UNIT-LEVEL AND PERSON PAIR-LEVEL SAMPLING WEIGHT CALIBRATION

Prepared for the 2011 Methodological Resource Book

RTI Project No. 0211838.207.005
Contract No. HHSS283200800004C
Deliverable No. 39

Authors:

Matthew Westlake
Patrick Chen
Harper Gordek

Project Director:

Thomas G. Virag

Prepared for:

Substance Abuse and Mental Health Services Administration
Rockville, Maryland 20857

Prepared by:

RTI International
Research Triangle Park, North Carolina 27709

January 2013

Preface

This report documents the method of weight calibration used for producing the final set of questionnaire dwelling unit (QDU) and pair weights for the National Survey on Drug Use and Health (NSDUH) data from 2011. The weighting team faced several challenges in this task and was able to address them by resorting to innovative modifications of certain basic statistical ideas, which are listed below.

- Under Brewer's method, high weights may occur due to 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 Chen et al. (2013), extreme value 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 persons aged 12 to 25 in the household, State, and clusters of States (see Section 5.2 for details).
- Control of extreme values in weights helps reduce instability of estimates to some extent, but there is a need for methods that do not introduce much bias. Following the famous suggestion of Hajek (1971) in his comments on Basu's fabled example of circus elephants, we performed ratio adjustment (a form of poststratification) to estimated totals obtained from the household data on the number of persons belonging to the pair domain of interest. This was implemented in a multivariate manner to get one set of final weights.
- In the absence of a suitable source of poststratification controls for the person pair-level weights and the household-level weights, the inherent two-phase nature of the survey design was capitalized upon to estimate these controls from the first phase of the large screener sample. The first-phase sample weight was poststratified to person-level U.S. Census Bureau counts to get more efficient estimated counts for pair and household data.
- The problem of multiplicities complicated the issue of providing one set of final weights. When dealing with person-level parameters involving drug-related behaviors among members of the same household, it is possible for an individual to manifest himself or herself in the pair sample through different pairs. To avoid overcounting, the pair weights have to be divided by multiplicity factors, which tend to be domain specific. For this reason, multiplicity factors for a key set of pair analysis domains also are produced along with a set of final calibrated pair weights.
- Missing items in the respondent questionnaire led to imputation for deriving pair relationships, multiplicity factors, and household counts for Hajek adjustments.

The calibration task described in this document has been in place, with minor modifications, since the 1999 version of NSDUH, which was then called the National Household Survey on Drug Abuse (NHSDA).[1] Results from this calibration applied to an earlier survey year were presented at the 2001 Joint Statistical Meetings. The procedures described in the proceedings papers from these presentations can serve as useful supplemental reference material on estimation in the presence of multiplicities and extreme weights (Chromy & Singh, 2001) and on GEM calibration of pair weights (Penne, Chen, & Singh, 2001). The experience of using GEM with person weights is described in an earlier proceedings paper (Chen, Penne, & Singh, 2000). This work was completed for the Substance Abuse and Mental Health Services Administration (SAMHSA), Center for Behavioral Health Statistics and Quality (CBHSQ), by RTI International (a trade name of Research Triangle Institute), Research Triangle Park, North Carolina, under Contract No. HHSS283200800004C. The authors would like to take this opportunity to thank a number of individuals for useful discussions and suggestions: Joe Gfroerer and Art Hughes of SAMHSA and Jim Chromy of RTI.

NSDUH Weighting Team
Ralph Folsom, Senior Advisor
Research Triangle Park, NC

[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.

Table of Contents

List of Tables

List of Exhibits

List of Terms and Abbreviations

DU	Dwelling unit.
Ev	Extreme value. See Sections 5.1 and 5.2 for more detail.
GEM	Generalized exponential model. See Chapter 3 for more detail.
Household-level person count	The number of pairs associated with a given domain in a given household. These counts are used as control totals in the poststratification step. See Chapter 10 in the imputation report (Frechtel et al., 2013) 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 Frechtel et al. (2013) for more detail.
Nr	Nonresponse.
Outwinsor	The proportion of weights trimmed after extreme value adjustment via winsorization.
Pair domain	A pair relationship where the target population is defined by one of the pair members, conditional on the attributes of the other pair member.
Pair relationship	The relationship between selected pair members.
Parent-child	A pair relationship where either both pair members identify the other as part of a parent-child relationship, or both pair members otherwise are determined to form a parent-child pair (either through other evidence or through imputation).
ps	Poststratification.
QDU	Questionnaire dwelling unit: a household where at least one member responded to the questionnaire.
res.pr.nr	Respondent pair nonresponse adjustment step. See Section 6.3.3 for more detail.
res.qdu.nr	Respondent questionnaire dwelling unit nonresponse adjustment step. See Section 6.2.3 for more detail.
res.pr.ev	Respondent pair extreme value adjustment step. See Section 6.3.5 for more detail.
res.qdu.ev	Respondent questionnaire dwelling unit extreme value adjustment step. See Section 6.2.5 for more detail.

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

1. Introduction

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

NSDUH allows for estimating characteristics at the person level, pair level, and household or QDU level. This report describes the weight calibration methods used for the pair- and QDU-level respondents. As described in the person-level report, NSDUH is an annual survey of about 67,500 persons selected from the civilian, noninstitutionalized population aged 12 or older from all 50 States and the District of Columbia. Unique to 2011 was a Gulf Coast Oversample (GCO) consisting of approximately 2,000 respondents in designated counties in Alabama, Florida, Louisiana, and Mississippi. This supplemental sample was used to study the impact of the April 2010 Deepwater Horizon oil spill on substance use and mental health and resulted in increasing the target national sample from 67,500 to 69,500. Based on a composite size measure, States were geographically partitioned into roughly equal-sized regions according to population. The 42 smaller States and the District of Columbia were partitioned into 12 State sampling (SS) regions, whereas the eight largest States were divided into 48 SS regions. Therefore, the partitioning of the United States resulted in the formation of a total of 900 SS regions. Under a stratified design with States serving as the primary strata and SS regions serving as the secondary strata, census tracts, segments within census tracts, and dwelling units (DUs) within segments were each selected using probability proportional to size sampling. NSDUH is sometimes referred to as a two-phase sample where the first phase consisted of a large number of screener dwelling units (SDUs, about 200,000) selected to ensure that various age groups (five in

[2] This report presents information from the 2011 National Survey on Drug Use and Health (NSDUH). Prior to 2002, the survey was called the National Household Survey on Drug Abuse (NHSDA).

all: 12 to 17, 18 to 25, 26 to 34, 35 to 49, and 50+) of eligible individuals were represented adequately in the second phase. Information collected from SDUs also provided estimates of population controls (as in two-phase sampling) for calibration at levels (such as pair and QDU) for which suitable U.S. Census Bureau-based controls were not available. The second phase consisted of the selection of zero, one, or two persons 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 persons for each of the 5 years (2007–2011). The distribution of pair data for different pairs of age groups may vary considerably (see Chapter 2 for details). It is seen that for certain age group domains, the realized sample size may not be sufficient to yield reliable estimates. Also, there may be problems of extreme weights due to small pair selection probabilities under Brewer's method that may cause instability of estimates. These and some other estimation issues related to pair data are discussed below, along with some adopted solutions.

Table 1.1 2007–2011 NSDUH Sample Sizes

Sample Unit		2007	2008	2009	2010	2011
SDU	Selected	158,377	160,114	161,377	166,532	179,293
	Completed	140,659	142,159	142,933	147,010	156,048
QDU	Selected	58,574	58,942	58,288	58,702	61,441
	Completed	47,709	48,180	48,088	48,113	50,133
Pair	Selected	26,696	26,769	26,497	26,295	27,095
	Completed	19,668	19,748	19,919	19,691	19,976
Person	Selected	85,270	85,711	84,785	84,997	88,536
	Completed	67,377	67,928	68,007	67,804	70,109

Note: The 2007–2010 sample sizes reflect the removal of falsified cases found in Pennsylvania and Maryland. The 2011 sample was not affected. For additional information, see Section B.3.5 in Appendix B of the *Results from the 2011 National Survey on Drug Use and Health: Summary of National Findings* (Center for Behavioral Health Statistics and Quality, 2012).

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

In estimation for pair domains, two major problems arise: one is that of multiplicities because, for a given domain defined by the pair relationship, when the parameter of interest is at

the person level, several pairs in the household could be associated with the same person, For example, analysts are interested in an outcome at the person level, the proportion of children who use drugs and whose parents report talking to them about drugs, where the focus is on the child in a parent-child pair. Several parent-child pairs in the household could be associated with the same child. If the household has two parents, the selected child has two inclusion possibilities (one with each parent) in the set of all such parent-child pairs (Frechtel et al., 2013). The other problem is that of extreme weights that may arise due to small selection probabilities for certain pair age groups, which may lead to unstable estimates. Each of these issues is discussed in turn.

If several pairs in the household are associated with the same person, it is necessary to use the average measure of behavior relationships for each member, which gives rise to multiplicities. Thus, the pair weights need to be divided by the person-level multiplicity factors for each domain of interest, and, therefore, multiplicity factors need to be produced along with the final set of calibrated weights. Because it is not straightforward to create these multiplicities, analyses would have to be necessarily limited to pair relationships where the multiplicities were produced a priori. It was anticipated that analyses of interest would be limited to 14 pair domains, listed in Table 1.2. Since 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, since the parameter of interest might be associated with only one member of the pair (the "focus" member), and the multiplicity would differ depending upon which pair member was the focus member.

Table 1.2 Pair Domains

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

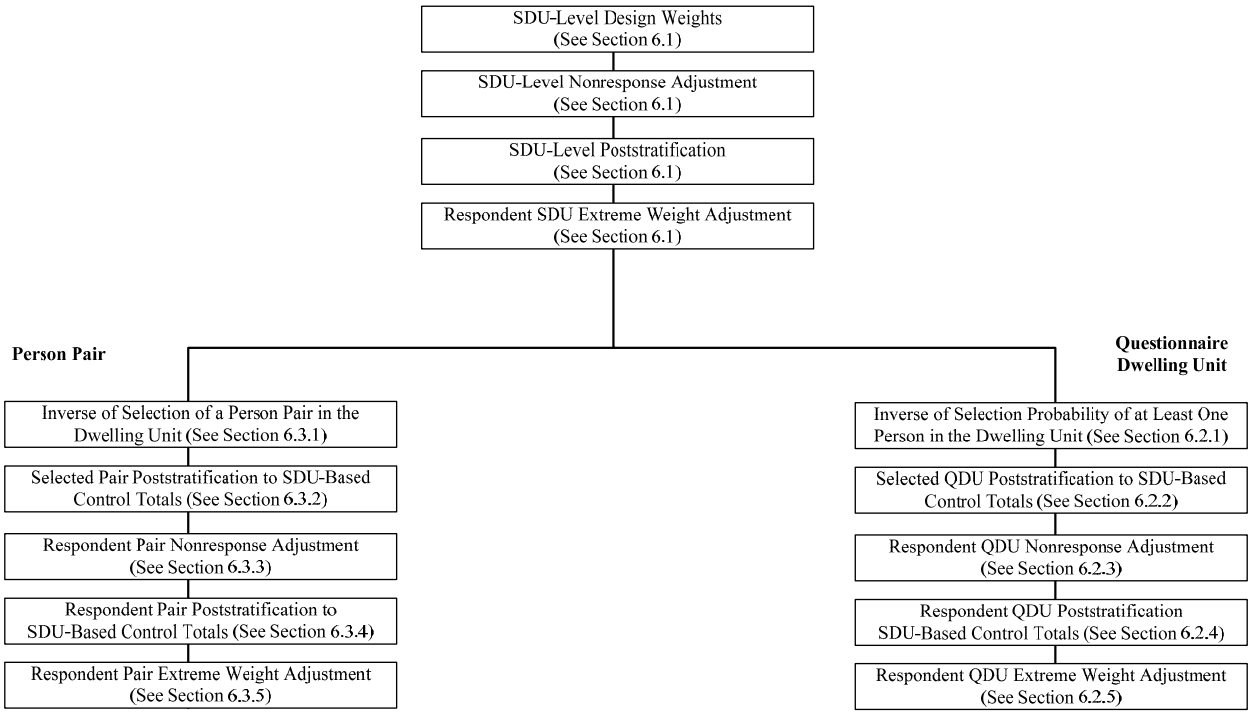
Some of the multiplicities, including counts of all possible pairs in a household for a given domain, were used for poststratification. Details are provided in Chapter 4. Prior editions of this report contained a chapter on editing and imputation of pair relationships, multiplicity factors, and household-level person counts for poststratification, but this information has been

removed in the interest of consolidating the imputation documentation and can now be found in the imputation report (Frechtel et al., 2013).

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 persons 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 persons selected from DUs). Note, however, that the multiplicity factor, being domain specific, renders the calibration adjustment factor domain specific. This raises the question of finding one set of calibration weights for use with all domains or outcome variables. To get around this problem, we performed a multivariate calibration with respect to a key set of pair domains. This type of poststratification then was followed by a repeat poststratification to further control the extreme weights by imposing separate bound restrictions on the initially identified extreme weights.

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

Exhibit 1.1 QDU and Pair Sampling Weight Calibration Steps



In dealing with extreme weights, it is assumed that they arise due to design (due to an imperfect frame, assignment of very small selection probabilities to some units, or a small weight adjustment factor after calibration) so that they make the sample representative of the population

and, hence, do not introduce bias. The only problem is that they may lead to highly unstable estimates similar to the problem of Basu's circus elephants[3] (Hajek, 1971). So, we need to perform some treatment (such as winsorization[4]) within suitably defined extreme weight domains such that these domains contain units possibly from different strata but with similar sample selection probabilities to avoid the occurrence of extreme weights due to a mix of different designs. The domains must be large enough (e.g., at least size 30) to be able to define extreme values according to the domain-specific weight distribution. Any extreme value treatment to increase precision of estimates would introduce some bias. However, this bias can be reduced considerably if the ev treatment is performed under calibration controls. This is what the built-in ev control in GEM tries to accomplish.

It follows that the definition of extreme weight domains should depend on factors that affect the selection probabilities of units in the sample, such as State- and age-specific sampling rates, segment selection probabilities, pair age-specific selection probabilities, and household composition. If one tries to define extreme weight domains by taking account of all these factors via cross-classification, it will lead to too many domains with insufficient observations. That is why it is difficult to define suitable extreme weight domains for pair data. In the case of person-level weights it was less difficult, since State by age group suitably captured the extreme weight domain requirements. The definition of extreme weight domains used in the 2011 survey was the same as the one used in the 1999–2010 surveys. The domains were defined as the cross-classification of State, pair age,[5] and number of persons 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 persons aged 12 to 25 were categorically defined as zero, one, and two or more. For more details, see Chapter 5.

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

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

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

2. Questionnaire Dwelling Unit and Pair Selection Probabilities

Similar to the 1999–2001 National Household Surveys on Drug Abuse (NHSDAs) and the 2002–2010 National Surveys on Drug Use and Health (NSDUHs),[6] the 2011 NSDUH had a two-phase design and used a computer-assisted interviewing (CAI) method. There were four stages of selection: census tracts, segments within census tracts, dwelling units (DUs) within segments, and persons within dwelling units. Any two survey eligible persons had some nonzero chance of being selected and, when both were selected, they formed a within household pair. This design feature is of interest to NSDUH researchers because, for example, it allows analysts to examine how the drug use propensity of an individual (in a family) relates to the drug use propensity of other members residing in the same dwelling unit (Morton, Martin, Shook-Sa, Chromy, & Hirsch, 2012).

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 persons 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 persons, then both persons 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 persons, no one was selected from that DU.

Starting with the 2002 NSDUH and continuing through 2011, the pair-sampling algorithm was modified to increase the number of pairs selected in the sample. Dwelling units with the sum of person selection probabilities greater than or equal to 2 were treated the same as in previous survey years. However, DUs where the sum of person-level selection probabilities was less than 2 received a slightly different treatment that increased the chance for selecting a pair of real persons. Section 2.1 describes the selection process for both types of DUs.

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

[6] This report presents information from the 2011 National Survey on Drug Use and Health (NSDUH). Prior to 2002, the survey was called the National Household Survey on Drug Abuse (NHSDA).

2.1 Pair Selection Probability

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

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

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

by setting K at

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

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

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

where age category 1 corresponds to persons aged 12 to 17, 2 to persons aged 18 to 25, 3 to persons aged 26 to 34, 4 to persons aged 35 to 49, and 5 to persons 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 the above scheme always selects a pair of two eligible persons.

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 to 2001 consisted of dividing $2 - S$ equally among the three dummy persons

added to the household, and then used Brewer's method to select a pair, as in Case I. However, if the household had two or more persons, we preferred a pair of real persons 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

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

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

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

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

by setting K' at

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

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

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

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

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

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

2.2 Questionnaire Dwelling Unit Selection Probability

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

The QDU selection probability was defined as

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

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

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

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

Table 2.1 Building Blocks of the QDU and Person Pair Samples: Dwelling Units and Persons in the 2007–2011 NSDUHs

	2007			2008			2009			2010			2011		
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	158,377	140,659	88.81	160,114	142,159	88.79	161,377	142,933	88.57	166,532	147,010	88.28	179,293	156,048	87.04
QDUs															
Total QDUs	58,574	47,709	81.45	58,942	48,180	81.74	58,288	48,088	82.50	58,702	48,113	81.96	61,441	50,133	81.60
Persons															
Total Persons	85,270	67,377	79.02	85,711	67,928	79.25	84,785	68,007	80.21	84,997	67,804	79.77	88,536	70,109	79.19
12-17	26,039	22,286	85.59	26,228	22,250	84.83	26,157	22,398	85.63	25,908	21,960	84.76	27,911	23,510	84.23
18-25	27,910	22,006	78.85	28,793	22,875	79.45	28,158	22,686	80.57	28,164	22,793	80.93	28,589	22,876	80.02
26-34	8,865	6,861	77.39	8,337	6,560	78.69	8,242	6,591	79.97	8,545	6,780	79.34	8,323	6,543	78.61
35-49	13,411	9,906	73.86	12,995	9,674	74.44	12,855	9,616	74.80	12,979	9,668	74.49	12,220	9,149	74.87
50+	9,045	6,318	69.85	9,358	6,569	70.20	9,373	6,716	71.65	9,401	6,603	70.24	11,493	8,031	69.88
Pairs															
Total Pairs[4]	26,696	19,668	73.67	26,769	19,748	73.77	26,497	19,919	75.17	26,295	19,691	74.88	27,095	19,976	73.73
0,0[5]	82,085	N/A	N/A	83,217	N/A	N/A	84,645	N/A	N/A	88,308	N/A	N/A	94,607	N/A	N/A
0, 12-17	8,416	7,966	94.65	8,634	8,094	93.75	8,432	7,936	94.12	8,595	7,906	91.98	9,402	8,651	92.01
0, 18-25	9,738	8,870	91.09	9,932	9,213	92.76	9,870	9,081	92.01	10,093	9,270	91.85	10,306	9,497	92.15
0, 26-34	3,972	3,412	85.90	3,771	3,303	87.59	3,798	3,318	87.36	3,914	3,418	87.33	3,930	3,386	86.16
0, 35-49	4,796	3,956	82.49	4,601	3,835	83.35	4,565	3,810	83.46	4,659	3,843	82.49	4,431	3,704	83.59
0, 50+	4,956	3,836	77.40	5,235	4,013	76.66	5,126	4,042	78.85	5,146	3,998	77.69	6,277	4,919	78.37
12-17, 12-17	4,347	3,691	84.91	4,319	3,654	84.60	4,288	3,648	85.07	4,160	3,525	84.74	4,649	3,885	83.57
12-17, 18-25	3,353	2,596	77.42	3,449	2,668	77.36	3,595	2,852	79.33	3,444	2,718	78.92	3,756	2,921	77.77
12-17, 26-34	917	743	81.03	897	690	76.92	872	724	83.03	922	752	81.56	834	685	82.13
12-17, 35-49	3,996	3,106	77.73	3,944	3,012	76.37	3,979	3,061	76.93	3,948	3,044	77.10	3,855	2,918	75.69
12-17, 50+	663	492	74.21	666	482	72.37	703	532	75.68	679	493	72.61	766	565	73.76
18-25, 18-25	5,602	4,082	72.87	5,872	4,261	72.56	5,588	4,168	74.59	5,502	4,165	75.70	5,476	4,015	73.32
18-25, 26-34	1,207	865	71.67	1,103	821	74.43	1,102	820	74.41	1,140	851	74.65	1,049	768	73.21
18-25, 35-49	1,666	1,058	63.51	1,706	1,112	65.18	1,576	1,059	67.20	1,639	1,098	66.99	1,469	994	67.67
18-25, 50+	742	454	61.19	859	546	63.56	839	547	65.20	844	537	63.63	1,057	666	63.01
26-34, 26-34	1,017	692	68.04	890	614	68.99	886	635	71.67	903	621	68.77	858	599	69.81
26-34, 35-49	517	344	66.54	545	380	69.72	447	312	69.80	512	354	69.14	492	314	63.82
26-34, 50+	218	113	51.83	241	145	60.17	251	149	59.36	251	162	64.54	302	192	63.58
35-49, 35-49	1,009	599	59.37	873	535	61.28	917	569	62.05	886	542	61.17	748	474	63.37
35-49, 50+	418	244	58.37	453	269	59.38	454	239	52.64	449	245	54.57	477	271	56.81
50+, 50+	1,024	589	57.52	952	559	58.72	1,000	604	60.40	1,016	584	57.48	1,307	709	54.25

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

Note: The 2007–2010 sample sizes reflect the removal of falsified cases found in Pennsylvania and Maryland. The 2011 sample was not affected. For additional information, see Section B.3.5 in Appendix B of the *Results from the 2011 National Survey on Drug Use and Health: Summary of National Findings* (Center for Behavioral Health Statistics and Quality, 2012).

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

[2] Respondent pairs are based on the questionnaire age and are comprised only of respondent persons.

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

[4] Total pairs excludes dummy person pairs.

[5] Among the completed screening dwelling units, no person was selected in this dwelling unit.

3. Brief Description of the Generalized Exponential Model

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

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

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

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

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

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

To implement GEM, several steps need to be followed: (1) define and create all the covariates; (2) define the extreme weights; (3) fit the GEM model. The details of practical aspects of GEM implementation can be found in Chapters 4 and 5 of this report and Chapter 4 of Chen et al. (2013).

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

We note that unlike the person-level weight calibration, the control totals for the questionnaire dwelling unit (QDU)-level and person pair-level poststratification are not available from the U.S. Census Bureau. A way around this problem is to take advantage of the two-phase nature of the design, in which the screener data provides 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 person-level weight calibration report (Chen et al., 2013).

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 generalized exponential model (GEM) were set to be common and to correspond to those used for the SDU nonresponse and poststratification adjustments. The variables are of two types: Those used for SDU nonresponse adjustment are 0/1 indicators, while those used for SDU poststratification adjustment are counting variables. The variables of the first type (0/1 indicators) are population density, group quarters, race/ethnicity of householder, percentage of persons in segment who are black or African American, percentage of persons in segment who are Hispanic or Latino, percentage of owner-occupied dwelling units (DUs) in segment, segment-combined median rent and housing value, and household type. Variables of the second type (counting variables) represent the number of eligible persons within each DU who fall into the various demographic categories of race, age group, Hispanicity, and gender. Note that the State and quarter variables are represented as both binary and counting variables. Thus, not only are DU counts within a specific State or quarter in the QDU sample controlled to the corresponding totals obtained from the SDU sample, but also counts of persons 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.

Exhibit 4.1 lists all predictor variables proposed for QDU-level calibration and identifies them as counting, binary, or both. Various main effects and higher level factor effects based on

the predictor variables were included in the GEM modeling. As stated previously, all adjustment steps at the QDU level used a common set of proposed predictor variables.

4.2 Pair Weight Calibration

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

In the poststratification of selected pairs and the nonresponse adjustment of respondent pairs, screener data were used in the definition of the pair-specific variables such as pair age, pair race/ethnicity, and pair gender, whereas in the poststratification and extreme weight adjustment of respondent pairs, these variables were obtained from the questionnaire. For the latter case, in addition to the variables described above, indicator covariates corresponding to selected pair domains were included to perform Hajek-type ratio adjustments via weight calibration, as mentioned in Chapter 1. The selected pair domains were limited to 10 of the 14 pair domains listed in Chapter 1. (Parent-child pairs where the child was in the 15- to 17-year-old age range and sibling-sibling-younger sibling focus pairs 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[b]
1: 12-17, 2: 18-25, 3: 26-34, 4: 35-49, 5: 50+[1]

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

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

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

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

Household Type (Ages of Persons Rostered within DU)[a]
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)[a]
1: 50-100%,[1] 2: 10->50%, 3: 0->10%

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

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

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

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

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

Race (5 Levels)[b]
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[a]
1: Hispanic or Latino White,[1] 2: Hispanic or Latino Black or African American, 3: Hispanic or Latino Other, 4: Non-Hispanic or Latino White, 5: Non-Hispanic or Latino Black or African American, 6: Non-Hispanic or Latino Other

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

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

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

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

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

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

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

Household Size
1: DU with 2 Persons,[1] 2: DU with 3 Persons, 3: DU with $\geq$ 4 Persons

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

State/Region

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

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

States[3]

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

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

Pair Relationship Associated with Multiplicity

1: Parent-Child (12-14)*
2: Parent-Child (12-17)*
3: Parent-Child (12-10)*
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; QDU = questionnaire dwelling unit.

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

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

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

* The pair member focused on.

5. Definition of Extreme Weights

An important feature of the generalized exponential model (GEM) is the built-in provision of extreme value (ev) treatment. Sampling weights are often classified as extreme (high or low) if they fall outside the interval, median ± 3 × interquartile range (IQR). The interval is set for prespecified domains defined usually by design variables corresponding to deep stratification.[7] Similar to previous NSDUHs, for the GEM modeling used in the 2011 NSDUH, a more conservative (narrower) interval was defined, median ± 2.5 × IQR. The narrower interval better prevents the adjusted weights from crossing the standard interval boundaries by treating weights near but not outside the commonly used boundaries (i.e., those that have the most potential to become extreme) as extreme as well.

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

5.1 Questionnaire Dwelling Unit Extreme Weight Definition

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

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

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

[7] Deep stratification refers to the stratification that was used in the sample design. In the case of the 2011 survey, deep stratification refers to the cross-classification of State sampling region by age group.

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

4. State sampling region by household type;

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

6. State sampling region by household type by household size.

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

5.2 Person Pair Extreme Weight Definition

The pair selection probability is a function of the selection probability of each person in the pair given by formula (2.1) or (2.6), depending on the sum of the person selection probabilities within the household as discussed in Section 2.1. This probability can be very small if the selection probabilities of individual members are small. For example, consider a particular selected dwelling unit (DU) from the 2011 survey. This DU gave rise to a selected pair of respondents, one aged between 35 and 49 and the other aged 50 or older. The selection probability in this DU was 0.14811 for a respondent aged between 35 and 49, and it was 0.08509 for a respondent aged 50 or older. Using the formula (2.6) in Chapter 2, the pair selection probability was computed to be 0.000487359. Therefore, the inverse of the selection probability, the pair-level design weight, was 2051.88. Thus, a small pair selection probability can create a high initial weight, which is the product of the screener dwelling unit (SDU) 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 1999–2011 surveys, the extreme weight definition was based on the following hierarchy of domains:

1. Pair age group[8] (with three age categories, 12 to 25, 26 to 49, and 50+) by number (0, 1, 2+) of persons aged 12 to 25 in the household;

2. State cluster (with five levels [explained below]) by pair age group by number (0, 1, 2+) of persons aged 12 to 25 in the household;

3. State cluster (with three levels [explained below]) by pair age group by number (0, 1, 2+) of persons aged 12 to 25 in the household; and

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

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

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

We now briefly introduce the considerations behind the above definition for extreme weight domains. The sample design prespecified the person-level selection probability within State by five age groups (12 to 17, 18 to 25, 26 to 34, 35 to 49, 50+). Age groups 12 to 17 and 18 to 25 have a relatively similar selection probability, and the same is true for age groups 26 to 34 and 35 to 49. The 50+ group, however, has a quite different selection probability from the other groups. Furthermore, since the 12 to 17 and 18 to 25 age groups have large selection probabilities, they have a very high chance of being selected if the household has persons in these age groups. Therefore, the number of persons 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 persons aged 12 to 25 in the household.

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

6. Weight Calibration at Questionnaire Dwelling Unit and Pair Levels

The 2011 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 basic sampling plan involved four stages of selection across two phases of design: within Phase I, (1) the selection of census tracts within each State sampling (SS) region, (2) the selection of subareas or segments (comprised of U.S. Census Bureau blocks) within SS regions; (3) the selection of dwelling units (DUs) within these subareas; and, finally, within Phase II, (4) the selection of eligible individuals within DUs. Specific details of the sample design and selection procedures for the sample can be found in the 2011 NSDUH sample design report (Morton, Martin, Shook-Sa, Chromy, & Hirsch, 2012).

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

The final sample weights for Phase I screener dwelling units (SDU) and Phase II QDU, person, and pair levels for the 2011 samples consisted of products of several factors, each representing either a probability of selection at some particular stage or some form of ev, nr, or ps calibration adjustment. In the following sections, we describe the QDU and pair weight components in greater detail. In summary, the first 10 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 2011 survey sample are the product of all weight components for each type of sample, illustrated in Exhibits 6.1 and 6.2.

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

For pair data, GEM modeling was initially applied by partitioning the pair data into four groups based on census regions. However, there were not enough observations in each group to fit a comprehensive model to reduce bias. Alternatively, a single model was attempted for the whole pair data, but it was rejected as not practical due to computational limitations. A compromise approach was adopted by combining census regions into two groups: Northeast with South and Midwest with West. This grouping proved both manageable and desirable as it

assisted in bias reduction, ease of modeling, and workload reduction. Exhibit 6.3 provides more details of the data partition for GEM modeling. The resulting sample sizes of selected and respondent units for the pair and QDU data partitions are shown for the 2007–2011 surveys in Table 6.1.

It may be noted that for the pair data in the 1999, 2000, and 2001 surveys, the built-in ev control feature of GEM was not used until the final respondent pair ev adjustment step. The reason for this is that the definition for ev domain was not finalized before the pair data calibration process was begun. However, for the 2002–2011 survey pair data, the built-in ev control feature was used for each adjustment step.

Exhibit 6.1 Summary of 2011 NSDUH QDU Sample Weight Components

Phase I Screener Dwelling Unit Level

Design Weight Components	
#1	Inverse Probability of Selecting Census Tract
#2	Inverse Probability of Selecting Segment
#3	Quarter Segment Weight Adjustment
#4	Subsegmentation Inflation Adjustment
#5	Inverse Probability of Selecting SDU
#6	Subsampling of Added SDU Adjustment
#7	SDU Release Adjustment
Weight Adjustment*	
#8	SDU Nonresponse Adjustment ***(res.sdu.nr)***
#9	SDU Poststratification Adjustment ***(res.sdu.ps)***
#10	SDU Extreme Value Adjustment ***(res.sdu.ev)***

Phase II Questionnaire Dwelling Unit Level

Design Weight Component	
#11	Inverse of Selection Probability of at Least One Person in the Dwelling Unit
Weight Adjustment*	
#12	Selected QDU Poststratification to SDU-Based Control Totals ***(sel.qdu.ps)***
#13	Respondent QDU Nonresponse Adjustment ***(res.qdu.nr)***
#14	Respondent QDU Poststratification to SDU-Based Control Totals ***(res.qdu.ps)***
#15	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 (Chen et al., 2013). For computational feasibility, all weight adjustments were done using the four model groups based on census regions defined in Exhibit 6.3.

Exhibit 6.2 Summary of 2011 NSDUH Person Pair Sample Weight Components

Phase I Screener Dwelling Unit Level

Design Weight Components
#1 Inverse Probability of Selecting Census Tract
#2 Inverse Probability of Selecting Segment
#3 Quarter Segment Weight Adjustment
#4 Subsegmentation Inflation Adjustment
#5 Inverse Probability of Selecting SDU
#6 Subsampling of Added SDU Adjustment
#7 SDU Release Adjustment
Weight Adjustment*
#8 SDU Nonresponse Adjustment ***(res.sdu.nr)***
#9 SDU Poststratification Adjustment ***(res.sdu.ps)***
#10 SDU Extreme Value Adjustment ***(res.sdu.ev)***

Phase II Person Pair Level

Design Weight Component
#11 Inverse of Selection Probability of a Person Pair in SDU
Weight Adjustment*
#12 Selected Pair Poststratification to SDU-Based Control Totals ***(sel.pr.ps)***
#13 Respondent Pair Nonresponse Adjustment ***(res.pr.nr)***
#14 Respondent Pair Poststratification Adjustment to SDU-Based Control Totals ***(res.per.ps)***
#15 Respondent Pair Extreme Value Adjustment ***(res.per.ev)***

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

* These adjustments use the generalized exponential model (GEM), which also involves pre- and post-processing in addition to running the GEM macro. See Exhibit 4.1 (Chen et al., 2013). For computational feasibility, all weight adjustments were done using the four model groups based on census regions defined in Exhibit 6.3.

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

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

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

Model Group	2007		2008		2009		2010		2011	
	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs
QDU										
Northeast	11,651	9,337	11,490	9,216	11,605	9,340	11,627	9,339	11,997	9,456
South	17,793	14,712	17,703	14,747	17,756	14,909	17,880	14,857	19,690	16,487
Midwest	16,652	13,590	16,763	13,650	16,382	13,491	16,670	13,686	17,045	13,752
West	12,478	10,070	12,986	10,567	12,545	10,348	12,525	10,231	12,709	10,438
Total	58,574	47,709	58,942	48,180	58,288	48,088	58,702	48,113	61,441	50,133
Model Group	**2007**		**2008**		**2009**		**2010**		**2011**	
	Selected Pairs	Completed Pairs	Selected Pairs	Completed Pairs	Selected Pairs	Completed Pairs	Selected Pairs	Completed Pairs	Selected Pairs	Completed Pairs
Pair										
Northeast + South	13,275	9,783	13,060	9,700	13,058	9,806	12,872	9,590	13,686	10,127
Midwest + West	13,421	9,885	13,709	10,048	13,439	10,113	13,423	10,101	13,409	9,849
Total	26,696	19,668	26,769	19,748	26,497	19,919	26,295	19,691	27,095	19,976

QDU = questionnaire dwelling unit.

Note: The 2007–2010 sample sizes reflect the removal of falsified cases found in Pennsylvania and Maryland. The 2011 sample was not affected. For additional information, see Section B.3.5 in Appendix B of the *Results from the 2011 National Survey on Drug Use and Health: Summary of National Findings* (Center for Behavioral Health Statistics and Quality, 2012).

6.1 Phase I SDU-Level Weight Components

A total of 10 weight components for the SDU level correspond to selection probabilities and nonresponse, poststratification, and extreme value adjustment factors. Note that this differs from previous NHSDAs and NSDUHs in that a new design-based component was incorporated at the beginning of the process so that corresponding weight component numbers are incremented by one when compared to previous survey years with an otherwise similar weighting scheme. The first seven components in the Phase I sample weights reflect the probability of selecting the DUs. These components were derived from (1) the probability of selecting the census tract within each State sampling (SS) region, (2) the probability of selecting the geographic segment within each SS region, (3) a quarter segment weight adjustment, (4) a subsegmentation inflation factor, (5) the probability of selecting a DU from within each counted and listed sampled segment, (6) the probability of inclusion of added DUs, and (7) DU percent release adjustment. The three remaining weight components, #8 through #10, are GEM calibration adjustments accounting for (8) DU nonresponse at the screening level, (9) DU poststratification to census controls, and (10) DU-level extreme value adjustment, although in 2009 extreme value adjustment at this stage was deemed unnecessary, and thus Weight Component #10 was set to one for all respondent DUs. The person-level, QDU-level, and person pair-level weights use the product of the above 10 weight components as the common initial weight before further adjustments. For more detailed information on Weight Components #1, #2, and #4 through #7, refer to the 2011 NSDUH sample design report (Morton et al., 2012), and for more detail on Weight Components #3 and #8 through #10, see the 2011 person-level sampling weight calibration report (Chen et al., 2013).

Note that from 2008 to 2011, there was an occasional second subsegmentation step when the initial partitioning of segments was insufficient due to 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 reproduced. Additional detail can be found in the 2011 NSDUH sample design report (Morton et al., 2012).

6.2 QDU Weight Components

6.2.1 QDU Weight Component #11: Inverse of Selection Probability of at Least One Person in the Dwelling Unit

The selection of a questionnaire dwelling unit from all completed SDUs is based on the outcome of a variant of Brewer's method, which may select zero, one, or two persons. Any pair of survey eligible residents within the dwelling unit had some known, nonzero chance of being selected for the survey. The value for Weight Component #11 is equal to the inverse of the probability that at least one person in the dwelling unit is selected (see Section 2.2 for details).

6.2.2 QDU Weight Component #12: Selected QDU Poststratification to SDU-Based Control Totals

This poststratification factor adjusts the weights for selected QDUs to the SDU-based control totals. The SDU-based control totals are obtained by using the calibrated SDU weights. This adjustment step provides more stable controls for the subsequent nonresponse adjustment (Weight Component #13). Exhibit 4.1 lists the initially proposed variables for GEM modeling. The predictor variables are either 0/1 indicators or counting variables representing the number of persons 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 persons, and so imputation was done to replace this missing data. For more details on the imputation of screener demographic information, see Chen et al. (2013).

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

6.2.3 QDU Weight Component #13: Respondent QDU Nonresponse Adjustment

This nonresponse adjustment step accounts for the failure to obtain respondent person(s) from each and every selected QDU. The same set of initially proposed predictor variables were used as for the previous adjustment (#12).

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

6.2.4 QDU Weight Component #14: Respondent QDU Poststratification to SDU-Based Control Totals

This final poststratification for all respondent QDUs utilized the same set of initially proposed predictor variables as previous adjustments. The corresponding control totals were obtained from the SDU-level sample, as was done for Weight Component #12.

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

6.2.5 QDU Weight Component #15: Respondent QDU Extreme Value Adjustment

The extreme weight proportions for the final poststratified weights were acceptably low, eliminating the need for the extreme value adjustment. Weight Component #15 was set to one for each responding QDU.

6.3 Pair-Level Weight Components

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

6.3.1 Pair Weight Component #11: Inverse of Selection Probability of a Person Pair in the Dwelling Unit

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

6.3.2 Pair Weight Component #12: Selected Pair Poststratification to SDU-Based Control Totals

Similar to QDU Weight Component #12, this step was motivated by the consideration that the larger sample of all possible pairs provides more stable control totals for the respondent pair nonresponse adjustment. The weights of selected pairs were poststratified to the control totals that derived from calibrated SDU weights of all possible pairs. The pair-level demographic variables for all selected pairs, such as pair age group, pair race/ethnicity, etc., were derived from screener demographic information.

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

6.3.3 Pair Weight Component #13: Respondent Pair Nonresponse Adjustment

If both persons 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. Due to 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 #12. Similar to Weight Component #12, the pair-level demographic variables for all selected pairs, such as pair age group, pair race/ethnicity, etc., were derived from screener demographic information.

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

6.3.4 Pair Weight Component #14: Respondent Pair Poststratification to SDU-Based Control Totals

This final poststratification utilized the same set of initially proposed predictor variables as previous adjustment steps. In addition, 10 pair relationship domain-level indicator variables were added to the set of covariates. The control totals for GEM calibration were derived from the SDU sample of all possible pairs of eligible persons, as was done for Weight Component #12. The calibration control totals for these 10 domains used household-level person counts and the final QDU weights. As mentioned in the introduction, use of these household-level count totals for pair relationship domains in GEM calibration provided Hajek-type weight adjustment in the interest of obtaining more stable estimates. In setting up calibration covariates, multiplicity factors were needed. These factors, as discussed in the introduction, are used in constructing estimates for person-level parameters based on pair-related drug behavior. The factors depend on the pair domains of interest. For a selected set of pair domains, multiplicity factors are provided along with the pair-level analysis weights. See Chapter 10 in the imputation report (Frechtel et al., 2013) for more detail on the creation of and imputation of missing values in the pair relationship, multiplicity, and household-level person counts. See Chapter 4 for more detail on the use of multiplicities and household-level person counts in poststratification.

Unlike Weight Components #12 and #13, demographic covariates were based on data from the questionnaire instead of information pulled from the dwelling unit screener.

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

6.3.5 Pair Weight Component #15: Respondent Pair Extreme Weight Adjustment

We checked the extreme weight proportions for the weights up to Weight Component #14, using the extreme weight domains (see Section 5.2). Even though the previous adjustment steps utilized the built-in extreme weight control feature of GEM, the extreme weight proportions were still high enough to cause concern that they might produce unreliable estimates. Therefore, the extreme weight adjustment via GEM was implemented, using the same final set of predictor variables kept in the model for Weight Component #15. This step was successful in reducing the extreme weight proportion in all model groups. For details, see Appendix J.

7. Evaluation of Calibration Weights

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

7.1 Response Rates

Table D.1 in Appendix D displays the final selected and responding questionnaire dwelling unit (QDU) sample sizes from the 2011 National Survey on Drug Use and Health (NSDUH) for various national domains. This table also shows the weighted response rates. Most domains reflect the overall 77.1 percent response rate, with most rates relatively close to 80 percent, although the highest response rate is 96.83 percent, from the Group level of the Group Quarters category. The lowest response rate came from the respondents in the Census Region Northeast, with 73.31 percent.

Table I.1 in Appendix I displays the final selected and responding pair-level sample sizes from the 2011 survey for various national domains. Due to 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 65.49 percent. The response rates range from a low of 44.32 percent in the pair race/ethnicity Hispanic or Latino and Other category to a high of 85.32 percent from the Pair Age Group 12-17,12-17. This extreme range of response rates is probably due to a combination of small sample sizes and response burden as a result of selection of pairs within households among various domains. Like at the QDU level, the top response rates are among the younger respondents (as measured by household type for the QDU data and pair age for the pair data). This pattern may be related to the relatively high response rates in the group level of the variable group quarters since it includes college dormitories.

7.2 Proportions of Extreme Values and Outwinsors

During the stages of modeling adjustments (i.e., nonresponse [nr] and poststratification [ps]), one major issue of concern when deciding the adequacy of a particular model was the extent of the resulting proportions of extreme value (ev) and outwinsor weights (see Sections 5.1 and 5.2 for these definitions). For each weight adjustment step, these proportions are computed before and after the step for various domains. Prior to adjustment, the product of all weight components is used to compute proportions of evs and outwinsors, while after the adjustment the product includes the new adjustment factor. If the proportion of evs and outwinsors are deemed high, a separate ev treatment step after ps could be performed. This was done for the pair-level weights. Details of this step are explained in Section 6.3.5. A separate ev treatment step was deemed unnecessary for the QDU-level weights.

Tables E.1 and E.2 and Tables J.1 through J.3 present percentages of evs at the QDU level and the pair level, respectively, for various domains. Unweighted percentages are the percentage of actual counts of units defined as evs relative to the total sample size. Weighted percentages reflect the percentage of total ev weights relative to the total sample weight, while

outwinsor percentages represent the total amount of residual weight when the weights are trimmed to the critical values (used for ev definition) relative to the total sample weight. For evaluation purposes, the outwinsor percentage is considered the most important of the three percentages, as this gave a measure of the impact of winsorization (or trimming) of ev weights (if we performed this treatment). See Sections 5.1 and 5.2 for the domains that were used to define extreme values.

7.3 Slippage Rates

The slippage rate for a given domain is defined as the relative percentage difference between the sampling weights and the external control totals, both before and after ps. The control totals for QDU and person pair ps are derived from the screener dwelling unit (SDU) weights, which were poststratified to U.S. Census Bureau population estimates (Chen et al., 2013). Table F.1 displays QDU national domain-specific weight sums for both before and after ps, as well as the desired totals to be met through ps. Table K.1 shows the same for the pair sample. These tables also show the relative percentage difference, or the amount of adjustment necessary (positive or negative) to meet the desired totals. The first relative difference is used explicitly during the ps modeling procedure to identify potential problems for convergence. Large differences in domains with relatively small sample sizes are indicative of potential large adjustment factors, which may cause problems in convergence while satisfying bound constraints. The reason is that adjustments required for one domain may have an adverse effect on another domain when a unit belongs to both.

As an example, consider that Table F.1, for the 2011 QDU domain household size of two, indicates a sample size of 17,500 with a total design-based weight of 54,645,502 and a census total of 54,660,454 with an initial slippage rate of 0.06 percent, which would imply a common weight adjustment approximately equal to 0.999358, if this were the only calibration control. Similarly, looking at pair data in Table K.1, the pair domain category of pair age 18-25, 18-25 has a sample size of 4,015, a design-based weight of 12,520,114, and a census total of 12,605,261, showing an initial slippage of -0.68 percent. The resultant required adjustment would be approximately equal to 1.006801, if this were the only control. However, in the generalized exponential model (GEM), all controls are simultaneously satisfied under a complex algorithm that allows for different adjustment factors for different units.

7.4 Weight Adjustment Summary Statistics

Tables G.1, G.2, and L.1 through L.3 display summary statistics on the product of weight components before and after all stages of adjustment for the QDU and person pair, respectively. The summary statistics include sample size (*n*), minimum (min), maximum (max), median (med), 25th percentile (Q1), 75th percentile (Q3), and the unequal weighting effect (UWE). Note that in Tables L.2 and L.3 the sample size for pair age group, pair race/ethnicity, and pair gender are slightly different. This is because those variables were defined using screening demographic information in the nonresponse adjustment of respondent pairs, while in the poststratification of respondent pairs, they were defined from questionnaire demographic information. Because UWE is directly affected by weight adjustment factors and extreme weights, these values—along with the percentage of extreme weights as noted in Section 7.2—were used as guidelines for determining model adequacy.

7.5 Sensitivity Analysis of Drug Use Estimates

It is known that, in general, there is a trade-off between bias reduction and variance reduction. For instance, with GEM (for nr or ps), enlarging a simple model (such as the one with only main effects) has the potential of further reducing the bias. At the same time, this enlargement also may be associated with a corresponding increase in the variance of the estimate due to additional variability caused by estimating the model parameters. To check for possible overfitting of the GEM model, we conducted a sensitivity analysis for respondent QDU poststratification for the QDU weights, respondent pair poststratification, and extreme weight adjustment for the person pair weights. A simple baseline model was fitted with the same bounds and maximum number of iterations as was used for the chosen (more complex) final model. We then looked for substantial changes in point estimates and standard errors (SEs). For the QDU weights, some household-level characteristics were selected such as family income, number of youths in the household, whether the household had health insurance coverage, and number of elders living in the household. The estimates and SEs are displayed in Table 7.1. For the person pair weights, selected licit and illicit drug use prevalence rates of 12- to 17-year-olds were calculated from parent-child pairs, and estimates and SEs of the estimates based on pair weights are shown in Tables 7.2a to 7.7b.

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

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

Table 7.1 Estimates of Totals and SEs for Domains of Interest Based on QDU Sample: 2011

Domain	2011				
	n	**Baseline (B)[1]**	**Final (F)[2]**	**(B-F)/F% (Estimate)**	**(B-F)/F% (SE)**
Households with Family Income					
$0 - < $10,000	5,264	10,225,504 (300,689)	10,220,071 (300,898)	0.05	-0.07
$10,000 - < $20,000	6,630	16,566,075 (411,979)	16,571,695 (412,869)	-0.03	-0.22
$20,000 - < $30,000	5,987	14,143,283 (348,921)	14,139,015 (348,740)	0.03	0.05
$30,000 - < $40,000	5,475	13,222,482 (342,331)	13,219,654 (343,080)	0.02	-0.22
$40,000 - < $50,000	5,265	12,686,674 (325,169)	12,686,431 (325,769)	0.00	-0.18
$50,000 - < $75,000	8,007	19,520,661 (415,263)	19,513,790 (415,364)	0.04	-0.02
$75,000 - < $100,000	5,474	12,789,005 (314,564)	12,791,799 (315,589)	-0.02	-0.32
$100,000+	8,031	19,785,122 (506,217)	19,796,351 (506,881)	-0.06	-0.13
Households with Number of Youths (< 18)					
0	20,823	77,445,018 (1,049,473)	77,430,006 (1,048,756)	0.02	0.07
1	11,954	17,559,778 (306,996)	17,568,957 (308,567)	-0.05	-0.51
2	10,135	15,097,704 (270,169)	15,110,440 (272,313)	-0.08	-0.79
3	4,736	6,103,358 (149,291)	6,100,784 (149,966)	0.04	-0.45
4+	2,485	2,732,948 (86,052)	2,728,618 (86,241)	0.16	-0.22
Households with Insurance Coverage					
Yes	40,226	98,762,981 (1,122,416)	98,742,418 (1,122,243)	0.02	0.02
No	9,907	20,175,824 (397,525)	20,196,387 (398,921)	-0.10	-0.35
Households with Number of Older Adults (65+)					
0	44,790	90,824,050 (1,015,817)	90,834,541 (1,015,753)	-0.01	0.01
1	3,593	18,305,226 (525,149)	18,286,936 (525,227)	0.10	-0.01
2	1,715	9,696,104 (335,980)	9,704,055 (336,373)	-0.08	-0.12
3+	35	113,426 (26,964)	113,274 (27,043)	0.13	-0.29

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

Table 7.2a Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Alcohol and Tobacco among Mother-Child (12 to 17) Pairs, by Mother Use: 2011

		2011		
Drug	**Mother User**	***n***	**Baseline[1]**	**Final[2]**
Alcohol				
Lifetime	Yes	2,031	36.79 (2.02)	36.53 (2.05)
	No	278	19.17 (3.65)	18.70 (3.65)
	Overall	2,309	34.10 (1.82)	33.79 (1.84)
Past Year	Yes	1,664	31.65 (2.09)	31.10 (2.07)
	No	645	17.54 (3.09)	17.25 (3.14)
	Overall	2,309	27.11 (1.77)	26.65 (1.77)
Past Month	Yes	1,212	17.85 (2.07)	17.88 (2.07)
	No	1,097	8.92 (1.57)	8.66 (1.60)
	Overall	2,309	13.30 (1.31)	13.18 (1.32)
Cigarettes				
Lifetime	Yes	1,531	21.20 (1.94)	21.41 (2.02)
	No	778	14.65 (2.54)	15.06 (2.66)
	Overall	2,309	18.73 (1.53)	19.03 (1.59)
Past Year	Yes	675	21.64 (3.03)	22.20 (3.21)
	No	1,634	10.68 (1.44)	10.76 (1.49)
	Overall	2,309	13.29 (1.34)	13.53 (1.41)
Past Month	Yes	604	15.13 (3.04)	15.75 (3.26)
	No	1,705	6.40 (1.13)	6.50 (1.16)
	Overall	2,309	8.27 (1.11)	8.53 (1.17)

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

Table 7.2b Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Alcohol and Tobacco among Father-Child (12 to 17) Pairs, by Father Use: 2011

		2011		
Drug	**Father User**	***n***	**Baseline[1]**	**Final[2]**
Alcohol				
Lifetime	Yes	1,448	32.69 (2.09)	32.17 (2.10)
	No	85	21.47 (7.73)	19.14 (6.73)
	Overall	1,533	31.84 (2.04)	31.17 (2.04)
Past Year	Yes	1,197	29.17 (2.21)	28.93 (2.24)
	No	336	16.02 (3.55)	15.44 (3.33)
	Overall	1,533	25.88 (1.91)	25.55 (1.91)
Past Month	Yes	998	13.05 (1.98)	12.66 (1.91)
	No	535	6.65 (1.54)	6.95 (1.79)
	Overall	1,533	10.79 (1.41)	10.65 (1.40)
Cigarettes				
Lifetime	Yes	1,129	15.61 (1.71)	15.34 (1.68)
	No	404	8.09 (1.94)	7.93 (1.91)
	Overall	1,533	13.36 (1.34)	13.13 (1.32)
Past Year	Yes	406	15.58 (2.83)	15.59 (2.83)
	No	1,127	6.26 (1.04)	6.18 (1.02)
	Overall	1,533	8.31 (1.03)	8.22 (1.02)
Past Month	Yes	354	10.98 (2.73)	11.01 (2.67)
	No	1,179	3.18 (0.66)	3.17 (0.65)
	Overall	1,533	4.59 (0.75)	4.56 (0.73)

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

Table 7.3a Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Any Illicit Drug or Marijuana among Mother-Child (12 to 17) Pairs, by Mother Use: 2011

		2011		
Drug	**Mother User**	***n***	**Baseline[1]**	**Final[2]**
Any Illicit				
Lifetime	Yes	1,232	31.83 (2.51)	32.00 (2.58)
	No	1,077	16.69 (1.81)	16.98 (1.89)
	Overall	2,309	24.32 (1.64)	24.59 (1.69)
Past Year	Yes	232	27.88 (4.63)	28.07 (4.71)
	No	2,077	16.92 (1.41)	17.07 (1.45)
	Overall	2,309	17.88 (1.35)	18.05 (1.38)
Past Month	Yes	135	13.51 (4.23)	13.56 (4.34)
	No	2,174	8.82 (1.06)	8.99 (1.10)
	Overall	2,309	9.05 (1.02)	9.21 (1.07)
Marijuana				
Lifetime	Yes	1,095	25.10 (2.55)	25.39 (2.63)
	No	1,214	10.69 (1.40)	10.96 (1.46)
	Overall	2,309	17.09 (1.45)	17.42 (1.49)
Past Year	Yes	150	23.24 (5.27)	22.91 (5.36)
	No	2,159	12.76 (1.24)	12.99 (1.29)
	Overall	2,309	13.27 (1.21)	13.49 (1.26)
Past Month	Yes	98	10.74 (4.46)	10.68 (4.58)
	No	2,211	7.10 (0.98)	7.29 (1.03)
	Overall	2,309	7.22 (0.96)	7.41 (1.01)

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

Table 7.3b Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Any Illicit Drug or Marijuana among Father-Child (12 to 17) Pairs, by Father Use: 2011

		2011		
Drug	**Father User**	***n***	**Baseline[1]**	**Final[2]**
Any Illicit				
Lifetime	Yes	918	26.11 (2.35)	24.95 (2.28)
	No	615	14.53 (2.02)	14.58 (2.05)
	Overall	1,533	21.27 (1.68)	20.58 (1.63)
Past Year	Yes	187	31.43 (5.45)	30.33 (5.29)
	No	1,346	12.10 (1.49)	11.63 (1.40)
	Overall	1,533	14.24 (1.48)	13.63 (1.40)
Past Month	Yes	104	17.88 (5.80)	19.55 (6.47)
	No	1,429	6.95 (1.25)	6.50 (1.16)
	Overall	1,533	7.65 (1.22)	7.29 (1.16)
Marijuana				
Lifetime	Yes	848	19.15 (2.29)	18.17 (2.18)
	No	685	7.48 (1.48)	7.44 (1.47)
	Overall	1,533	13.76 (1.45)	13.17 (1.37)
Past Year	Yes	142	18.70 (4.50)	17.25 (4.08)
	No	1,391	9.82 (1.41)	9.34 (1.32)
	Overall	1,533	10.53 (1.35)	9.95 (1.26)
Past Month	Yes	89	8.15 (3.59)	7.87 (3.53)
	No	1,444	4.79 (1.05)	4.38 (0.94)
	Overall	1,533	4.96 (1.01)	4.55 (0.92)

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

Table 7.4 Percentages of Youths (12 to 17) Living with a Parent Reporting Lifetime, Past Year, and Past Month Use of Alcohol and Tobacco among Parent-Child (12 to 17) Pairs, Asked Whether Their Parents Had Spoken to Them about the Dangers of Tobacco, Alcohol, or Drug Use within the Past 12 Months: 2011

Drug	Parent Talked about Dangers with Child	2011		
		n	Baseline[1]	Final[2]
Alcohol				
Lifetime	Yes	2,237	33.00 (2.03)	32.48 (2.04)
	No	1,552	34.63 (2.13)	34.20 (2.12)
	Overall	3,789	33.69 (1.48)	33.22 (1.49)
Past Year	Yes	2,237	27.66 (1.99)	27.15 (1.98)
	No	1,552	26.53 (2.01)	26.11 (1.98)
	Overall	3,789	27.17 (1.44)	26.70 (1.43)
Past Month	Yes	2,237	12.49 (1.33)	12.26 (1.33)
	No	1,552	12.28 (1.47)	12.29 (1.47)
	Overall	3,789	12.40 (0.99)	12.27 (0.99)
Cigarettes				
Lifetime	Yes	2,237	16.37 (1.53)	16.22 (1.57)
	No	1,552	18.20 (1.72)	18.57 (1.78)
	Overall	3,789	17.16 (1.13)	17.22 (1.16)
Past Year	Yes	2,237	11.61 (1.28)	11.52 (1.30)
	No	1,552	11.68 (1.37)	12.00 (1.44)
	Overall	3,789	11.64 (0.94)	11.72 (0.96)
Past Month	Yes	2,237	6.92 (1.04)	6.98 (1.07)
	No	1,552	7.22 (1.14)	7.49 (1.20)
	Overall	3,789	7.05 (0.76)	7.20 (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.

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: 2011

Drug	Parent Talked about Dangers with Child	2011		
		n	Baseline[1]	Final[2]
Any Illicit				
Lifetime	Yes	2,237	21.87 (1.57)	21.27 (1.56)
	No	1,552	25.09 (1.87)	25.49 (1.93)
	Overall	3,789	23.25 (1.22)	23.08 (1.24)
Past Year	Yes	2,237	16.21 (1.37)	15.82 (1.36)
	No	1,552	17.40 (1.63)	17.60 (1.67)
	Overall	3,789	16.72 (1.06)	16.58 (1.07)
Past Month	Yes	2,237	8.14 (1.09)	7.99 (1.09)
	No	1,552	9.63 (1.21)	9.70 (1.22)
	Overall	3,789	8.77 (0.81)	8.72 (0.81)
Marijuana				
Lifetime	Yes	2,237	15.93 (1.45)	15.60 (1.43)
	No	1,552	16.65 (1.62)	16.89 (1.67)
	Overall	3,789	16.24 (1.09)	16.15 (1.10)
Past Year	Yes	2,237	12.96 (1.28)	12.71 (1.28)
	No	1,552	12.26 (1.42)	12.33 (1.45)
	Overall	3,789	12.66 (0.96)	12.55 (0.97)
Past Month	Yes	2,237	6.37 (1.01)	6.25 (1.01)
	No	1,552	6.95 (1.11)	6.98 (1.11)
	Overall	3,789	6.62 (0.74)	6.56 (0.73)

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

Table 7.6a Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Alcohol and Tobacco among Mother-Child (12 to 17) Pairs, for Mother in the Pair, Asked Whether She Had Spoken to Her Children about the Dangers of Tobacco, Alcohol, or Drug Use within the Past 12 Months: 2011

Drug	Mother Talked about Dangers with Child	2011		
		n	Baseline[1]	Final[2]
Alcohol				
Lifetime	0 times	161	23.25 (4.96)	22.49 (4.75)
	1-2 times	319	37.01 (5.25)	35.62 (5.15)
	A few times	584	33.37 (4.07)	32.79 (4.09)
	Many times	1,141	35.10 (2.45)	35.35 (2.53)
	Overall	2,205	33.90 (1.87)	33.59 (1.89)
Past Year	0 times	161	18.59 (4.71)	17.62 (4.39)
	1-2 times	319	29.88 (5.22)	28.82 (5.07)
	A few times	584	28.64 (4.10)	28.21 (4.09)
	Many times	1,141	26.38 (2.22)	26.14 (2.28)
	Overall	2,205	26.87 (1.81)	26.42 (1.80)
Past Month	0 times	161	7.06 (3.14)	6.26 (2.58)
	1-2 times	319	16.68 (4.72)	16.74 (4.75)
	A few times	584	13.37 (2.75)	13.11 (2.73)
	Many times	1,141	13.19 (1.75)	13.13 (1.80)
	Overall	2,205	13.25 (1.34)	13.11 (1.35)
Cigarettes				
Lifetime	0 times	161	5.10 (1.84)	5.30 (1.93)
	1-2 times	319	15.18 (4.52)	15.16 (4.55)
	A few times	584	18.30 (3.30)	18.57 (3.38)
	Many times	1,141	22.04 (2.26)	22.45 (2.40)
	Overall	2,205	18.48 (1.57)	18.79 (1.64)
Past Year	0 times	161	3.08 (1.29)	3.24 (1.33)
	1-2 times	319	7.18 (2.20)	7.36 (2.23)
	A few times	584	12.18 (2.84)	12.17 (2.85)
	Many times	1,141	16.81 (2.17)	17.21 (2.31)
	Overall	2,205	12.85 (1.38)	13.10 (1.45)
Past Month	0 times	161	1.42 (0.82)	1.56 (0.89)
	1-2 times	319	5.70 (2.07)	5.94 (2.11)
	A few times	584	8.06 (2.62)	8.07 (2.61)
	Many times	1,141	10.15 (1.73)	10.52 (1.87)
	Overall	2,205	8.13 (1.15)	8.37 (1.20)

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

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: 2011

Drug	Father Talked about Dangers with Child	2011		
		n	Baseline[1]	Final[2]
Alcohol				
Lifetime	0 times	170	31.88 (6.68)	30.85 (6.70)
	1-2 times	263	34.16 (5.21)	34.10 (5.26)
	A few times	482	31.65 (3.48)	30.57 (3.38)
	Many times	508	32.91 (3.68)	32.40 (3.72)
	Overall	1,423	32.59 (2.13)	31.90 (2.13)
Past Year	0 times	170	23.61 (5.95)	23.72 (6.06)
	1-2 times	263	26.66 (4.79)	26.80 (4.88)
	A few times	482	25.88 (3.37)	24.72 (3.20)
	Many times	508	27.98 (3.58)	27.97 (3.63)
	Overall	1,423	26.44 (2.01)	26.10 (2.00)
Past Month	0 times	170	11.44 (4.27)	11.00 (4.18)
	1-2 times	263	11.46 (3.57)	11.76 (3.89)
	A few times	482	7.76 (1.87)	7.66 (1.76)
	Many times	508	12.89 (2.78)	12.58 (2.71)
	Overall	1,423	10.76 (1.47)	10.61 (1.46)
Cigarettes				
Lifetime	0 times	170	16.24 (4.50)	16.11 (4.52)
	1-2 times	263	10.04 (2.60)	9.67 (2.48)
	A few times	482	10.98 (2.17)	10.71 (2.08)
	Many times	508	16.00 (2.48)	15.87 (2.48)
	Overall	1,423	13.31 (1.40)	13.10 (1.39)
Past Year	0 times	170	6.92 (2.28)	7.17 (2.39)
	1-2 times	263	6.17 (1.66)	6.09 (1.67)
	A few times	482	6.88 (1.70)	6.73 (1.59)
	Many times	508	10.89 (2.23)	10.77 (2.21)
	Overall	1,423	8.16 (1.07)	8.10 (1.06)
Past Month	0 times	170	3.84 (1.76)	3.93 (1.83)
	1-2 times	263	3.84 (1.34)	3.90 (1.40)
	A few times	482	4.11 (1.47)	4.10 (1.39)
	Many times	508	5.19 (1.33)	5.05 (1.30)
	Overall	1,423	4.40 (0.76)	4.38 (0.75)

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

Table 7.7a Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Any Illicit Drug and Marijuana among Mother-Child (12 to 17) Pairs, for Mother in the Pair, Asked Whether She Had Spoken to Her Child about the Dangers of Tobacco, Alcohol, or Drug Use within the Past 12 Months: 2011

Drug	Mother Talked about Dangers with Child	2011		
		n	Baseline[1]	Final[2]
Any Illicit				
Lifetime	0 times	161	16.51 (3.56)	17.44 (3.74)
	1-2 times	319	18.20 (3.63)	18.08 (3.61)
	A few times	584	22.27 (3.36)	22.08 (3.38)
	Many times	1,141	28.81 (2.50)	29.26 (2.61)
	Overall	2,205	24.29 (1.68)	24.53 (1.73)
Past Year	0 times	161	13.15 (3.27)	13.88 (3.45)
	1-2 times	319	13.80 (3.26)	13.91 (3.26)
	A few times	584	16.91 (2.64)	16.97 (2.71)
	Many times	1,141	20.20 (2.10)	20.26 (2.18)
	Overall	2,205	17.69 (1.37)	17.82 (1.41)
Past Month	0 times	161	6.84 (2.33)	7.38 (2.49)
	1-2 times	319	6.54 (2.14)	6.91 (2.22)
	A few times	584	8.33 (1.75)	8.44 (1.81)
	Many times	1,141	10.48 (1.73)	10.49 (1.81)
	Overall	2,205	8.96 (1.04)	9.10 (1.09)
Marijuana				
Lifetime	0 times	161	8.27 (2.61)	8.92 (2.80)
	1-2 times	319	11.69 (3.00)	12.06 (3.09)
	A few times	584	15.06 (2.78)	15.02 (2.79)
	Many times	1,141	21.44 (2.32)	21.84 (2.41)
	Overall	2,205	17.03 (1.48)	17.33 (1.53)
Past Year	0 times	161	5.84 (2.00)	6.28 (2.15)
	1-2 times	319	9.82 (2.75)	10.34 (2.87)
	A few times	584	11.56 (2.12)	11.59 (2.18)
	Many times	1,141	16.37 (1.97)	16.51 (2.05)
	Overall	2,205	13.12 (1.24)	13.32 (1.28)
Past Month	0 times	161	2.88 (1.50)	3.22 (1.66)
	1-2 times	319	5.96 (2.11)	6.31 (2.18)
	A few times	584	6.64 (1.64)	6.81 (1.71)
	Many times	1,141	8.54 (1.65)	8.63 (1.74)
	Overall	2,205	7.13 (0.98)	7.30 (1.03)

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

Table 7.7b Percentages of Youths (12 to 17) Reporting Lifetime, Past Year, and Past Month Use of Any Illicit Drug and Marijuana among Father-Child (12 to 17) Pairs, for Father in the Pair, Asked Whether He Had Spoken to His Child about the Dangers of Tobacco, Alcohol, or Drug Use within the Past 12 Months: 2011

Drug	Father Talked about Dangers with Child	2011		
		n	Baseline[1]	Final[2]
Any Illicit				
Lifetime	0 times	170	17.60 (4.38)	16.34 (3.95)
	1-2 times	263	27.19 (4.42)	26.48 (4.39)
	A few times	482	16.09 (2.85)	15.66 (2.75)
	Many times	508	23.19 (3.08)	22.65 (3.03)
	Overall	1,423	20.84 (1.76)	20.20 (1.71)
Past Year	0 times	170	9.30 (3.41)	8.03 (2.61)
	1-2 times	263	12.83 (3.01)	12.20 (2.89)
	A few times	482	12.82 (2.64)	12.26 (2.49)
	Many times	508	17.70 (2.91)	17.23 (2.82)
	Overall	1,423	14.04 (1.56)	13.40 (1.48)
Past Month	0 times	170	3.11 (1.38)	3.03 (1.38)
	1-2 times	263	8.40 (2.85)	7.64 (2.71)
	A few times	482	4.68 (1.81)	4.59 (1.73)
	Many times	508	12.16 (2.77)	11.64 (2.64)
	Overall	1,423	7.77 (1.30)	7.41 (1.24)
Marijuana				
Lifetime	0 times	170	6.96 (2.61)	7.03 (2.74)
	1-2 times	263	12.68 (3.36)	11.82 (3.15)
	A few times	482	13.29 (2.72)	12.70 (2.58)
	Many times	508	16.34 (2.82)	15.61 (2.68)
	Overall	1,423	13.36 (1.53)	12.76 (1.45)
Past Year	0 times	170	3.56 (1.68)	3.55 (1.72)
	1-2 times	263	9.87 (2.89)	9.26 (2.74)
	A few times	482	10.79 (2.50)	10.08 (2.29)
	Many times	508	12.80 (2.70)	12.04 (2.52)
	Overall	1,423	10.31 (1.42)	9.70 (1.32)
Past Month	0 times	170	1.78 (1.14)	1.79 (1.19)
	1-2 times	263	4.19 (1.69)	3.39 (1.32)
	A few times	482	3.54 (1.59)	3.34 (1.43)
	Many times	508	7.83 (2.39)	7.20 (2.20)
	Overall	1,423	4.92 (1.07)	4.49 (0.97)

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

References

Center for Behavioral Health Statistics and Quality. (2012). *Results from the 2011 National Survey on Drug Use and Health: Summary of national findings* (HHS Publication No. SMA 12-4713, NSDUH Series H-44). Rockville, MD: Substance Abuse and Mental Health Services Administration. Available at http://www.samhsa.gov/data/NSDUH.aspx

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

Chen, P., Cribb, D., Dai, L., Gordek, H., Laufenberg, J., Sathe, N., & Westlake, M. (2013). Person-level sampling weight calibration. In *2011 National Survey on Drug Use and Health: Methodological resource book* (Section 12, prepared for the Substance Abuse and Mental Health Services Administration, Contract No. HHSS283200800004C, Phase II, Deliverable No. 39, RTI/0211838.207.004). Research Triangle Park, NC: RTI International.

Chen, P., Penne, M. A., & Singh, A. C. (2000). Experience with the generalized exponential model (GEM) for weight calibration for NHSDA. In *Proceedings of the 2000 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Indianapolis, IN* (pp. 604-607). Alexandria, VA: American Statistical Association. [Available as a PDF at http://www.amstat.org/sections/srms/proceedings/]

Chromy, J. R., & Singh, A. C. (2001). Estimation for person-pair drug-related characteristics in the presence of pair multiplicities and extreme sampling weights. In *Proceedings of the 2001 Joint Statistical Meetings*, *American Statistical Association, Survey Research Methods Section, Atlanta, GA* [CD-ROM]. Alexandria, VA: American Statistical Association. [Available as a PDF at http://www.amstat.org/sections/srms/proceedings/]

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

Folsom, R. E., & Singh, A. C. (2000). The generalized exponential model for sampling weight calibration for extreme values, nonresponse, and poststratification. In *Proceedings of the 2000 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Indianapolis, IN* (pp. 598-603). Alexandria, VA: American Statistical Association. [Available as a PDF at http://www.amstat.org/sections/srms/proceedings/]

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

Frechtel, P., Archambault, H., Carpenter, L., Cummiskey, C., Edwards, S., Laufenberg, J., Martin, P., Moore, A., & Scott, V. (2013). Imputation report. In *2011 National Survey on Drug Use and Health: Methodological resource book* (Section 11, prepared for the Substance Abuse and Mental Health Services Administration, Contract No. HHSS283200800004C, Deliverable No. 39, RTI/0211838.207.006.007). Research Triangle Park, NC: RTI International.

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

Morton, K. B., Martin, P. C., Shook-Sa, B. E., Chromy, J. R., & Hirsch, E. L. (2012). Sample design report. In *2011 National Survey on Drug Use and Health: Methodological resource book* (Section 2, prepared for the Substance Abuse and Mental Health Services Administration, Contract No. HHSS283200800004C, Phase I, Deliverable No. 8, RTI/0211838.203.004). Research Triangle Park, NC: RTI International.

Penne, M., Chen, P., & Singh, A.C. (2001) Person-pair sampling weight calibration using the generalized exponential model for the National Household Survey on Drug Abuse. In *Proceedings of the 2001 Joint Statistical Meetings*, *American Statistical Association, Survey Research Methods Section, Atlanta, GA* [CD-ROM]. Alexandria, VA: American Statistical Association. [Available as a PDF at http://www.amstat.org/sections/srms/proceedings/]

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

Appendix A: Technical Details about the Generalized Exponential Model

Appendix A: Technical Details about the Generalized Exponential Model

A.1 Distance Function

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

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

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

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

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

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

Note that the number of parameters in the GEM should be $\le n$, 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 usual raking ratio method (Singh & Mohl, 1996) of weight adjustment is a special case of the GEM, noting that for $\ell_k=0$, $u_k=\infty$, $c_k=1$, and $k=1,\ldots,n$, we have

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

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

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

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

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

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

For the poststratification adjustment, the following weights are set: for nonextreme weights, $\ell_k = \ell_2, \; c_k = c_2 = 1,$ and $u_k = u_2;$ for high extreme weights, $\ell_k = \ell_1 m_k, c_k = m_k,$ and $u_k = u_1 m_k;$ and similarly, for low extreme weights, $\ell_k = \ell_3 m_k, c_k = m_k,$ and $u_k = u_3 m_k.$ The extreme value adjustment is identical to poststratifcation, 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 $\lambda^{(0)} = 1$.

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 or not. If not, a half step[1] is used in the iteration increment.

A.4 Scaled Constrained Exponential Model

In National Household Surveys on Drug Abuse (NHSDAs)[2] prior to 1999, constrained exponential models (CEMs) were used for poststratification, and scaled CEMs were used for nonresponse adjustments. The CEM refers to the logit model of Deville and Särndal (1992), in which lower and upper bounds do not vary with k; that is, $\ell_k = \ell$, $u_k = u$, and $c_k = c = 1$, such that $\ell < 1 < u$. Thus, the CEM is a special case of the GEM. For the nonresponse adjustment, Folsom and Witt (1994) modified the CEM estimating equations by a scaling factor (ρ^{-1}, the inverse of the overall response propensity), such that $1 < \rho^{-1}a_k < \rho^{-1}u$. This implies that choosing ℓ in the CEM as ρ ensures that the scaled adjustment factor for nonresponse is at least one.

[1] A half step refers to halving the increment in the Newton-Raphson iterative process for fitting GEM.

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

Appendix B: Derivation of Poststratification Control Totals

Appendix B: Derivation of Poststratification Control Totals

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

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

B.1 Derivation of QDU-Level Poststratification Controls

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

B.1.1 Person Level

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

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

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

B.1.1.2 Adjusting SDU Data to the Control Totals

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

B.1.2 QDU Level

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

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

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

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

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

B.2 Derivation of Person Pair-Level Poststratification Controls

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

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

For most pair data domains—those other than the 10 pair domains based on relationship—the control totals for the poststratification adjustments were obtained from SDU data and were based on the number of possible pairs within SDUs. In order to obtain these pair counts belonging to various sociodemographic domains, the screener roster information was used to calculate all possible pairs within SDUs. For example, consider an SDU with two persons aged 12 to 17 and three persons aged 26 to 34. From this household composition, one can construct one pair of persons aged 12 to 17, three pairs of persons aged 26 to 34, and six pairs of persons 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 persons 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 persons selected from dwelling units) and were calculated at the QDU (household) level. The pair weights were adjusted by the appropriate multiplicity. See Chapter 10 in Frechtel and Laufenberg (2012) for details on the multiplicity counts and household-level control totals, which are referred to as household-level person counts.

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

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

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

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

Introduction

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

For 2011, modeling involved taking four model groups through three adjustment steps: (1) selected dwelling unit poststratification, (2) respondent dwelling unit nonresponse adjustment, and (3) respondent dwelling unit poststratification. After the final poststratification, the adjusted sampling weights were reasonably distributed and did not require the additional treatment of the extreme value step.

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

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

C.1 Final Model Explanatory Variables

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

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

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

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

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

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

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

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

Household Type (Ages of Persons Rostered within DU)[a]
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)[a]
1: 50-100%,[1] 2: 10->50%, 3: 0->10%

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

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

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

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

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

Race (5 Levels)[b]
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[a]
1: Hispanic or Latino White,[1] 2: Hispanic or Latino Black or African American, 3: Hispanic or Latino Other, 4: Non-Hispanic or Latino White, 5: Non-Hispanic or Latino Black or African American, 6: Non-Hispanic or Latino Other

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

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

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

State[a,b,3]

Model Group 1: 1: Connecticut, 2: Maine, 3: Massachusetts,[1] 4: New Hampshire, 5: New Jersey, 6: New York, 7: Pennsylvania, 8: Rhode Island, 9: Vermont

Model Group 2: 1: Illinois, 2: Indiana, 3: Iowa, 4: Kansas, 5: Michigan, 6: Minnesota, 7: Missouri, 8: Nebraska, 9: North Dakota, 10: Ohio, 11: South Dakota, 12: Wisconsin[1]

Model Group 3: 1: Alabama, 2: Arkansas, 3: Delaware, 4: District of Columbia, 5: Florida, 6: Georgia, 7: Kentucky, 8: Louisiana, 9: Maryland, 10: Mississippi, 11: North Carolina,[1] 12: Oklahoma, 13: South Carolina, 14: Tennessee, 15: Texas, 16: Virginia, 17: West Virginia

Model Group 4: 1: Alaska, 2: Arizona,[1] 3: California, 4: Colorado, 5: Idaho, 6: Hawaii, 7: Montana, 8: Nevada, 9: New Mexico, 10: Oregon, 11: Utah, 12: Washington, 13: Wyoming

State/Region[a,3]

Model Group 1: 1: New York, 2: Pennsylvania, 3: Other[1]

Model Group 2: 1: Illinois, 2: Michigan, 3: Ohio, 4: Other[1]

Model Group 3: 1: Florida, 2: Texas, 3: Other[1]

Model Group 4: 1: California, 2: Other[1]

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

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

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

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

[a] Binary variable.

[b] Counting variable. A count of all persons in the household.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

3. Construct the cross-classification table.

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

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

Shading indicates the reference-level set.

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

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

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

Shading indicates the reference-level set.

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

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

Shading indicates the reference-level set.

The number of respondents in the class State × Age × Race (3 Levels) at this stage of modeling would appear within each cell of the table. Construction of the other cross-

classification tables follows the same logic and is only necessary to the point of providing understanding of the final table.

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

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

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

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

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

Shading indicates the reference-level set.

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

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

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

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

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

Shading indicates the reference-level set.

Exhibit C.2 Covariates for 2011 NSDUH Questionnaire Dwelling Unit Weights

Variables	Binary	Counting	Level	Proposed
One-Factor Effects		**76**	**76**	
Intercept	Yes		1	1
Population Density	Yes		4	3
Group Quarter	Yes		3	2
Race/Ethnicity of Householder	Yes		6	5
Rent/Housing	Yes		5	4
Segment % Black or African American	Yes		3	2
Segment % Hispanic or Latino	Yes		3	2
Segment % Owner-Occupied	Yes		3	2
Household Type	Yes		7	6
State	Yes	Yes	Model-specific	
Quarter	Yes	Yes	4	3
Age Group		Yes	5	4
Race		Yes	5	4
Hispanicity		Yes	2	1
Gender		Yes	2	1
Household Size		Yes	1	1
Two-Factor Effects				
Age × Race (3 Levels)		Yes	5 × 3	8
Age × Hispanicity		Yes	5 × 2	4
Age × Gender		Yes	5 × 2	4
Race (3 Levels) × Hispanicity		Yes	3 × 2	2
Race (3 Levels) × Gender		Yes	3 × 2	2
Hispanicity × Gender		Yes	2 × 2	1
State × Age		Yes	Model-specific	
State × Race (5 Levels)		Yes	Model-specific	
State × Gender		Yes	Model-specific	
State × Hispanicity		Yes	Model-specific	
% Black or African American × % Owner-Occupied	Yes		3 × 3	4
% Black or African American × Rent/Housing		Yes	3 × 5	8
% Hispanicity × % Owner-Occupied		Yes	3 × 3	4
% Hispanicity × Rent/Housing		Yes	3 × 5	8
% Owner × Rent/Housing	Yes		3 × 5	8
Three-Factor Effects				
Race (3 Levels) × Age × Gender		Yes	8	8
State/Region × Age × Gender		Yes		
State/Region × Age × Hispanicity		Yes		
State/Region × Age × Race (3 Levels)		Yes		
State/Region × Hispanicity × Gender		Yes		
State/Region × Race (3 Levels) × Hispanicity		Yes		
State/Region × Race (3 Levels) × Gender		Yes		

Appendix C.1: Model Group 1: Northeast

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

Table C.1a 2011 QDU Weight GEM Modeling Summary (Model Group 1: Northeast)

Modeling Step[1]	Extreme Weight Proportions			UWE[2]	# Covariates[3]	Bounds[4]	
	% Unweighted	% Weighted	% Outwinsor			Nominal	Realized
sel.qdu.ps	2.40	4.54	1.32	3.0155	243	(0.58, 2.94)	(0.59, 2.94)
	1.80	2.97	0.68	2.8292	243	(0.45, 2.37)	(0.46, 2.29)
						(0.90, 1.66)	(0.90, 1.66)
res.qdu.nr	1.79	2.87	0.78	2.9050	243	(1.00, 2.70)	(1.00, 2.70)
	1.48	3.24	0.73	3.2831	242	(1.00, 3.19)	(1.00, 3.11)
						(1.30, 5.00)	(1.30, 1.35)
res.qdu.ps	1.48	3.24	0.73	3.2831	243	(0.20, 2.10)	(0.96, 2.09)
	1.68	3.29	0.52	3.2776	242	(0.20, 5.00)	(0.89, 1.46)
						(0.90, 5.00)	(0.96, 1.06)

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

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

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

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

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

Table C.1b 2011 Distribution of Weight Adjustment Factors and Weight Products (Model Group 1: Northeast)

	SDU Weight	QDU Design Weight		sel.qdu.ps[1]		res.qdu.nr[1]		res.qdu.ps[1]	
	1-10	duwght11	1-11	duwght12	1-12	duwght13	1-13	duwght14	1-14
Minimum	28	1.00	28	0.29	26	0.44	28	0.60	27
1%	81	1.00	92	0.60	89	1.00	101	0.93	101
5%	120	1.00	159	0.75	159	1.01	170	0.97	170
10%	186	1.00	235	0.82	231	1.04	251	0.98	252
25%	327	1.00	549	0.91	528	1.12	575	0.99	572
Median	649	1.15	918	0.99	921	1.23	1,089	1.00	1,089
75%	875	3.67	2,045	1.09	2,052	1.37	2,349	1.01	2,365
90%	1,281	6.48	4,401	1.23	4,394	1.52	5,754	1.02	5,733
95%	1,625	8.34	6,536	1.34	6,525	1.66	8,794	1.03	8,777
99%	2,524	12.47	12,081	1.60	12,230	2.10	16,719	1.11	16,873
Maximum	8,047	15.68	62,767	2.91	26,059	3.85	51,507	1.46	52,502
n	11,997	-	11,997	-	11,997	-	9,456	-	9,456
Mean	712	2.60	1,806	1.01	1,799	1.27	2,282	1.00	2,282
Max/Mean	11	-	35	-	15	-	23	-	23

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

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

Model Group 1 Overview

Selected Questionnaire Dwelling Unit-Level Poststratification

All 243 proposed effects were kept in the model.

Respondent Questionnaire Dwelling Unit-Level Nonresponse

Out of 243 proposed effects, 242 were kept in the model, with the exception of State by Race, which combined American Indian or Alaska Native and Asian for New Hampshire.

Respondent Questionnaire Dwelling Unit-Level Poststratification

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

Exhibit C.1.1 Covariates for 2011 NSDUH Questionnaire Dwelling Unit Weights (sel.qdu.ps) Model Group 1: Northeast

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

Exhibit C.1.2 Covariates for 2011 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 1: Northeast

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		**60**	**60**	
Intercept	1	1	1	All levels present.
Group Quarter	3	2	2	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Household Type	7	6	6	All levels present.
Household Size	1	1	1	All levels present.
Rent/Housing	5	4	4	All levels present.
Population Density	4	3	3	All levels present.
% Black or African American	3	2	2	All levels present.
% Hispanic or Latino	35	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
State (Count)	9	8	8	All levels present.
State (Binary)	9	8	8	All levels present.
Quarter (Count)	4	3	3	All levels present.
Quarter (Binary)	4	3	3	All levels present.
Age Group	5	4	4	All levels present.
Race	5	4	4	All levels present.
Hispanicity	2	1	1	All levels present.
Gender	2	1	1	All levels present.
Two-Factor Effects		**133**	**132**	
Age × Race (3 Levels)	5 × 3	8	8	All levels present.
Age × Hispanicity	5 × 2	4	4	All levels present.
Age × Gender	5 × 2	4	4	All levels present.
Race (3 Levels) × Hispanicity	3 × 2	2	2	All levels present.
Race (3 Levels) × Gender	3 × 2	2	2	All levels present.
Hispanicity × Gender	2 × 2	1	1	All levels present.
State × Age	9 × 5	32	32	All levels present.
State × Race	9 × 5	32	31	Coll. (4,3) & (4,4); conv.
State × Gender	9 × 2	8	8	All levels present.
State × Hispanicity	9 × 2	8	8	All levels present.
% Black or African American × % Owner-Occupied	3 × 3	4	4	All levels present.
% Black or African American × Rent/Housing	3 × 5	8	8	All levels present.
% Hispanicity × % Owner-Occupied	3 × 3	4	4	All levels present.
% Hispanicity × Rent/Housing	3 × 5	8	8	All levels present.
% Owner-Occupied × Rent/Housing	3 × 5	8	8	All levels present.
Three-Factor Effects		**50**	**50**	
Race (3 Levels) × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Hispanicity	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Race (3 Levels)	3 × 5 × 3	16	16	All levels present.
State/Region × Hispanicity × Gender	3 × 2 × 2	2	2	All levels present.
State/Region × Race (3 Levels) × Hispanicity	3 × 3 × 2	4	4	All levels present.
State/Region × Race (3 Levels) × Gender	3 × 3 × 2	4	4	All levels present.
Total		**243**	**242**	

Exhibit C.1.3 Covariates for 2011 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 1: Northeast

This step used the same variables as the respondent questionnaire dwelling unit-level nonresponse adjustment step in Exhibit C.1.2.

Appendix C.2: Model Group 2: Midwest

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

Table C.2a 2011 QDU Weight GEM Modeling Summary (Model Group 2: Midwest)

Modeling Step[1]	Extreme Weight Proportions			UWE[2]	# Covariates[3]	Bounds[4]	
	% Unweighted	% Weighted	% Outwinsor			Nominal	Realized
sel.qdu.ps	1.57	1.97	0.50	2.5453	300	(0.40, 2.00)	(0.40, 2.00)
	1.27	1.57	0.31	2.5042	300	(0.53, 4.20)	(0.54, 4.18)
						(0.90, 2.49)	(0.90, 2.49)
res.qdu.nr	1.48	1.86	0.34	2.5621	300	(1.00, 2.90)	(1.00, 2.90)
	1.01	1.17	0.31	2.7489	295	(1.00, 5.00)	(1.00, 5.00)
						(1.00, 4.66)	(1.00, 4.65)
res.qdu.ps	1.01	1.17	0.31	2.7489	300	(0.20, 2.10)	(0.94, 2.09)
	1.08	1.16	0.24	2.7531	299	(0.20, 5.00)	(0.69, 1.63)
						(0.90, 5.00)	(0.91, 1.99)

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

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

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

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

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

Table C.2b 2011 Distribution of Weight Adjustment Factors and Weight Products (Model Group 2: Midwest)

	SDU Weight	QDU Design Weight		sel.qdu.ps[1]		res.qdu.nr[1]		res.qdu.ps[1]	
	1-10	duwght11	1-11	duwght12	1-12	duwght13	1-13	duwght14	1-14
Minimum	21	1.00	21	0.23	40	0.56	44	0.57	41
1%	96	1.00	118	0.68	115	1.00	125	0.92	125
5%	139	1.00	199	0.81	196	1.03	216	0.98	215
10%	253	1.00	372	0.86	361	1.06	390	0.99	392
25%	447	1.00	525	0.93	528	1.12	613	1.00	613
Median	558	1.14	745	1.00	761	1.21	893	1.00	894
75%	722	3.61	1,790	1.08	1,792	1.32	2,047	1.01	2,045
90%	1,131	6.57	3,841	1.18	3,849	1.45	5,015	1.01	4,984
95%	1,350	7.62	5,408	1.26	5,484	1.53	7,155	1.02	7,165
99%	1,935	11.07	9,555	1.56	9,126	1.80	12,674	1.09	12,765
Maximum	6,263	13.34	23,020	4.18	22,231	6.07	29,765	1.99	29,668
n	17,045	-	17,045	-	17,045	-	13,752	-	13,752
Mean	636	2.50	1,550	1.02	1,546	1.24	1,917	1.00	1,917
Max/Mean	10	-	15	-	14	-	16	-	15

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

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

Model Group 2 Overview

Selected Questionnaire Dwelling Unit-Level Poststratification

All 300 proposed effects were kept in the model.

Respondent Questionnaire Dwelling Unit-Level Nonresponse

All main effects were maintained at proposed levels. Two-factor effects were modified for State by Race, combining American Indian or Alaska Native and Asian for Missouri and Ohio. Three-factor effects were modified for State by Race by Hispanicity, combining Black or African American and Other for Illinois and Ohio, and were modified for State by Age by Race, combining Black or African American and Other for Ohio.

Respondent Questionnaire Dwelling Unit-Level Poststratification

Out of 300 proposed effects, 299 were kept in the model. Two-factor effects were modified by State by Race, combining American Indian or Alaska Native and Asian for Missouri.

Exhibit C.2.1 Covariates for 2011 NSDUH Questionnaire Dwelling Unit Weights (sel.qdu.ps) Model Group 2: Midwest

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

Exhibit C.2.2 Covariates for 2011 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 2: Midwest

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		**66**	**66**	
Intercept	1	1	1	All levels present.
Group Quarter	3	2	2	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Household Type	7	6	6	All levels present.
Household Size	1	1	1	All levels present.
Rent/Housing	5	4	4	All levels present.
Population Density	4	3	3	All levels present.
% Black or African American	3	2	2	All levels present.
% Hispanic or Latino	35	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
State (Count)	9	11	11	All levels present.
State (Binary)	9	11	11	All levels present.
Quarter (Count)	4	3	3	All levels present.
Quarter (Binary)	4	3	3	All levels present.
Age Group	5	4	4	All levels present.
Race	5	4	4	All levels present.
Hispanicity	2	1	1	All levels present.
Gender	2	1	1	All levels present.
Two-Factor Effects		**163**	**161**	
Age × Race (3 Levels)	5 × 3	8	8	All levels present.
Age × Hispanicity	5 × 2	4	4	All levels present.
Age × Gender	5 × 2	4	4	All levels present.
Race (3 Levels) × Hispanicity	3 × 2	2	2	All levels present.
Race (3 Levels) × Gender	3 × 2	2	2	All levels present.
Hispanicity × Gender	2 × 2	1	1	All levels present.
State × Age	12 × 5	44	44	All levels present.
State × Race	12 × 5	44	42	Coll (10,3) & (10,4), (7,3) & (7,4); conv.
State × Gender	12 × 2	11	11	All levels present.
State × Hispanicity	12 × 2	11	11	All levels present.
% Black or African American × % Owner-Occupied	3 × 3	4	4	All levels present.
% Black or African American × Rent/Housing	3 × 5	8	8	All levels present.
% Hispanicity × % Owner-Occupied	3 × 3	4	3	All levels present.
% Hispanicity × Rent/Housing	3 × 5	8	8	All levels present.
% Owner-Occupied × Rent/Housing	3 × 5	8	8	All levels present.
Three-Factor Effects		**71**	**68**	
Race (3 Levels) × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Gender	4 × 5 × 2	12	12	All levels present.
State/Region × Age × Hispanicity	4 × 5 × 2	12	12	All levels present.
State/Region × Age × Race (3 Levels)	4 × 5 × 3	24	23	Coll (10,5,2) & (10,5,3); conv.
State/Region × Hispanicity × Gender	4 × 2 × 2	3	3	All levels present.
State/Region × Race (3 Levels) × Hispanicity	4 × 3 × 2	6	4	Coll (1,2,1) & (1,3,1); (10,2,1) & (10,3,1); conv.
State/Region × Race (3 Levels) × Gender	4 × 3 × 2	6	6	All levels present.
Total		**300**	**295**	

Exhibit C.2.3 Covariates for 2011 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 2: Midwest

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		**66**	**66**	
Intercept	1	1	1	All levels present.
Group Quarter	3	2	2	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Household Type	7	6	6	All levels present.
Household Size	1	1	1	All levels present.
Rent/Housing	5	4	4	All levels present.
Population Density	4	3	3	All levels present.
% Black or African American	3	2	2	All levels present.
% Hispanic or Latino	35	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
State (Count)	9	11	11	All levels present.
State (Binary)	9	11	11	All levels present.
Quarter (Count)	4	3	3	All levels present.
Quarter (Binary)	4	3	3	All levels present.
Age Group	5	4	4	All levels present.
Race	5	4	4	All levels present.
Hispanicity	2	1	1	All levels present.
Gender	2	1	1	All levels present.
Two-Factor Effects		**163**	**162**	
Age × Race (3 Levels)	5 × 3	8	8	All levels present.
Age × Hispanicity	5 × 2	4	4	All levels present.
Age × Gender	5 × 2	4	4	All levels present.
Race (3 Levels) × Hispanicity	3 × 2	2	2	All levels present.
Race (3 Levels) × Gender	3 × 2	2	2	All levels present.
Hispanicity × Gender	2 × 2	1	1	All levels present.
State × Age	12 × 5	44	44	All levels present.
State × Race	12 × 5	44	43	Coll. (7,3) & (7,4); conv.
State × Gender	12 × 2	11	11	All levels present.
State × Hispanicity	12 × 2	11	11	All levels present.
% Black or African American × % Owner-Occupied	3 × 3	4	4	All levels present.
% Black or African American × Rent/Housing	3 × 5	8	8	All levels present.
% Hispanicity × % Owner-Occupied	3 × 3	4	3	All levels present.
% Hispanicity × Rent/Housing	3 × 5	8	8	All levels present.
% Owner-Occupied × Rent/Housing	3 × 5	8	8	All levels present.
Three-Factor Effects		**71**	**71**	
Race (3 Levels) × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Gender	4 × 5 × 2	12	12	All levels present.
State/Region × Age × Hispanicity	4 × 5 × 2	12	12	All levels present.
State/Region × Age × Race (3 Levels)	4 × 5 × 3	24	24	All levels present.
State/Region × Hispanicity × Gender	4 × 2 × 2	3	3	All levels present.
State/Region × Race (3 Levels) × Hispanicity	4 × 3 × 2	6	6	All levels present.
State/Region × Race (3 Levels) × Gender	4 × 3 × 2	6	6	All levels present.
Total		**300**	**299**	

Appendix C.3: Model Group 3: South

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

Table C.3a 2011 QDU Weight GEM Modeling Summary (Model Group 3: South)

Modeling Step[1]	Extreme Weight Proportions			UWE[2]	# Covariates[3]	Bounds[4]	
	% Unweighted	% Weighted	% Outwinsor			Nominal	Realized
sel.qdu.ps	2.03	3.41	0.90	2.8988	339	(0.50, 1.40)	(0.50, 1.40)
	1.42	2.55	0.42	2.7532	338	(0.36, 3.10)	(0.36, 3.08)
						(0.90, 1.96)	(0.90, 1.96)
res.qdu.nr	1.39	2.38	0.42	2.8273	339	(1.00, 1.80)	(1.00, 1.80)
	1.32	2.42	0.38	2.9966	338	(1.00, 4.35)	(1.00, 4.26)
						(1.20, 5.00)	(1.20, 1.23)
res.qdu.ps	1.32	2.42	0.38	2.9966	339	(0.61, 2.88)	(0.62, 2.88)
	1.22	2.11	0.30	2.9971	338	(0.52, 3.34)	(0.53, 3.32)
						(0.90, 1.05)	(0.90, 1.04)

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

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

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

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

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

Table C.3b 2011 Distribution of Weight Adjustment Factors and Weight Products (Model Group 3: South)

	SDU Weight	QDU Design Weight		sel.qdu.ps[1]		res.qdu.nr[1]		res.qdu.ps[1]	
	1-10	duwght11	1-11	duwght12	1-12	duwght13	1-13	duwght14	1-14
Minimum	11	1.00	12	0.26	11	0.46	11	0.46	9
1%	58	1.00	75	0.61	72	1.00	76	0.86	76
5%	142	1.00	211	0.79	211	1.01	234	0.97	230
10%	227	1.00	311	0.85	312	1.04	346	0.99	345
25%	474	1.00	688	0.93	680	1.08	744	1.00	743
Median	824	1.17	1,188	1.00	1,197	1.16	1,338	1.00	1,338
75%	1,205	3.49	2,401	1.09	2,437	1.26	2,688	1.01	2,689
90%	1,621	6.58	5,613	1.19	5,697	1.40	7,111	1.01	7,094
95%	2,029	7.85	7,972	1.29	8,452	1.50	10,683	1.03	10,687
99%	3,260	11.83	13,754	1.68	14,342	1.77	18,301	1.09	18,327
Maximum	11,649	18.71	104,265	3.85	46,883	4.26	45,483	3.32	47,840
n	19,690	-	19,690	-	19,690	-	16,487	-	16,487
Mean	922	2.55	2,239	1.02	2,263	1.20	2,702	1.00	2,702
Max/Mean	13	-	47	-	21	-	17	-	18

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

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

Model Group 3 Overview

Selected Questionnaire Dwelling Unit-Level Poststratification

Out of 339 proposed effects, 338 were kept in the model. Two-factor effects were modified for State by Race, combining Asian and Two or More Races for West Virginia.

Respondent Questionnaire Dwelling Unit-Level Nonresponse

Out of 339 proposed effects, 338 were kept in the model. Three-factor effects were modified for State by Race by Hispanicity, combining Black or African American with Other for Florida.

Respondent Questionnaire Dwelling Unit-Level Poststratification

Out of 339 proposed effects, 338 were kept in the model. Two-factor effects were modified for State by Race, combining American Indian or Alaska Native and Asian for West Virginia.

Exhibit C.3.1 Covariates for 2011 NSDUH Questionnaire Dwelling Unit Weights (sel.qdu.ps) Model Group 3: South

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

Exhibit C.3.2 Covariates for 2011 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 3: South

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		**76**	**76**	
Intercept	1	1	1	All levels present.
Group Quarter	3	2	2	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Household Type	7	6	6	All levels present.
Household Size	1	1	1	All levels present.
Rent/Housing	5	4	4	All levels present.
Population Density	4	3	3	All levels present.
% Black or African American	3	2	2	All levels present.
% Hispanic or Latino	35	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
State (Count)	17	16	16	All levels present.
State (Binary)	17	16	16	All levels present.
Quarter (Count)	4	3	3	All levels present.
Quarter (Binary)	4	3	3	All levels present.
Age Group	5	4	4	All levels present.
Race	5	4	4	All levels present.
Hispanicity	2	1	1	All levels present.
Gender	2	1	1	All levels present.
Two-Factor Effects		**213**	**213**	
Age × Race (3 Levels)	5 × 3	8	8	All levels present.
Age × Hispanicity	5 × 2	4	4	All levels present.
Age × Gender	5 × 2	4	4	All levels present.
Race (3 Levels) × Hispanicity	3 × 2	2	2	All levels present.
Race (3 Levels) × Gender	3 × 2	2	2	All levels present.
Hispanicity × Gender	2 × 2	1	1	All levels present.
State × Age	17 × 5	64	64	All levels present.
State × Race	17 × 5	64	64	All levels present.
State × Gender	17 × 2	16	16	All levels present.
State × Hispanicity	17 × 2	16	16	All levels present.
% Black or African American × % Owner-Occupied	3 × 3	4	4	All levels present.
% Black or African American × Rent/Housing	3 × 5	8	8	All levels present.
% Hispanicity × % Owner-Occupied	3 × 3	4	4	All levels present.
% Hispanicity × Rent/Housing	3 × 5	8	8	All levels present.
% Owner-Occupied × Rent/Housing	3 × 5	8	8	All levels present.
Three-Factor Effects		**50**	**49**	
Race (3 Levels) × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Hispanicity	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Race (3 Levels)	3 × 5 × 3	16	16	All levels present.
State/Region × Hispanicity × Gender	3 × 2 × 2	2	2	All levels present.
State/Region × Race (3 Levels) × Hispanicity	3 × 3 × 2	4	3	Coll (5,2,1) & (5,3,1); conv.
State/Region × Race (3 levels) × Gender	3 × 3 × 2	4	4	All levels present.
Total		**339**	**338**	

Exhibit C.3.3 Covariates for 2011 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 3: South

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

Appendix C.4: Model Group 4: West

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

Table C.4a 2011 QDU Weight GEM Modeling Summary (Model Group 4: West)

Modeling Step[1]	Extreme Weight Proportions			UWE[2]	# Covariates[3]	Bounds[4]	
	% Unweighted	% Weighted	% Outwinsor			Nominal	Realized
sel.qdu.ps	2.10	4.69	1.50	3.2191	270	(0.50, 2.99)	(0.52, 2.99)
	1.57	3.52	0.87	3.1934	267	(0.32, 2.52)	(0.33, 2.46)
						(0.90, 1.56)	(0.90, 1.56)
res.qdu.nr	1.67	3.55	1.00	3.3065	270	(1.00, 2.70)	(1.00, 2.70)
	1.52	3.52	0.73	3.6981	266	(1.00, 2.86)	(1.00, 2.84)
						(1.30, 5.00)	(1.30, 1.30)
res.qdu.ps	1.52	3.52	0.73	3.6981	270	(0.20, 2.10)	(0.94, 2.10)
	1.52	3.33	0.51	3.7066	267	(0.20, 5.00)	(0.70, 1.74)
						(0.90, 5.00)	(0.96, 0.99)

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

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

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

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

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

Table C.4b 2011 Distribution of Weight Adjustment Factors and Weight Products (Model Group 4: West)

	SDU Weight	QDU Design Weight		sel.qdu.ps[1]		res.qdu.nr[1]		res.qdu.ps[1]	
	1-10	duwght11	1-11	duwght12	1-12	duwght13	1-13	duwght14	1-14
Minimum	18	1.00	18	0.30	29	0.55	30	0.60	28
1%	74	1.00	78	0.60	75	1.00	81	0.91	80
5%	101	1.00	117	0.76	120	1.02	136	0.97	135
10%	122	1.00	166	0.83	166	1.04	186	0.98	186
25%	233	1.00	388	0.91	383	1.09	414	0.99	415
Median	660	1.13	1,069	1.00	1,054	1.18	1,188	1.00	1,187
75%	1,424	3.16	2,181	1.09	2,203	1.30	2,491	1.01	2,484
90%	1,952	6.10	5,386	1.20	5,336	1.45	6,169	1.02	6,150
95%	2,272	7.92	8,345	1.29	8,507	1.56	10,876	1.03	10,840
99%	3,339	12.21	15,987	1.54	15,566	1.87	22,021	1.07	22,097
Maximum	10,206	15.89	48,334	3.19	33,865	2.84	46,269	1.74	48,746
n	12,709	-	12,709	-	12,709	-	10,438	-	10,438
Mean	899	2.41	2,079	1.01	2,081	1.22	2,534	1.00	2,534
Max/Mean	11	-	23	-	16	-	18	-	19

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

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

Model Group 4 Overview

Selected Questionnaire Dwelling Unit-Level Poststratification

All main effects were maintained in full. Two-factor effects were modified for Rent/Housing by Percent Black or African American, combining 50-100% and 10-<50% for the first and the fourth quintiles, and were modified for Percent Owner Occupied by Percent Black or African American, combining 50-100% Black or African American and 10-<50% Black or African American for 50-100% owner occupied.

Respondent Questionnaire Dwelling Unit-Level Nonresponse

All main effects were maintained in full. Two-factor effects were modified for Rent/Housing by Percent Black or African American, combining 50-100% and 10-<50% for the first and the fourth quintiles, were modified for Percent Owner Occupied by Percent Black or African American, combining 50-100% Black or African American and 10-<50% Black or African American for 50-100% owner occupied, and were modified for State by Race, combining American Indian or Alaska Native and Asian for Wyoming.

Respondent Questionnaire Dwelling Unit-Level Poststratification

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

Exhibit C.4.1 Covariates for 2011 NSDUH Questionnaire Dwelling Unit Weights (sel.qdu.ps) Model Group 4: West

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		**68**	**68**	
Intercept	1	1	1	All levels present.
Group Quarter	3	2	2	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Household Type	7	6	6	All levels present.
Household Size	1	1	1	All levels present.
Rent/Housing	5	4	4	All levels present.
Population Density	4	3	3	All levels present.
% Black or African American	3	2	2	All levels present.
% Hispanic or Latino	35	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
State (Count)	13	12	12	All levels present.
State (Binary)	13	12	12	All levels present.
Quarter (Count)	4	3	3	All levels present.
Quarter (Binary)	4	3	3	All levels present.
Age Group	5	4	4	All levels present.
Race	5	4	4	All levels present.
Hispanicity	2	1	1	All levels present.
Gender	2	1	1	All levels present.
Two-Factor Effects		**173**	**170**	
Age × Race (3 Levels)	5 × 3	8	8	All levels present.
Age × Hispanicity	5 × 2	4	4	All levels present.
Age × Gender	5 × 2	4	4	All levels present.
Race (3 Levels) × Hispanicity	3 × 2	2	2	All levels present.
Race (3 Levels) × Gender	3 × 2	2	2	All levels present.
Hispanicity × Gender	2 × 2	1	1	All levels present.
State × Age	13 × 5	48	48	All levels present.
State × Race	13 × 5	48	48	All levels present.
State × Gender	13 × 2	12	12	All levels present.
State × Hispanicity	13 × 2	12	12	All levels present.
% Black or African American × % Owner-Occupied	3 × 3	4	3	Coll. (1,2) & (2,2); sing.
% Black or African American × Rent/Housing	3 × 5	8	6	Coll. (1,1) & (2,1); zero. Coll. (1,4) & (2,4); sing.
% Hispanicity × % Owner-Occupied	3 × 3	4	4	All levels present.
% Hispanicity × Rent/Housing	3 × 5	8	8	All levels present.
% Owner-Occupied × Rent/Housing	3 × 5	8	8	All levels present.
Three-Factor Effects		**29**	**29**	
Race (3 Levels) × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Gender	2 × 5 × 2	4	4	All levels present.
State/Region × Age × Hispanicity	2 × 5 × 2	4	4	All levels present.
State/Region × Age × Race (3 Levels)	2 × 5 × 3	8	8	All levels present.
State/Region × Hispanicity × Gender	2 × 2 × 2	1	1	All levels present.
State/Region × Race (3 Levels) × Hispanicity	2 × 3 × 2	2	2	All levels present.
State/Region × Race (3 Levels) × Gender	2 × 3 × 2	2	2	All levels present.
Total		**270**	**267**	

Exhibit C.4.2 Covariates for 2011 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 4: West

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		**68**	**68**	
Intercept	1	1	1	All levels present.
Group Quarter	3	2	2	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Household Type	7	6	6	All levels present.
Household Size	1	1	1	All levels present.
Rent/Housing	5	4	4	All levels present.
Population Density	4	3	3	All levels present.
% Black or African American	3	2	2	All levels present.
% Hispanic or Latino	35	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
State (Count)	13	12	12	All levels present.
State (Binary)	13	12	12	All levels present.
Quarter (Count)	4	3	3	All levels present.
Quarter (Binary)	4	3	3	All levels present.
Age Group	5	4	4	All levels present.
Race	5	4	4	All levels present.
Hispanicity	2	1	1	All levels present.
Gender	2	1	1	All levels present.
Two-Factor Effects		**173**	**169**	
Age × Race (3 Levels)	5 × 3	8	8	All levels present.
Age × Hispanicity	5 × 2	4	4	All levels present.
Age × Gender	5 × 2	4	4	All levels present.
Race (3 Levels) × Hispanicity	3 × 2	2	2	All levels present.
Race (3 Levels) × Gender	3 × 2	2	2	All levels present.
Hispanicity × Gender	2 × 2	1	1	All levels present.
State × Age	13 × 5	48	48	All levels present.
State × Race	13 × 5	48	47	Coll (13,3) & (13,4); conv
State × Gender	13 × 2	12	12	All levels present.
State × Hispanicity	13 × 2	12	12	All levels present.
% Black or African American × % Owner-Occupied	3 × 3	4	3	Coll. (1,2) & (2,2); sing.
% Black or African American × Rent/Housing	3 × 5	8	6	Coll. (1,1) & (2,1); zero. Coll. (1,4) & (2,4); sing.
% Hispanicity × % Owner-Occupied	3 × 3	4	4	All levels present.
% Hispanicity × Rent/Housing	3 × 5	8	8	All levels present.
% Owner-Occupied × Rent/Housing	3 × 5	8	8	All levels present.
Three-Factor Effects		**29**	**29**	
Race (3 Levels) × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Gender	2 × 5 × 2	4	4	All levels present.
State/Region × Age × Hispanicity	2 × 5 × 2	4	4	All levels present.
State/Region × Age × Race (3 Levels)	2 × 5 × 3	8	8	All levels present.
State/Region × Hispanicity × Gender	2 × 2 × 2	1	1	All levels present.
State/Region × Race (3 Levels) × Hispanicity	2 × 3 × 2	2	2	All levels present.
State/Region × Race (3 Levels) × Gender	2 × 3 × 2	2	2	All levels present.
Total		**270**	**266**	

Exhibit C.4.3 Covariates for 2011 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 4: West

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

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

Table D.1 2011 NSDUH QDU-Level Response Rates

Domain	Selected QDU	Respondent QDU	% Interview Response Rate[1]
Total	61,441	50,133	77.10
Census Region			
Northeast	11,997	9,456	73.31
South	19,690	16,487	79.26
Midwest	17,045	13,752	76.79
West	12,709	10,438	76.85
Quarter			
Quarter 1	14,360	11,633	76.00
Quarter 2	16,559	13,462	77.03
Quarter 3	15,796	12,992	77.90
Quarter 4	14,726	12,046	77.48
Household Type			
12-17, 18-25, 26+	5,958	5,208	87.63
12-17, 18-25	96	86	89.50
12-17, 26+	17,812	15,210	85.59
18-25, 26+	12,697	10,373	81.30
12-17	24	18	82.51
18-25	6,307	5,450	85.32
26+	18,547	13,788	73.69
Race/Ethnicity of Householder			
Hispanic or Latino White	7,747	6,528	80.11
Hispanic or Latino Black or African American	149	122	79.98
Hispanic or Latino Other	418	358	81.34
Non-Hispanic or Latino White	41,252	33,138	75.91
Non-Hispanic or Latino Black or African American	7,558	6,532	82.24
Non-Hispanic or Latino Other	4,317	3,455	74.13
% Hispanic or Latino in Segment			
50-100%	4,183	3,519	80.70
10-50%	10,536	8,590	76.91
<10%	46,722	38,024	76.79
% Black or African American in Segment			
50-100%	4,960	4,266	80.87
10-50%	9,474	7,936	80.09
<10%	47,007	37,931	76.06
% Owner-Occupied DUs in Segment			
50-100%	46,376	37,607	76.49
10-50%	11,674	9,661	78.48
<10%	3,391	2,865	80.68
Combined Median Rent/Housing Value			
1st Quintile	10,483	8,834	80.49
2nd Quintile	12,594	10,444	78.29
3rd Quintile	13,988	11,435	77.71
4th Quintile	13,178	10,601	75.69
5th Quintile	11,198	8,819	74.63
Population Density			
Large MSA	25,558	20,440	75.68
Medium to Small MSA	30,848	25,494	78.23
Non-MSA, Urban	1,489	1,269	81.87
Non-MSA, Rural	3,546	2,930	80.16
Group Quarters			
Group	687	659	96.83
Non-Group	60,754	49,474	76.99
Household Size			
One	7,324	5,768	75.52
Two	22,122	17,500	75.27
Three	17,557	14,554	80.08
Four or More	14,438	12,311	84.05

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

[1] The weight used for calculating the response rate includes SDU- and QDU-level design weights, SDU nonresponse and poststratification adjustments, and selected QDU poststratification adjustment. This weight is the product of WT1*...*WT10*DUWT11*DUWT12.

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

Table E.1 2011 NSDUH Selected QDU-Level Proportions of Extreme Values and Outwinsors

Domain	*n*	SDU-Level Weights[1] (SDUWT: WT1*...*WT10)			Before sel.qdu.ps[1] (SDUWT*DUWT11)			After sel.qdu.ps[1] (SDUWT*DUWT11*DUWT12)		
		% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]
Total	61,441	2.51	6.13	1.79	1.99	3.58	1.02	1.48	2.63	0.54
Census Region										
Northeast	11,997	2.91	7.28	2.25	2.40	4.54	1.32	1.80	2.97	0.68
South	19,690	2.50	5.88	1.61	2.03	3.41	0.90	1.42	2.55	0.42
Midwest	17,045	2.09	4.18	1.12	1.57	1.97	0.50	1.27	1.57	0.31
West	12,709	2.72	7.52	2.39	2.10	4.69	1.50	1.57	3.52	0.87
Quarter										
Quarter 1	14,360	2.69	6.52	1.86	1.99	3.88	1.13	1.62	3.00	0.61
Quarter 2	16,559	2.20	4.84	1.32	1.64	2.57	0.70	1.21	1.72	0.37
Quarter 3	15,796	2.27	5.52	1.62	1.93	3.88	1.07	1.49	2.82	0.55
Quarter 4	14,726	2.95	7.61	2.37	2.44	4.01	1.19	1.64	2.96	0.64
Household Type										
12-17, 18-25, 26+	5,958	2.80	7.20	1.89	2.80	7.20	1.89	2.35	6.89	1.73
12-17, 18-25	96	2.08	4.27	1.05	0.00	0.00	0.00	0.00	0.00	0.00
12-17, 26+	17,812	2.32	5.79	1.74	2.32	5.81	1.73	1.67	4.62	1.17
18-25, 26+	12,697	2.69	6.96	2.13	2.58	6.63	2.09	2.31	5.85	1.29
12-17	24	4.17	2.65	0.78	0.00	0.00	0.00	0.00	0.00	0.00
18-25	6,307	3.57	7.09	1.72	3.36	7.00	1.80	1.65	4.33	1.07
26+	18,547	2.12	5.13	1.59	0.55	2.23	0.61	0.42	1.36	0.20
Race/Ethnicity of Householder										
Hispanic or Latino White	7,747	2.32	4.78	1.29	1.82	3.31	0.86	1.21	2.09	0.36
Hispanic or Latino Black or African American	149	65.10	89.46	42.84	58.39	79.68	31.21	46.31	50.39	16.94
Hispanic or Latino Other	418	30.62	69.82	29.01	21.77	44.28	17.05	21.29	40.96	11.59
Non-Hispanic or Latino White	41,252	1.29	2.94	0.72	1.00	1.66	0.39	0.62	1.23	0.15
Non-Hispanic or Latino Black or African American	7,558	3.94	8.76	1.97	3.37	4.14	0.92	3.00	4.00	0.78
Non-Hispanic or Latino Other	4,317	7.09	13.71	3.50	5.49	9.31	2.57	4.10	6.55	1.65

Table E.1 2011 NSDUH Selected QDU-Level Proportions of Extreme Values and Outwinsors (continued)

Domain	*n*	SDU-Level Weights[1] (SDUWT: WT1*...*WT10)			Before sel.qdu.ps[1] (SDUWT*DUWT11)			After sel.qdu.ps[1] (SDUWT*DUWT11*DUWT12)		
		% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]
% Hispanic or Latino in Segment										
50-100%	4,183	2.06	5.45	1.92	1.67	3.71	1.01	1.51	3.24	0.94
10-50%	10,536	3.27	8.42	2.81	2.61	5.12	1.73	2.21	3.91	0.98
<10%	46,722	2.38	5.48	1.45	1.88	3.13	0.82	1.32	2.19	0.37
% Black or African American in Segment										
50-100%	4,960	3.06	9.14	2.85	2.54	6.99	2.22	2.36	4.61	0.93
10-50%	9,474	3.31	7.48	2.25	2.77	4.19	1.22	2.05	3.44	0.87
<10%	47,007	2.29	5.47	1.57	1.77	3.10	0.86	1.28	2.25	0.43
% Owner-Occupied DUs in Segment										
50-100%	46,376	2.19	5.74	1.69	1.77	3.45	1.00	1.25	2.33	0.44
10-50%	11,674	3.28	6.77	2.02	2.54	4.15	1.14	1.94	3.18	0.83
<10%	3,391	4.25	8.91	2.28	3.04	3.49	0.91	3.10	4.71	0.91
Combined Median Rent/Housing Value										
1st Quintile	10,483	2.84	6.29	1.93	2.31	3.98	1.18	1.64	2.81	0.58
2nd Quintile	12,594	2.09	4.97	1.52	1.83	2.95	0.83	1.29	1.98	0.54
3rd Quintile	13,988	2.49	6.06	1.72	1.99	2.93	0.80	1.31	2.02	0.52
4th Quintile	13,178	2.83	7.56	2.37	2.12	4.90	1.44	1.61	3.42	0.58
5th Quintile	11,198	2.33	5.60	1.41	1.71	3.12	0.86	1.63	2.83	0.49
Population Density										
Large MSA[1]	25,558	2.71	6.67	1.93	2.05	3.60	1.07	1.87	3.07	0.67
Medium to Small MSA[1]	30,848	2.45	5.91	1.76	2.01	3.93	1.07	1.26	2.32	0.44
Non-MSA,[1] Urban	1,489	1.41	1.42	0.44	1.41	0.83	0.22	0.74	0.56	0.16
Non-MSA,[1] Rural	3,546	2.06	3.76	0.97	1.58	1.36	0.38	0.99	1.47	0.17
Group Quarters										
Group	687	6.40	13.65	2.32	6.11	17.94	3.00	1.16	2.51	1.21
Non-Group	60,754	2.47	6.06	1.79	1.94	3.50	1.01	1.49	2.63	0.54
Household Size										
One	7,324	2.02	5.95	1.96	1.37	2.85	0.77	0.72	1.34	0.22
Two	22,122	2.56	5.78	1.76	1.74	2.96	0.85	1.09	1.98	0.31
Three	17,557	2.54	6.06	1.84	2.28	4.90	1.53	1.89	4.35	1.08
Four or More	14,438	2.65	6.76	1.72	2.33	6.00	1.62	1.98	5.87	1.48

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

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

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

Table E.2 2011 NSDUH Respondent QDU-Level Proportions of Extreme Values and Outwinsors

Domain	*n*	Before res.qdu.nr[1] (SDUWT*DUWT11*DUWT12) % Unweighted	Before res.qdu.nr[1] (SDUWT*DUWT11*DUWT12) % Weighted[2]	Before res.qdu.nr[1] (SDUWT*DUWT11*DUWT12) % Outwinsor[3]	After res.qdu.nr[1] (SDUWT*DUWT11*...*DUWT13) % Unweighted	After res.qdu.nr[1] (SDUWT*DUWT11*...*DUWT13) % Weighted[2]	After res.qdu.nr[1] (SDUWT*DUWT11*...*DUWT13) % Outwinsor[3]	Final Weight: After res.qdu.ps[1] (SDUWT*DUWT11*...*DUWT14) % Unweighted	Final Weight: After res.qdu.ps[1] (SDUWT*DUWT11*...*DUWT14) % Weighted[2]	Final Weight: After res.qdu.ps[1] (SDUWT*DUWT11*...*DUWT14) % Outwinsor[3]
Total	50,133	1.55	2.61	0.59	1.31	2.54	0.51	1.33	2.38	0.37
Census Region										
Northeast	9,456	1.79	2.87	0.78	1.48	3.24	0.73	1.68	3.29	0.52
South	16,487	1.39	2.38	0.42	1.32	2.42	0.38	1.22	2.11	0.30
Midwest	13,752	1.48	1.86	0.34	1.01	1.17	0.31	1.08	1.16	0.24
West	10,438	1.67	3.55	1.00	1.52	3.52	0.73	1.52	3.33	0.51
Quarter										
Quarter 1	11,633	1.62	2.71	0.63	1.57	3.02	0.54	1.50	2.67	0.39
Quarter 2	13,462	1.24	1.80	0.41	0.92	1.80	0.37	1.00	1.83	0.29
Quarter 3	12,992	1.52	2.71	0.57	1.24	2.48	0.54	1.30	2.41	0.41
Quarter 4	12,046	1.84	3.22	0.75	1.55	2.85	0.58	1.57	2.63	0.41
Household Type										
12-17, 18-25, 26+	5,208	2.32	6.99	1.79	1.73	5.50	1.36	1.88	5.51	1.08
12-17, 18-25	86	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12-17, 26+	15,210	1.74	4.79	1.25	1.60	4.85	1.21	1.72	4.87	0.84
18-25, 26+	10,373	2.35	5.77	1.33	1.70	5.12	1.21	1.71	5.01	0.94
12-17	18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18-25	5,450	1.71	4.50	1.20	1.61	4.87	0.88	1.45	4.05	0.73
26+	13,788	0.38	1.11	0.19	0.42	1.36	0.19	0.37	1.20	0.13
Race/Ethnicity of Householder										
Hispanic or Latino White	6,528	1.27	2.03	0.38	1.18	2.17	0.41	1.16	1.98	0.27
Hispanic or Latino Black or African American	122	47.54	48.94	19.10	44.26	50.10	15.14	44.26	50.43	14.14
Hispanic or Latino Other	358	20.67	39.04	11.18	15.08	34.23	9.95	15.64	34.29	8.49
Non-Hispanic or Latino White	33,138	0.65	1.12	0.17	0.59	1.11	0.13	0.54	0.96	0.08
Non-Hispanic or Latino Black or African American	6,532	3.11	4.35	0.83	1.65	2.74	0.52	1.84	2.41	0.39
Non-Hispanic or Latino Other	3,455	4.05	6.67	1.79	4.78	9.74	2.17	5.27	9.89	1.39
% Hispanic or Latino in Segment										
50-100%	3,519	1.71	3.87	1.06	1.48	3.89	0.84	1.65	4.03	0.72
10-50%	8,590	2.31	3.99	1.07	2.32	4.56	1.06	2.41	4.34	0.75
<10%	38,024	1.36	2.07	0.40	1.06	1.81	0.31	1.06	1.65	0.23

Table E.2 2011 NSDUH Respondent QDU-Level Proportions of Extreme Values and Outwinsors (continued)

Domain	*n*	Before res.qdu.nr[1] (SDUWT*DUWT11*DUWT12)			After res.qdu.nr[1] (SDUWT*DUWT11*...*DUWT13)			Final Weight: After res.qdu.ps[1] (SDUWT*DUWT11*...*DUWT14)		
		% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]
% Black or African American in Segment										
50-100%	4,266	2.30	4.40	0.95	1.57	4.02	0.90	1.52	3.41	0.58
10-50%	7,936	2.17	3.58	0.96	1.85	4.04	0.91	2.04	4.06	0.80
<10%	37,931	1.33	2.19	0.47	1.16	2.06	0.38	1.16	1.91	0.26
% Owner-Occupied DUs in Segment										
50-100%	37,607	1.30	2.27	0.47	1.13	2.24	0.43	1.11	2.04	0.30
10-50%	9,661	2.03	3.27	0.92	1.74	3.20	0.79	1.94	3.22	0.65
<10%	2,865	3.11	4.75	1.03	2.16	4.33	0.61	2.13	4.14	0.51
Combined Median Rent/Housing Value										
1st Quintile	8,834	1.71	2.73	0.63	1.06	2.10	0.52	1.13	2.00	0.38
2nd Quintile	10,444	1.33	2.12	0.62	1.24	2.26	0.53	1.23	2.08	0.44
3rd Quintile	11,435	1.39	2.24	0.57	1.01	1.80	0.40	1.05	1.81	0.33
4th Quintile	10,601	1.66	3.22	0.60	1.52	3.13	0.57	1.62	3.05	0.43
5th Quintile	8,819	1.70	2.71	0.55	1.77	3.18	0.52	1.67	2.77	0.30
Population Density										
Large MSA[1]	20,440	1.98	3.22	0.76	1.78	3.33	0.67	1.79	3.13	0.49
Medium to Small MSA[1]	25,494	1.31	2.17	0.47	1.05	1.90	0.37	1.06	1.74	0.26
Non-MSA,[1] Urban	1,269	0.71	0.57	0.09	0.47	0.40	0.02	0.63	0.56	0.04
Non-MSA,[1] Rural	2,930	0.96	1.22	0.19	0.58	0.60	0.17	0.82	0.94	0.23
Group Quarters										
Group	659	1.06	2.58	1.26	0.46	2.01	0.91	0.61	2.89	1.37
Non-Group	49,474	1.55	2.61	0.59	1.32	2.54	0.51	1.34	2.38	0.37
Household Size										
One	5,768	0.92	1.22	0.29	0.78	1.20	0.15	0.45	0.87	0.10
Two	17,500	1.15	1.84	0.32	1.04	2.00	0.33	1.09	1.88	0.26
Three	14,554	1.89	4.46	1.18	1.39	3.91	0.97	1.42	3.86	0.72
Four or More	12,311	1.99	5.88	1.44	1.84	5.98	1.42	1.99	5.92	1.01

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

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

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

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

Table F.1 2011 NSDUH QDU-Level Slippage Rates

Domain	*n*	Initial Total (*I*)[1]	Final Total (*F*)[2]	Control from SDU Weights (*C*)	(*I* - *C*)/*C*%	(*F* - *C*)/*C*%
Total	50,133	118,938,806	118,938,806	118,938,806	0.00	-0.00
Census Region						
Northeast	9,456	21,577,098	21,577,098	21,577,098	0.00	-0.00
South	16,487	44,555,849	44,555,849	44,555,849	0.00	-0.00
Midwest	13,752	26,357,805	26,357,805	26,357,805	0.00	-0.00
West	10,438	26,448,054	26,448,054	26,448,054	0.00	-0.00
Quarter						
Quarter 1	11,633	29,779,692	29,779,692	29,779,692	0.00	-0.00
Quarter 2	13,462	29,831,927	29,831,927	29,831,927	0.00	-0.00
Quarter 3	12,992	29,435,547	29,435,547	29,435,547	0.00	-0.00
Quarter 4	12,046	29,891,640	29,891,640	29,891,640	0.00	-0.00
Household Type						
12-17, 18-25, 26+	5,208	5,053,361	5,053,361	5,053,361	0.00	-0.00
12-17, 18-25	86	82,609	82,609	82,609	0.00	-0.00
12-17, 26+	15,210	13,705,155	13,705,155	13,705,155	0.00	-0.00
18-25, 26+	10,373	13,678,867	13,678,867	13,678,867	0.00	-0.00
12-17	18	13,747	13,747	13,747	0.00	0.00
18-25	5,450	5,726,024	5,726,024	5,726,024	0.00	-0.00
26+	13,788	80,679,043	80,679,043	80,679,043	0.00	-0.00
Race/Ethnicity of Householder						
Hispanic or Latino White	6,528	13,320,882	13,320,882	13,320,882	0.00	-0.00
Hispanic or Latino Black or African American	122	688,735	688,735	688,735	0.00	-0.00
Hispanic or Latino Other	358	1,028,089	1,028,089	1,028,089	0.00	-0.00
Non-Hispanic or Latino White	33,138	82,109,696	82,109,696	82,109,696	0.00	-0.00
Non-Hispanic or Latino Black or African American	6,532	14,324,393	14,324,392	14,324,392	0.00	-0.00
Non-Hispanic or Latino Other	3,455	7,467,011	7,467,011	7,467,011	0.00	-0.00
% Hispanic or Latino in Segment						
50-100%	3,519	8,710,000	8,710,000	8,710,000	0.00	-0.00
10-50%	8,590	24,895,079	24,895,079	24,895,079	0.00	-0.00
<10%	38,024	85,333,727	85,333,727	85,333,727	0.00	-0.00
% Black or African American in Segment						
50-100%	4,266	8,993,798	8,993,798	8,993,798	0.00	0.00
10-50%	7,936	19,924,839	19,924,839	19,924,839	0.00	-0.00
<10%	37,931	90,020,168	90,020,168	90,020,168	0.00	-0.00
% Owner-Occupied DUs in Segment						
50-100%	37,607	89,831,597	89,831,597	89,831,597	0.00	-0.00
10-50%	9,661	22,375,435	22,375,435	22,375,435	0.00	-0.00
<10%	2,865	6,731,773	6,731,773	6,731,773	0.00	-0.00

Table F.1 2011 NSDUH QDU-Level Slippage Rates (continued)

Domain	*n*	Initial Total (*I*)[1]	Final Total (*F*)[2]	Control from SDU Weights (*C*)	(*I* - *C*)/*C*%	(*F* - *C*)/*C*%
Combined Median Rent/Housing Value						
1st Quintile	8,834	17,780,836	17,780,836	17,780,836	0.00	-0.00
2nd Quintile	10,444	22,640,554	22,640,554	22,640,554	0.00	-0.00
3rd Quintile	11,435	25,284,953	25,284,953	25,284,953	0.00	-0.00
4th Quintile	10,601	26,856,340	26,856,340	26,856,340	0.00	-0.00
5th Quintile	8,819	26,376,123	26,376,123	26,376,123	0.00	-0.00
Population Density						
Large MSA	20,440	60,318,520	60,318,519	60,318,519	0.00	-0.00
Medium to Small MSA	25,494	50,780,342	50,780,342	50,780,342	0.00	-0.00
Non-MSA, Urban	1,269	2,255,190	2,255,190	2,255,190	0.00	-0.00
Non-MSA, Rural	2,930	5,584,755	5,584,755	5,584,755	0.00	-0.00
Group Quarters						
Group	659	665,602	665,602	665,602	0.00	-0.00
Non-Group	49,474	118,273,204	118,273,204	118,273,204	0.00	-0.00
Household Size						
One	5,768	32,050,378	32,030,125	32,003,504	0.15	0.08
Two	17,500	54,645,502	54,660,454	54,610,438	0.06	0.09
Three	14,554	18,543,320	18,554,954	18,688,007	-0.77	-0.71
Four or More	12,311	13,699,606	13,693,273	13,636,858	0.46	0.41

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

[1] WT1*...*WT10*DUWT11*...*DUWT13 (before QDU poststratification).

[2] WT1*...*WT10*DUWT11*...*DUWT14 (after QDU poststratification).

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

Table G.1 2011 NSDUH Selected QDU-Level Weight Summary Statistics

Domain	*n*	SDU-Level Weights[1] (SDUWT: WT1*...*WT10)						Before sel.qdu.ps[1] (SDUWT*DUWT11)						After sel.qdu.ps[1] (SDUWT*DUWT11*DUWT12)					
		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	61,441	11	409	653	1,039	11,649	1.62	12	543	981	2,096	104,265	2.99	11	540	985	2,116	46,883	2.89
Census Region																			
Northeast	11,997	28	327	649	875	8,047	1.55	28	549	918	2,045	62,767	3.02	26	528	921	2,052	26,059	2.83
South	19,690	11	474	824	1,205	11,649	1.52	12	688	1,188	2,401	104,265	2.90	11	680	1,197	2,437	46,883	2.75
Midwest	17,045	21	447	558	722	6,263	1.36	21	525	745	1,790	23,020	2.55	40	528	761	1,792	22,231	2.50
West	12,709	18	233	660	1,424	10,206	1.84	18	388	1,069	2,181	48,334	3.22	29	383	1,054	2,203	33,865	3.19
Quarter																			
Quarter 1	14,360	12	421	716	1,125	9,758	1.62	12	583	1,057	2,270	104,265	3.05	20	573	1,055	2,291	43,562	2.91
Quarter 2	16,559	11	398	614	960	8,724	1.54	12	501	923	1,943	73,586	2.99	11	503	915	1,953	34,806	2.83
Quarter 3	15,796	16	400	608	1,018	11,649	1.64	16	521	965	2,017	62,767	3.02	21	522	977	2,052	40,244	2.86
Quarter 4	14,726	15	436	666	1,060	10,206	1.65	16	579	991	2,178	35,769	2.87	30	578	997	2,190	46,883	2.92
Household Type																			
12-17, 18-25, 26+	5,958	12	432	675	1,109	10,165	1.65	12	432	675	1,109	10,165	1.65	20	419	668	1,113	10,566	1.70
12-17, 18-25	96	30	290	702	1,101	4,009	1.84	30	290	703	1,101	4,009	1.84	43	267	688	1,083	4,756	1.90
12-17, 26+	17,812	12	389	627	1,000	8,724	1.60	12	391	630	1,006	8,747	1.60	11	381	628	1,013	8,181	1.60
18-25, 26+	12,697	11	436	700	1,141	11,226	1.66	22	512	870	1,419	12,835	1.63	29	507	872	1,432	13,533	1.62
12-17	24	45	208	456	985	1,408	1.52	45	210	458	990	1,408	1.52	42	198	574	957	1,143	1.43
18-25	6,307	15	323	643	977	6,058	1.67	36	380	765	1,197	6,579	1.65	32	383	768	1,195	11,549	1.64
26+	18,547	22	420	648	1,015	11,649	1.54	50	1,860	3,355	5,626	104,265	1.80	47	1,818	3,321	5,686	46,883	1.73
Race/Ethnicity of Householder																			
Hispanic or Latino White	7,747	12	516	802	1,322	8,010	1.42	12	624	1,142	1,866	33,654	2.42	11	610	1,146	1,885	25,646	2.40
Hispanic or Latino Black or African American	149	12	777	2,014	3,299	11,649	1.82	12	1,052	3,080	5,708	104,265	5.05	20	1,166	2,528	5,340	40,244	2.74
Hispanic or Latino Other	418	11	129	496	1,681	8,724	2.96	15	167	760	2,773	29,958	4.09	36	214	844	2,702	33,865	4.12
Non-Hispanic or Latino White	41,252	21	396	630	965	8,073	1.52	21	544	967	2,177	56,218	2.88	26	543	967	2,206	43,562	2.86
Non-Hispanic or Latino Black or African American	7,558	32	454	691	1,075	10,206	1.64	35	570	975	1,970	31,643	2.90	21	564	981	2,009	33,424	2.94
Non-Hispanic or Latino Other	4,317	18	217	554	1,076	7,003	2.01	18	302	864	2,053	73,586	3.84	28	290	841	2,004	46,883	3.43
% Hispanic or Latino in Segment																			
50-100%	4,183	22	619	980	1,454	8,568	1.39	22	738	1,372	2,164	35,769	2.32	45	730	1,354	2,191	33,249	2.39
10-50%	10,536	11	560	849	1,368	10,206	1.54	12	697	1,297	2,560	62,767	2.69	11	696	1,285	2,589	33,865	2.65
<10%	46,722	12	347	602	914	11,649	1.62	12	503	889	1,979	104,265	3.14	20	499	892	1,982	46,883	2.99

Table G.1 2011 NSDUH Selected QDU-Level Weight Summary Statistics (continued)

Domain	*n*	SDU-Level Weights[1] (SDUWT: WT1*...*WT10)						Before sel.qdu.ps[1] (SDUWT*DUWT11)						After sel.qdu.ps[1] (SDUWT*DUWT11*DUWT12)					
		Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
% Black or African American in Segment																			
50-100%	4,960	11	427	669	998	11,649	1.80	12	546	940	1,921	104,265	3.78	11	534	947	1,938	40,244	3.03
10-50%	9,474	21	491	738	1,189	8,724	1.57	34	629	1,119	2,199	73,586	2.85	21	623	1,134	2,253	28,462	2.74
<10%	47,007	18	382	637	1,010	8,568	1.60	18	524	961	2,087	62,767	2.95	26	523	959	2,101	46,883	2.91
% Owner-Occupied DUs[1] in Segment																			
50-100%	46,376	18	411	642	1,018	11,649	1.62	18	543	975	2,117	104,265	3.00	21	541	975	2,129	43,562	2.89
10-50%	11,674	11	392	686	1,068	8,568	1.63	12	539	995	2,015	73,586	3.05	11	534	1,007	2,047	46,883	2.92
<10%	3,391	15	432	709	1,208	7,889	1.60	16	559	1,054	2,081	29,958	2.71	36	545	1,074	2,203	33,865	2.81
Combined Median Rent/Housing Value																			
1st Quintile	10,483	17	327	592	902	11,649	1.71	17	472	866	1,831	104,265	3.28	29	471	876	1,833	40,244	2.92
2nd Quintile	12,594	23	380	620	940	10,206	1.64	23	514	913	1,957	50,367	2.88	29	522	924	1,986	27,792	2.81
3rd Quintile	13,988	11	350	616	981	8,724	1.65	12	504	918	1,997	35,769	2.91	11	495	928	1,994	33,865	2.89
4th Quintile	13,178	19	442	691	1,080	8,573	1.63	19	578	1,036	2,167	73,586	3.22	28	563	1,031	2,165	46,883	3.02
5th Quintile	11,198	32	522	799	1,227	7,653	1.45	32	653	1,215	2,596	33,654	2.62	32	646	1,207	2,645	27,547	2.66
Population Density																			
Large MSA[1]	25,558	11	581	820	1,274	10,206	1.45	12	721	1,272	2,633	73,586	2.61	11	712	1,277	2,651	33,865	2.58
Medium to Small MSA[1]	30,848	19	295	557	841	11,649	1.71	19	438	801	1,762	104,265	3.35	21	437	809	1,757	46,883	3.12
Non-MSA,[1] Urban	1,489	18	213	506	774	3,992	1.68	18	345	753	1,737	18,195	2.81	32	336	735	1,687	16,021	2.91
Non-MSA,[1] Rural	3,546	23	208	457	787	7,877	1.79	23	325	726	1,661	23,764	3.26	31	340	737	1,715	21,646	3.14
Group Quarters																			
Group	687	32	253	651	891	4,274	1.60	50	294	774	1,130	22,096	4.12	42	305	751	1,137	15,523	2.79
Non-Group	60,754	11	410	654	1,040	11,649	1.62	12	545	985	2,112	104,265	2.98	11	542	990	2,136	46,883	2.88
Household Size																			
One	7,324	15	353	619	958	11,649	1.64	45	975	2,683	6,179	104,265	2.29	42	982	2,712	6,172	46,883	2.16
Two	22,122	15	403	644	1,003	11,226	1.58	16	679	1,471	3,382	62,767	2.23	29	691	1,466	3,329	25,729	2.20
Three	17,557	11	422	658	1,044	8,573	1.60	19	462	751	1,289	27,148	2.10	28	459	751	1,287	24,091	2.11
Four or More	14,438	12	425	687	1,127	10,206	1.66	12	439	719	1,222	12,835	1.77	11	427	703	1,226	16,322	1.87

[1] DU = dwelling unit, MSA = metropolitan statistical area, ps = poststratification adjustment, QDU = questionnaire dwelling unit, SDU = screener dwelling unit, sel = selected.
[2] Q1 and Q3 refer to the first and third quartile of the weight distribution.
[3] Unequal weighting effect (UWE) is defined as $1 + [(n - 1)/n]*CV^2$, where CV = coefficient of variation of weights.

Table G.2 2011 NSDUH Respondent QDU-Level Weight Summary Statistics

Domain	*n*	Before res.qdu.nr[1] (SDUWT*DUWT11*DUWT12)						After res.qdu.nr[1] (SDUWT*DUWT11*...*DUWT13)						Final Weight: After res.qdu.ps[1] (SDUWT*DUWT11*...*DUWT14)					
		Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
Total	50,133	11	521	943	1,948	46,883	2.97	11	608	1,128	2,411	51,507	3.22	9	608	1,130	2,413	52,502	3.22
Census Region																			
Northeast	9,456	26	498	881	1,829	25,221	2.91	28	575	1,089	2,349	51,507	3.28	27	572	1,089	2,365	52,502	3.28
South	16,487	11	658	1,147	2,215	46,883	2.83	11	744	1,338	2,688	45,483	3.00	9	743	1,338	2,689	47,840	3.00
Midwest	13,752	40	518	737	1,651	22,231	2.56	44	613	893	2,047	29,765	2.75	41	613	894	2,045	29,668	2.75
West	10,438	29	365	999	2,039	33,865	3.31	30	414	1,188	2,491	46,269	3.70	28	415	1,187	2,484	48,746	3.71
Quarter																			
Quarter 1	11,633	20	551	1,007	2,075	37,183	3.00	21	645	1,221	2,567	51,507	3.30	20	645	1,223	2,567	52,502	3.30
Quarter 2	13,462	11	486	878	1,804	34,806	2.93	11	568	1,054	2,229	44,926	3.16	9	567	1,054	2,235	44,661	3.17
Quarter 3	12,992	21	508	936	1,916	40,244	2.91	21	590	1,127	2,375	46,269	3.13	18	589	1,128	2,376	48,746	3.13
Quarter 4	12,046	30	561	958	2,013	46,883	2.99	30	651	1,131	2,466	41,667	3.24	27	650	1,135	2,477	39,840	3.24
Household Type																			
12-17, 18-25, 26+	5,208	20	407	666	1,119	10,566	1.72	21	451	775	1,288	10,647	1.70	18	450	774	1,293	9,854	1.68
12-17, 18-25	86	43	252	676	1,081	4,756	1.95	45	295	836	1,225	4,868	1.85	44	293	823	1,210	4,820	1.86
12-17, 26+	15,210	11	378	627	1,018	8,181	1.61	11	428	742	1,182	9,384	1.61	9	428	743	1,189	8,350	1.59
18-25, 26+	10,373	29	497	869	1,432	13,533	1.63	29	587	1,072	1,740	14,122	1.66	27	586	1,069	1,748	13,842	1.65
12-17	18	56	238	603	985	1,143	1.35	212	328	716	1,107	1,528	1.31	208	327	694	1,079	1,528	1.30
18-25	5,450	32	377	762	1,186	11,549	1.63	33	437	867	1,398	11,387	1.67	34	437	869	1,398	11,382	1.66
26+	13,788	47	1,755	3,275	5,641	46,883	1.76	47	2,331	4,434	7,644	51,507	1.78	48	2,323	4,426	7,651	52,502	1.78
Race/Ethnicity of Householder																			
Hispanic or Latino White	6,528	11	601	1,116	1,807	19,967	2.34	11	686	1,292	2,152	32,838	2.64	9	688	1,289	2,150	32,822	2.65
Hispanic or Latino Black or African American	122	20	1,250	2,568	5,289	40,244	2.79	46	1,415	2,935	6,072	45,052	2.97	114	1,513	2,920	5,721	47,840	3.01
Hispanic or Latino Other	358	36	208	789	2,511	33,865	4.44	37	237	888	3,005	45,483	5.17	37	237	872	3,009	46,513	5.15
Non-Hispanic or Latino White	33,138	26	522	922	2,002	35,324	2.96	32	620	1,126	2,537	51,507	3.16	31	620	1,127	2,542	52,502	3.16
Non-Hispanic or Latino Black or African American	6,532	21	551	953	1,884	33,424	2.99	21	608	1,071	2,185	46,269	3.30	18	605	1,072	2,179	48,746	3.32
Non-Hispanic or Latino Other	3,455	28	263	789	1,846	46,883	3.57	28	319	967	2,324	41,667	3.93	27	315	964	2,382	39,840	3.94
% Hispanic or Latino in Segment																			
50-100%	3,519	45	715	1,331	2,093	33,249	2.42	46	821	1,507	2,481	45,483	2.70	45	817	1,506	2,483	42,552	2.71
10-50%	8,590	11	673	1,235	2,389	33,865	2.72	11	775	1,458	2,994	51,507	3.06	9	775	1,457	3,012	52,502	3.06
<10%	38,024	20	479	853	1,816	46,883	3.08	21	557	1,029	2,252	46,269	3.29	18	555	1,029	2,258	48,746	3.29

Table G.2 2011 NSDUH Respondent QDU-Level Weight Summary Statistics (continued)

Domain	*n*	Before res.qdu.nr[1] (SDUWT*DUWT11*DUWT12)						After res.qdu.nr[1] (SDUWT*DUWT11*...*DUWT13)						Final Weight: After res.qdu.ps[1] (SDUWT*DUWT11*...*DUWT14)					
		Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
% Black or African American in Segment																			
50-100%	4,266	11	516	907	1,796	40,244	3.08	11	583	1,036	2,114	45,052	3.32	9	581	1,037	2,122	47,840	3.33
10-50%	7,936	21	608	1,094	2,126	28,462	2.81	21	683	1,260	2,590	44,841	3.07	18	679	1,258	2,597	46,513	3.08
<10%	37,931	26	503	915	1,928	46,883	2.99	28	593	1,113	2,403	51,507	3.24	27	592	1,116	2,406	52,502	3.24
% Owner-Occupied DUs[1] in Segment																			
50-100%	37,607	21	523	931	1,952	40,244	2.97	21	615	1,124	2,435	46,269	3.22	18	614	1,125	2,440	48,746	3.22
10-50%	9,661	11	507	962	1,904	46,883	3.01	11	584	1,145	2,299	51,507	3.27	9	583	1,147	2,299	52,502	3.27
<10%	2,865	36	534	1,030	2,073	33,865	2.83	37	584	1,144	2,478	40,553	3.14	37	576	1,135	2,465	40,365	3.17
Combined Median Rent/Housing Value																			
1st Quintile	8,834	29	451	847	1,723	40,244	3.01	30	499	995	2,034	45,483	3.20	28	498	992	2,039	47,840	3.21
2nd Quintile	10,444	29	507	891	1,823	27,792	2.89	29	585	1,061	2,220	38,774	3.11	27	584	1,061	2,222	39,165	3.10
3rd Quintile	11,435	11	479	886	1,879	33,865	2.96	11	561	1,066	2,315	46,269	3.19	9	561	1,064	2,314	48,746	3.20
4th Quintile	10,601	28	549	985	1,984	46,883	3.13	28	643	1,182	2,493	51,507	3.39	27	642	1,185	2,494	52,502	3.39
5th Quintile	8,819	32	624	1,148	2,430	27,547	2.72	35	763	1,416	3,176	39,406	2.95	33	764	1,415	3,171	39,534	2.95
Population Density																			
Large MSA[1]	20,440	11	693	1,227	2,419	33,865	2.64	11	826	1,471	3,146	51,507	2.91	9	825	1,472	3,146	52,502	2.91
Medium to Small MSA[1]	25,494	21	420	779	1,615	46,883	3.22	21	478	928	1,979	45,052	3.41	18	477	927	1,986	47,840	3.40
Non-MSA,[1] Urban	1,269	32	311	685	1,625	15,676	2.96	32	352	805	1,866	18,777	3.13	30	354	798	1,868	18,649	3.13
Non-MSA,[1] Rural	2,930	31	336	718	1,663	21,646	3.16	31	378	842	1,983	27,276	3.28	29	379	840	1,997	27,107	3.27
Group Quarters																			
Group	659	43	312	759	1,137	15,523	2.81	45	315	777	1,193	15,523	2.74	44	316	775	1,186	15,509	2.76
Non-Group	49,474	11	523	948	1,969	46,883	2.96	11	611	1,135	2,440	51,507	3.21	9	611	1,139	2,442	52,502	3.21
Household Size																			
One	5,768	56	910	2,419	5,905	46,883	2.26	87	1,119	3,119	7,800	51,507	2.31	93	1,119	3,120	7,804	52,502	2.31
Two	17,500	29	655	1,360	3,143	24,457	2.26	30	774	1,658	4,189	41,689	2.42	29	775	1,663	4,189	38,929	2.42
Three	14,554	28	448	738	1,262	15,391	2.04	28	512	885	1,515	30,137	2.31	27	512	887	1,516	25,786	2.30
Four or More	12,311	11	422	697	1,216	16,322	1.85	11	475	822	1,417	19,498	1.99	9	473	823	1,421	20,096	1.99

[1] DU = dwelling unit, MSA = metropolitan statistical area, nr = nonresponse adjustment, ps = poststratification adjustment, QDU = questionnaire dwelling unit, res = respondent, SDU = screener dwelling unit.
[2] Q1 and Q3 refer to the first and third quartile of the weight distribution.
[3] Unequal weighting effect (UWE) is defined as $1 + [(n - 1)/n]*CV^2$, where CV = coefficient of variation of weights.

Appendix H: GEM Modeling Summary for the Pair Weights

Appendix H: GEM Modeling Summary for the Pair Weights

Introduction

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

For 2011, modeling involved taking two model groups through four adjustment steps: (1) selected pair poststratification, (2) pair nonresponse adjustment, (3) responding pair poststratification, and (4) responding pair extreme value adjustment.

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

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

Final Model Explanatory Variables

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

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

Group Quarter Indicator

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

Household Size

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

Pair Age (15 Levels)

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

Pair Age (6 Levels)

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

Pair Age (3 Levels)

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

Pair Gender

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

Pair Race/Ethnicity (10 Levels)

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

Pair Race/Ethnicity (5 Levels)

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

Pair Race/Ethnicity (4 Levels)

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

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

1: 50-100%[1], 2: 10->50%, 3: 0->10%

Percentage of Segments That Are Black or African American

1: 50-100%, 2: 10->50%, 3: 0->10%[1]

Percentage of Segments That Are Hispanic or Latino

1: 50-100%, 2: 10->50%, 3: 0->10%[1]

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

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

Population Density

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

Quarter

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

Race/Ethnicity of Householder

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

State/Region

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

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

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

States[3]

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

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

Pair Relationship Associated with Multiplicity

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

DU = dwelling unit, MSA = metropolitan statistical area.

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

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

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

* The pair member focused on.

Exhibit H.2 Covariates for 2011 NSDUH Pair Weights

Variables	Level	Proposed
One-Factor Effects		
Intercept	1	1
State	Model-specific	
Quarter	4	3
Population Density	3	2
Group Quarter	3	2
Household Size	3	2
Pair Age	15	14
Pair Gender	4	2
Pair Race/Ethnicity	10	9
Race/Ethnicity of Householder	6	5
Rent/Housing	5	4
Segment % Black or African American	3	2
Segment % Hispanic or Latino	3	2
% Owner-Occupied	3	2
Pair Relationship	Model-specific	
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

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

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

Table H.1a 2011 Pair Weight GEM Modeling Summary (Model Group 1: Northeast and South)

Modeling Step[1]	Extreme Weight Proportions			UWE[2]	# Covariates[3]	Bounds[4]	
	% Unweighted	% Weighted	% Winsorized			Nominal	Realized
sel.pr.ps	5.29	22.46	10.48	56.5114	213	(0.23, 1.90)	(0.23, 1.90)
	2.64	9.52	1.84	8.7323	199	(0.20, 3.42)	(0.20, 3.42)
						(0.99, 1.85)	(0.99, 1.85)
res.pr.nr	2.48	10.50	2.19	8.7323	213	(1.00, 1.60)	(1.00, 1.60)
	2.05	9.58	2.32	9.2240	213	(1.00, 5.00)	(1.00, 5.00)
						(1.50, 4.17)	(3.97, 4.17)
res.pr.ps	2.03	8.81	2.29	9.2240	223	(0.33, 1.10)	(0.33, 1.10)
	1.44	4.56	0.94	8.9416	211	(0.26, 4.08)	(0.27, 4.07)
						N/A	N/A
res.pr.ev	1.44	4.56	0.94	8.9416	223	(0.97, 2.18)	(0.98, 2.18)
	1.08	3.02	0.44	8.9492	211	(0.87, 2.27)	(0.91, 2.26)
						N/A	N/A

GEM = generalized exponential model.

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

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

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

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

Table H.1b 2011 Distribution of Weight Adjustment Factors and Weight Products (Model Group 1: Northeast and South)

	SDU Weight	Pair Selection Prob		sel.pr.ps[1]		res.pr.nr[1]		res.pr.ps[1]		res.pr.ev[1]	
	1-10	pairwt11	1-11	pairwt12	1-12	pairwt13	1-13	pairwt14	1-14	pairwt15	1-15
Minimum	11	1.02	14	0.05	7	0.29	7	0.14	8	0.45	8
1%	63	1.11	147	0.20	94	0.86	90	0.34	68	0.89	67
5%	137	1.20	348	0.34	242	1.00	247	0.46	204	0.96	202
10%	215	1.30	588	0.46	430	1.00	449	0.56	372	0.97	367
25%	444	1.53	1,224	0.70	1,035	1.01	1,099	0.76	983	0.99	984
Median	758	5.34	3,036	1.00	2,916	1.10	3,199	0.95	2,909	1.00	2,901
75%	1,120	11.46	8,040	1.33	8,223	1.36	9,477	1.12	9,083	1.01	9,102
90%	1,589	22.66	17,108	1.71	20,489	1.92	27,507	1.40	27,040	1.03	26,861
95%	1,987	28.96	28,005	1.98	35,023	2.53	49,324	1.75	51,338	1.05	51,990
99%	3,242	56.17	58,490	2.73	78,149	4.08	128,263	2.45	136,251	1.08	137,978
Maximum	8,047	2,051.88	6,644,771	4.06	1,046,187	6.76	1,471,845	4.07	1,141,198	2.26	1,135,896
n	13,686	-	13,686	-	13,686	-	10,127	-	10,127	-	10,127
Mean	868	9.20	8,061	1.06	8,578	1.33	11,593	0.99	11,593	1.00	11,593
Max/Mean	9	-	824	-	122	-	127	-	98	-	98

SDU = screener dwelling unit.

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

Model Group 1 Overview

Selected Pair-Level Poststratification

In the selected pair-level poststratification step, 199 of 213 proposed factors were retained in the final model. Of the 76 proposed main effects, 74 were included in the model. Group Quarter levels "college dorm" and "other group quarter" were combined. Race/Ethnicity of Householder was reduced by combining "Hispanic or Latino Black or African American" with "Hispanic or Latino Other" due to convergence problems. In addition, all 125 proposed two-factor effects were included in the model. All three-factor effects were dropped due to convergence problems.

Respondent Pair-Level Nonresponse

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

Respondent Pair-Level Poststratification

In the respondent pair-level poststratification step, 211 of 223 proposed factors were retained in the final model. All main and two-factor effects were retained. All three-factor effects were dropped due to convergence problems.

Respondent Pair-Level Extreme Value Adjustment

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

Exhibit H.1.1 Covariates for 2011 NSDUH Pair Weights (sel.pr.ps) Model Group 1: Northeast and South

Variables	Level	Proposed	Final	Comments
One-Factor Effects		**76**	**74**	
Intercept	1	1	1	All levels present.
State	26	25	25	All levels present.
Quarter	4	3	3	All levels present.
Population Density	4	3	3	All levels present.
Group Quarter	3	2	1	Coll. (1) & (2); conv.
Household Size	3	2	2	All levels present.
Pair Age	15	14	14	All levels present.
Pair Gender	3	2	2	All levels present.
Pair Race/Ethnicity	10	9	9	All levels present.
Race/Ethnicity of Householder	6	5	4	Coll. (2) & (3); conv.
Rent/Housing	5	4	4	All levels present.
Segment % Black or African American	3	2	2	All levels present.
Segment % Hispanic or Latino	3	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
Two-Factor Effects		**125**	**125**	
Pair Race/Ethnicity (5 Levels) × Pair Age (6 Levels)	5 × 6	20	20	All levels present.
Pair Race/Ethnicity (5 Levels) × Pair Gender	5 × 3	8	8	All levels present.
Pair Gender × Pair Age (6 Levels)	3 × 6	10	10	All levels present.
State/Region × Pair Race/Ethnicity (5 Levels)	6 × 5	20	20	All levels present.
State/Region × Pair Age (6 Levels)	6 × 6	25	25	All levels present.
State/Region × Pair Gender	6 × 3	10	10	All levels present.
Rent/Housing × % Black or African American	5 × 3	8	8	All levels present.
Rent/Housing × % Hispanic or Latino	5 × 3	8	8	All levels present.
Rent/Housing × % Owner-Occupied	5 × 3	8	8	All levels present.
% Owner-Occupied × % Black or African American	3 × 3	4	4	All levels present.
% Owner-Occupied × % Hispanic or Latino	3 × 3	4	4	All levels present.
Three-Factor Effects		**12**	**0**	
Pair Race/Ethnicity (4 Levels) × Pair Gender × Pair Age (3 Levels)	4 × 3 × 3	12	0	Drop all; conv.
Total		**213**	**199**	

Exhibit H.1.2 Covariates for 2011 NSDUH Pair Weights (res.pr.nr) Model Group 1: Northeast and South

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

Exhibit H.1.3 Covariates for 2011 NSDUH Pair Weights (res.pr.ps) Model Group 1: Northeast and South

Variables	Level	Proposed	Final	Comments
One-Factor Effects		**86**	**86**	
Intercept	1	1	1	All levels present.
State	26	25	25	All levels present.
Quarter	4	3	3	All levels present.
Population Density	4	3	3	All levels present.
Group Quarter	3	2	2	All levels present.
Household Size	3	2	2	All levels present.
Pair Age	15	14	14	All levels present.
Pair Gender	3	2	2	All levels present.
Pair Race/Ethnicity	10	9	9	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Rent/Housing	5	4	4	All levels present.
Segment % Black or African American	3	2	2	All levels present.
Segment % Hispanic or Latino	3	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
Pair Relationship	10	10	10	All levels present.
Two-Factor Effects		**125**	**125**	
Pair Race/Ethnicity (5 Levels) × Pair Age (6 Levels)	5 × 6	20	20	All levels present.
Pair Race/Ethnicity (5 Levels) × Pair Gender	5 × 3	8	8	All levels present.
Pair Gender × Pair Age (6 Levels)	3 × 6	10	10	All levels present.
State/Region × Pair Race/Ethnicity (5 Levels)	6 × 5	20	20	All levels present.
State/Region × Pair Age (6 Levels)	6 × 6	25	25	All levels present.
State/Region × Pair Gender	6 × 3	10	10	All levels present.
Rent/Housing × % Black or African American	5 × 3	8	8	All levels present.
Rent/Housing × % Hispanic or Latino	5 × 3	8	8	All levels present.
Rent/Housing × % Owner-Occupied	5 × 3	8	8	All levels present.
% Owner-Occupied × % Black or African American	3 × 3	4	4	All levels present.
% Owner-Occupied × % Hispanic or Latino	3 × 3	4	4	All levels present.
Three-Factor Effects		**12**	**0**	
Pair Race/Ethnicity (4 Levels) × Pair Gender × Pair Age (3 Levels)	4 × 3 × 3	12	0	Drop all; conv.
Total		**223**	**211**	

Exhibit H.1.4 Covariates for 2011 NSDUH Pair Weights (res.pr.ev) Model Group 1: Northeast and South

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

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

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

Table H.2a 2011 Pair Weight GEM Modeling Summary (Model Group 2: Midwest and West)

Modeling Step[1]	Extreme Weight Proportions			UWE[2]	# Covariates[3]	Bounds[4]	
	% Unweighted	% Weighted	% Winsorized			Nominal	Realized
sel.pr.ps	5.53	25.70	12.03	27.6797	212	(0.29, 2.40)	(0.30, 2.40)
	3.53	12.37	2.43	9.9334	200	(0.20, 4.32)	(0.21, 4.31)
						(0.99, 1.71)	(0.99, 1.71)
res.pr.nr	3.18	12.21	1.9%	10.7473	212	(1.01, 2.50)	(1.01, 2.50)
	3.24	14.38	3.53	13.8499	212	(1.00, 5.00)	(1.00, 4.98)
						N/A	N/A
res.pr.ps	3.25	14.50	3.81	13.8499	212	(0.33, 1.60)	(0.34, 1.60)
	2.31	10.86	1.67	11.5371	210	(0.26, 1.77)	(0.26, 1.75)
						N/A	N/A
res.pr.ev	2.31	10.86	1.67	11.5371	212	(0.94, 2.00)	(0.99, 1.81)
	0.85	6.37	0.86	11.3802	210	(0.71, 2.00)	(0.87, 1.19)
						N/A	N/A

GEM = generalized exponential model.

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

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

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

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

Table H.2b 2011 Distribution of Weight Adjustment Factors and Weight Products (Model Group 2: Midwest and West)

	SDU Weight	Pair Selection		sel.pr.ps[1]		res.pr.nr[1]		res.pr.ps[1]		res.pr.ev[1]	
	1-10	pairwt11	1-11	pairwt12	1-12	pairwt13	1-13	pairwt14	1-14	pairwt15	1-15
Minimum	18	1.02	22	0.11	26	0.60	27	0.17	18	0.63	16
1%	75	1.09	142	0.29	133	0.90	148	0.35	103	0.88	97
5%	115	1.21	283	0.49	267	1.00	290	0.54	240	0.94	233
10%	158	1.30	514	0.59	467	1.01	510	0.66	440	0.95	434
25%	406	1.47	916	0.79	922	1.06	1,052	0.86	1,028	0.98	1,019
Median	585	5.11	2,504	1.04	2,493	1.16	2,820	1.02	2,862	1.00	2,864
75%	982	10.91	6,752	1.32	6,666	1.41	8,008	1.15	7,908	1.02	7,898
90%	1,636	20.93	15,294	1.67	17,111	1.86	22,535	1.31	22,549	1.04	22,728
95%	2,008	27.15	24,079	1.94	27,915	2.29	39,846	1.41	39,345	1.05	39,616
99%	2,875	52.32	59,354	2.49	70,008	3.53	113,708	1.58	115,142	1.08	117,003
Maximum	10,206	2,192.84	2,681,904	4.31	1,047,876	4.98	2,063,702	1.75	1,232,038	1.22	1,203,872
n	13,409	-	13,409	-	13,409	-	9,849	-	9,849	-	9,849
Mean	776	9.12	7,296	1.10	7,503	1.34	10,215	1.00	10,215	1.00	10,215
Max/Mean	13	-	368	-	140	-	202	-	121	-	118

SDU = screener dwelling unit.

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

Model Group 2 Overview

Selected Pair-Level Poststratification

In the selected pair-level poststratification step, 200 of 212 proposed factors were retained in the final model. All main and two-factor effects were retained at proposed levels. None of the 12 three-factor effects were kept in the model due to convergence problems.

Respondent Pair-Level Nonresponse

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

Respondent Pair-Level Poststratification

In the respondent pair-level poststratification step, 210 of 222 proposed factors were retained in the final model, as in the selected pair-level poststratification step.

Respondent Pair-Level Extreme Value Adjustment

In the respondent pair-level extreme value adjustment step, 210 of 222 proposed factors were retained in the final model, as in the respondent pair-level poststratification step.

Exhibit H.2.1 Covariates for 2011 NSDUH Pair Weights (sel.pr.ps) Model Group 2: Midwest and West

Variables	Level	Proposed	Final	Comments
One-Factor Effects		**75**	**75**	
Intercept	1	1	1	All levels present.
State	25	24	24	All levels present.
Quarter	4	3	3	All levels present.
Population Density	4	3	3	All levels present.
Group Quarter	3	2	2	All levels present.
Household Size	3	2	2	All levels present.
Pair Age	15	14	14	All levels present.
Pair Gender	3	2	2	All levels present.
Pair Race/Ethnicity	10	9	9	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Rent/Housing	5	4	4	All levels present.
Segment % Black or African American	3	2	2	All levels present.
Segment % Hispanic or Latino	3	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
Two-Factor Effects		**125**	**125**	
Pair Race/Ethnicity (5 Levels) × Pair Age (6 Levels)	5 × 6	20	20	All levels present.
Pair Race/Ethnicity (5 Levels) × Pair Gender	5 × 3	8	8	All levels present.
Pair Gender × Pair Age (6 Levels)	3 × 6	10	10	All levels present.
State/Region × Pair Race/Ethnicity (5 Levels)	6 × 5	20	20	All levels present.
State/Region × Pair Age (6 Levels)	6 × 6	25	25	All levels present.
State/Region × Pair Gender	6 × 3	10	10	All levels present.
Rent/Housing × % Black or African American	5 × 3	8	8	All levels present.
Rent/Housing × % Hispanic or Latino	5 × 3	8	8	All levels present.
Rent/Housing × % Owner-Occupied	5 × 3	8	8	All levels present.
% Owner-Occupied × % Black or African American	3 × 3	4	4	All levels present.
% Owner-Occupied × % Hispanic or Latino	3 × 3	4	4	All levels present.
Three-Factor Effects		**12**	**0**	
Pair Race/Ethnicity (4 Levels) × Pair Gender × Pair Age (3 Levels)	4 × 3 × 3	12	0	Drop all; conv.
Total		**212**	**200**	

Exhibit H.2.2 Covariates for 2011 NSDUH Pair Weights (res.pr.nr) Model Group 2: Midwest and West

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

Exhibit H.2.3 Covariates for 2011 NSDUH Pair Weights (res.pr.ps) Model Group 2: Midwest and West

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

Exhibit H.2.4 Covariates for 2011 NSDUH Pair Weights (res.pr.ev) Model Group 2: Midwest and West

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

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

Table I.1 2011 NSDUH Person Pair-Level Response Rates

Domain	Selected Pairs	Respondent Pairs	% Interview Response Rate[1]
Total	27,095	19,976	65.49
Pair Age Group			
12-17, 12-17	4,649	3,886	85.32
12-17, 18-25	3,756	2,934	79.24
12-17, 26-34	834	674	80.79
12-17, 35-49	3,855	2,938	76.54
12-17, 50+	766	561	72.32
18-25, 18-25	5,476	4,084	75.57
18-25, 26-34	1,049	718	66.13
18-25, 35-49	1,469	995	68.74
18-25, 50+	1,057	663	61.33
26-34, 26-34	858	583	65.35
26-34, 35-49	492	305	60.13
26-34, 50+	302	179	65.99
35-49, 35-49	748	477	57.74
35-49, 50+	477	273	56.81
50+, 50+	1,307	706	51.00
Pair Race/Ethnicity			
Hispanic or Latino	4,001	3,052	68.74
Black or African American	2,912	2,343	69.74
White	16,243	11,729	65.23
Other	1,770	1,204	52.23
White & Black or African American	236	182	75.85
White & Hispanic or Latino	858	639	65.08
White & Other	732	553	66.64
Black or African American & Hispanic or Latino	111	90	74.97
Black or African American & Other	97	79	72.32
Hispanic or Latino & Other	135	105	44.32
Pair Gender			
Male, Male	6,020	4,381	64.22
Female, Female	5,895	4,549	69.61
Male, Female	15,180	11,046	64.70
Household Size			
Two	6,903	4,731	59.30
Three	7,438	5,492	64.60
Four or More	12,754	9,753	69.10

Table I.1 2011 NSDUH Person Pair-Level Response Rates (continued)

Domain	Selected Pairs	Respondent Pairs	% Interview Response Rate[1]
Census Region			
Northeast	5,254	3,634	56.74
South	8,432	6,493	70.30
Midwest	7,525	5,506	64.90
West	5,884	4,343	65.61
Quarter			
Quarter 1	6,282	4,662	66.35
Quarter 2	7,314	5,340	65.53
Quarter 3	7,038	5,199	66.84
Quarter 4	6,461	4,775	63.22
% Hispanic or Latino in Segment			
50-100%	2,128	1,624	69.38
10-50%	4,792	3,504	62.63
<10%	20,175	14,848	65.82
% Black or African American in Segment			
50-100%	2,101	1,668	71.45
10-50%	4,075	3,127	68.41
<10%	20,919	15,181	64.23
% Owner-Occupied DUs in Segment			
50-100%	20,773	15,244	65.33
10-50%	4,996	3,728	65.75
<10%	1,326	1,004	70.72
Combined Median Rent/Housing Value			
1st Quintile	4,650	3,589	73.55
2nd Quintile	5,587	4,211	69.58
3rd Quintile	6,126	4,600	68.45
4th Quintile	5,804	4,155	60.69
5th Quintile	4,928	3,421	58.32
Population Density			
Large MSA	11,353	8,111	63.27
Medium to Small MSA	13,593	10,252	67.91
Non-MSA, Urban	592	458	75.37
Non-MSA, Rural	1,557	1,155	67.35
Group Quarters			
Group	349	288	80.00
Non-Group	26,746	19,688	65.44

DU = dwelling unit, MSA = metropolitan statistical area.

[1] The weight used for calculating the response rate includes screener dwelling unit (SDU)- and pair-level design weights, SDU nonresponse and poststratification adjustments, and selected pair poststratification adjustment. This weight is the product of WT1*…*WT10*PRWT11*PRWT12.

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

Table J.1 2011 NSDUH Selected Pair-Level Proportions of Extreme Values and Outwinsors

Domain	*n*	SDU-Level Weights[1] (SDUWT: WT1*...*WT10)			Before sel.pr.ps[1] (SDUWT*PRWT11)			After sel.pr.ps[1] (SDUWT*PRWT11*PRWT12)		
		% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]
Total	27,095	2.13	5.80	1.67	5.56	28.23	16.74	3.24	17.00	7.16
Pair Age Group										
12-17, 12-17	4,649	1.40	4.29	1.34	3.79	17.35	6.60	0.92	5.79	1.12
12-17, 18-25	3,756	1.84	5.29	1.34	7.37	29.69	12.09	2.10	8.32	1.31
12-17, 26-34	834	3.00	6.35	2.38	2.40	8.61	1.52	1.44	6.03	0.88
12-17, 35-49	3,855	1.84	5.18	1.49	2.39	14.26	6.37	0.91	4.55	0.78
12-17, 50+	766	1.96	5.40	1.53	2.61	16.43	7.59	1.04	5.13	0.70
18-25, 18-25	5,476	2.57	6.56	1.73	9.13	32.30	13.35	5.75	17.78	2.74
18-25, 26-34	1,049	4.77	12.18	3.71	4.10	19.36	7.63	2.57	12.28	2.96
18-25, 35-49	1,469	2.31	6.50	2.09	6.60	31.74	13.63	3.88	11.37	1.47
18-25, 50+	1,057	1.51	4.91	1.97	3.60	17.13	6.34	2.74	7.66	1.25
26-34, 26-34	858	3.03	6.64	1.22	3.03	14.34	5.30	6.29	23.07	9.22
26-34, 35-49	492	4.27	10.88	3.54	8.13	49.56	34.25	8.33	36.73	18.74
26-34, 50+	302	3.31	6.83	1.28	4.64	26.38	19.51	1.66	13.84	10.09
35-49, 35-49	748	2.01	6.90	2.41	7.89	61.52	47.20	7.22	46.35	30.25
35-49, 50+	477	1.68	4.83	1.31	5.03	52.53	44.49	2.31	16.94	11.18
50+, 50+	1,307	0.92	1.79	0.52	6.20	21.26	11.06	8.26	24.20	8.11
Pair Race/Ethnicity										
Hispanic or Latino	4,001	3.50	9.92	3.52	6.10	31.65	17.80	3.65	21.50	9.43
Black or African American	2,912	2.71	6.43	1.44	7.04	30.37	15.55	5.08	24.56	8.79
White	16,243	0.72	1.66	0.36	4.48	19.59	10.70	2.17	13.06	6.19
Other	1,770	5.93	14.54	3.52	8.19	49.01	30.16	5.37	21.84	8.40
White & Black or African American	236	9.32	16.68	3.78	12.29	32.50	11.51	6.36	17.25	2.37
White & Hispanic or Latino	858	4.08	11.15	3.35	7.23	21.52	7.14	5.83	18.15	3.87
White & Other	732	4.23	10.57	3.60	7.38	21.39	6.84	5.46	11.35	2.65
Black or African American & Hispanic or Latino	111	19.82	46.47	18.72	21.62	93.46	87.89	18.92	52.98	22.34
Black or African American & Other	97	4.12	4.29	1.16	3.09	2.47	0.31	2.06	1.59	0.40
Hispanic or Latino & Other	135	17.04	35.45	9.44	9.63	51.39	29.16	6.67	54.75	26.42
Pair Gender										
Male, Male	6,020	2.28	6.08	1.75	7.34	22.30	8.71	4.02	14.38	3.11
Female, Female	5,895	2.12	5.59	1.42	5.60	22.78	10.39	3.14	12.08	3.52
Male, Female	15,180	2.08	5.77	1.74	4.84	31.22	20.52	2.97	19.14	9.34
Household Size										
Two	6,903	2.16	5.44	1.49	1.00	3.41	1.00	0.65	1.88	0.37
Three	7,438	2.00	5.24	1.71	2.64	30.79	24.20	3.05	18.52	7.01
Four or More	12,754	2.20	6.28	1.74	9.74	39.98	20.98	4.75	23.90	10.70

Table J.1 2011 NSDUH Selected Pair-Level Proportions of Extreme Values and Outwinsors (continued)

Domain	*n*	SDU-Level Weights[1] (SDUWT: WT1*...*WT10)			Before sel.pr.ps[1] (SDUWT*PRWT11)			After sel.pr.ps[1] (SDUWT*PRWT11*PRWT12)		
		% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]
Census Region										
Northeast	5,254	2.26	6.59	1.93	5.41	20.98	9.24	3.41	14.50	4.14
South	8,432	2.21	4.99	1.29	5.41	28.51	18.16	2.36	15.39	7.53
Midwest	7,525	1.57	4.11	1.11	5.51	26.73	15.65	3.75	18.26	7.54
West	5,884	2.63	7.93	2.57	5.98	33.72	20.44	3.70	20.03	8.52
Quarter										
Quarter 1	6,282	2.56	6.71	1.87	6.24	28.54	16.68	3.53	18.76	8.90
Quarter 2	7,314	1.63	4.74	1.35	4.89	21.07	9.82	2.73	14.41	5.26
Quarter 3	7,038	1.80	4.42	1.29	5.23	34.42	23.68	2.97	16.86	7.01
Quarter 4	6,461	2.65	7.35	2.19	6.02	27.87	15.64	3.82	18.01	7.50
% Hispanic or Latino in Segment										
50-100%	2,128	2.35	6.46	2.24	6.02	33.50	18.36	4.65	26.46	10.85
10-50%	4,792	3.01	8.35	2.64	6.49	27.64	12.84	4.24	14.79	4.41
<10%	20,175	1.90	4.81	1.25	5.29	27.58	17.70	2.86	16.16	7.48
% Black or African American in Segment										
50-100%	2,101	3.05	8.81	3.16	8.33	39.67	23.68	5.52	25.58	9.81
10-50%	4,075	2.82	7.59	2.20	6.16	34.01	24.24	4.69	15.82	4.86
<10%	20,919	1.91	5.05	1.39	5.17	25.56	14.16	2.73	16.37	7.39
% Owner-Occupied DUs[1] in Segment										
50-100%	20,773	1.84	5.24	1.48	5.19	28.56	17.77	2.87	16.68	7.38
10-50%	4,996	2.98	7.23	2.20	7.01	28.12	13.33	5.28	19.91	6.69
<10%	1,326	3.47	8.42	2.51	5.96	21.75	9.10	1.28	4.43	1.25
Combined Median Rent/Housing Value										
1st Quintile	4,650	2.60	6.31	1.87	6.24	27.82	13.66	3.31	15.90	6.82
2nd Quintile	5,587	1.72	5.14	1.69	5.76	38.89	28.15	3.45	21.22	9.49
3rd Quintile	6,126	1.99	5.70	1.63	5.37	23.85	12.43	2.89	16.43	6.30
4th Quintile	5,804	2.29	6.39	1.84	5.57	20.21	7.94	3.67	13.58	3.46
5th Quintile	4,928	2.15	5.51	1.40	4.93	28.81	18.92	2.86	18.11	9.97
Population Density										
Large MSA[1]	11,353	2.55	6.86	2.04	5.83	28.65	16.36	3.48	17.77	7.82
Medium to Small MSA[1]	13,593	1.85	4.78	1.32	5.47	27.84	17.24	3.27	15.90	5.77
Non-MSA,[1] Urban	592	1.52	2.01	0.54	4.56	26.94	14.50	1.69	23.25	16.10
Non-MSA,[1] Rural	1,557	1.80	4.29	1.17	4.82	27.44	17.35	1.80	15.38	8.59
Group Quarters										
Group	349	8.88	18.21	3.74	12.89	43.49	19.98	15.76	38.34	7.75
Non-Group	26,746	2.05	5.65	1.65	5.47	28.17	16.73	3.08	16.92	7.16

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

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

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

Table J.2 2011 NSDUH Respondent Pair-Level Proportions of Extreme Values and Outwinsors

Domain	*n*	Before res.pr.nr[1] (SDUWT*PRWT11*PRWT12)			After res.pr.nr[1] (SDUWT*PRWT11*...*PRWT13)		
		% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]
Total	19,976	2.95	16.42	7.07	2.76	16.65	6.85
Pair Age Group							
12-17, 12-17	3,886	0.87	5.65	1.13	0.44	3.09	0.51
12-17, 18-25	2,934	2.04	8.06	1.15	1.40	6.14	1.17
12-17, 26-34	674	1.63	4.88	0.88	1.48	4.82	0.86
12-17, 35-49	2,938	0.82	3.13	0.52	0.88	4.23	0.87
12-17, 50+	561	2.32	14.90	3.68	0.53	3.00	0.40
18-25, 18-25	4,084	5.53	16.30	2.51	5.02	18.17	3.73
18-25, 26-34	718	2.37	11.90	3.02	4.87	16.67	3.41
18-25, 35-49	995	4.02	12.69	1.92	4.92	14.62	2.65
18-25, 50+	663	2.87	11.32	3.54	1.36	4.59	0.68
26-34, 26-34	583	5.15	19.35	8.11	5.49	23.41	7.92
26-34, 35-49	305	8.52	34.25	18.39	12.13	47.10	19.64
26-34, 50+	179	3.35	22.98	15.66	3.35	17.83	8.83
35-49, 35-49	477	5.45	41.25	25.94	8.60	50.84	30.53
35-49, 50+	273	2.56	25.03	17.64	4.40	28.25	13.94
50+, 50+	706	7.22	22.20	8.09	3.97	12.33	4.85
Pair Race/Ethnicity							
Hispanic or Latino	3,052	2.92	20.98	8.88	3.70	25.09	9.77
Black or African American	2,343	4.74	21.94	6.71	2.09	12.40	4.62
White	11,729	2.00	12.97	6.24	2.12	13.12	5.60
Other	1,204	4.73	22.98	13.44	4.73	28.67	15.87
White & Black or African American	182	6.04	20.36	4.34	7.14	16.71	4.97
White & Hispanic or Latino	639	5.79	14.97	3.31	5.16	18.38	4.13
White & Other	553	5.79	12.33	1.76	3.80	8.81	1.14
Black or African American & Hispanic or Latino	90	14.44	46.55	26.83	10.00	45.89	12.92
Black or African American & Other	79	2.53	2.20	0.46	1.27	9.51	2.35
Hispanic or Latino & Other	105	3.81	24.22	6.08	5.71	39.93	11.84
Pair Gender							
Male, Male	4,381	3.65	11.48	2.48	3.54	11.86	2.80
Female, Female	4,549	2.77	9.57	1.93	2.07	9.97	1.99
Male, Female	11,046	2.75	19.90	9.93	2.73	19.90	9.37
Household Size							
Two	4,731	0.55	1.74	0.42	0.72	4.26	0.93
Three	5,492	2.64	15.66	6.17	2.86	17.18	7.03
Four or More	9,753	4.30	23.20	10.42	3.69	22.69	9.76

Table J.2 2011 NSDUH Respondent Pair-Level Proportions of Extreme Values and Outwinsors (continued)

Domain	*n*	Before res.pr.nr[1] (SDUWT*PRWT11*PRWT12)			After res.pr.nr[1] (SDUWT*PRWT11*...*PRWT13)		
		% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]
Census Region							
Northeast	3,634	2.70	10.40	2.12	2.01	9.21	2.40
South	6,493	2.54	17.98	8.85	2.22	15.89	7.87
Midwest	5,506	3.27	15.43	6.18	3.27	19.19	6.60
West	4,343	3.38	18.65	8.23	3.55	21.03	8.83
Quarter							
Quarter 1	4,662	3.26	18.77	9.81	3.26	20.26	10.21
Quarter 2	5,340	2.40	14.33	5.77	2.15	17.21	6.79
Quarter 3	5,199	2.60	15.96	6.29	2.65	14.13	4.94
Quarter 4	4,775	3.66	16.66	6.44	3.06	15.13	5.53
% Hispanic or Latino in Segment							
50-100%	1,624	3.88	24.51	10.11	3.45	25.05	8.73
10-50%	3,504	3.97	14.04	3.95	4.45	19.17	5.83
<10%	14,848	2.61	15.75	7.55	2.28	14.36	6.87
% Black or African American in Segment							
50-100%	1,668	4.98	24.28	10.08	2.82	19.97	9.41
10-50%	3,127	4.41	15.12	3.43	2.65	13.21	2.92
<10%	15,181	2.43	15.82	7.58	2.77	17.07	7.45
% Owner-Occupied DUs[1] in Segment							
50-100%	15,244	2.68	16.34	7.23	2.53	16.33	6.94
10-50%	3,728	4.53	18.21	6.96	4.37	19.96	7.07
<10%	1,004	1.20	4.12	1.43	0.20	0.40	0.11
Combined Median Rent/Housing Value							
1st Quintile	3,589	3.54	19.46	8.10	1.67	13.81	6.31
2nd Quintile	4,211	3.02	18.75	9.25	2.35	18.60	8.57
3rd Quintile	4,600	2.50	16.48	6.88	2.26	14.04	5.30
4th Quintile	4,155	3.10	11.58	2.39	3.90	15.71	3.86
5th Quintile	3,421	2.69	16.40	9.10	3.68	20.44	10.26
Population Density							
Large MSA[1]	8,111	3.24	17.86	8.14	3.37	19.38	8.18
Medium to Small MSA[1]	10,252	2.95	14.60	5.34	2.55	12.94	4.37
Non-MSA,[1] Urban	458	1.09	25.69	21.30	0.44	24.80	18.73
Non-MSA,[1] Rural	1,155	1.73	12.80	5.01	1.30	14.07	8.92
Group Quarters							
Group	288	14.58	32.90	6.76	8.68	21.48	5.08
Non-Group	19,688	2.78	16.34	7.08	2.67	16.64	6.85

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

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

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

Table J.3 2011 NSDUH Respondent Pair-Level Proportions of Extreme Values and Outwinsors

Domain	*n*	Before res.pr.ps[1] (SDUWT*PRWT11*...*PRWT13)			After res.pr.ps[1] (SDUWT*PRWT11*...*PRWT14)			Final Weight: After res.pr.ev[1] (SDUWT*PRWT11*...*PRWT15)		
		% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]
Total	19,976	2.63	11.44	2.99	1.87	7.47	1.27	0.97	4.57	0.63
Pair Age Group										
12-17, 12-17	3,885	0.44	3.05	0.54	0.39	2.61	0.38	0.23	1.50	0.27
12-17, 18-25	2,921	1.57	7.27	1.62	1.16	4.51	0.48	0.41	1.24	0.20
12-17, 26-34	685	1.31	3.82	0.83	1.90	5.87	0.73	0.58	1.84	0.39
12-17, 35-49	2,918	0.99	4.77	0.94	0.79	3.10	0.38	0.34	1.51	0.14
12-17, 50+	565	0.53	2.99	0.39	0.53	2.77	0.58	0.35	2.38	0.29
18-25, 18-25	4,015	5.01	19.00	4.66	3.29	11.71	2.29	1.97	7.29	1.21
18-25, 26-34	768	4.30	15.77	3.18	2.86	8.52	1.23	1.30	3.78	0.26
18-25, 35-49	994	4.43	13.17	2.70	2.41	6.02	1.03	0.80	3.04	0.42
18-25, 50+	666	1.50	4.89	0.81	1.50	4.50	0.33	0.45	0.72	0.09
26-34, 26-34	599	4.67	11.40	2.45	3.17	9.58	2.42	2.00	6.97	0.98
26-34, 35-49	314	10.83	21.33	6.27	7.01	8.47	1.16	2.23	2.79	0.39
26-34, 50+	192	1.04	2.68	0.21	1.56	8.89	1.30	0.52	6.36	0.58
35-49, 35-49	474	6.75	24.00	10.38	3.16	12.50	3.44	2.95	12.49	2.52
35-49, 50+	271	3.32	19.86	4.51	3.69	9.54	1.11	0.74	1.56	0.09
50+, 50+	709	4.09	12.48	2.96	4.09	10.13	1.56	2.82	7.82	0.82
Pair Race/Ethnicity										
Hispanic or Latino	3,112	3.25	13.92	3.48	1.93	5.09	0.99	1.16	3.08	0.51
Black or African American	2,262	1.99	6.43	1.48	2.39	5.53	0.71	1.24	3.34	0.34
White	11,400	2.01	8.53	1.87	1.32	5.38	0.82	0.57	2.61	0.29
Other	1,167	4.80	25.06	10.59	5.31	26.36	5.10	2.83	20.18	3.16
White & Black or African American	209	6.70	20.18	5.10	1.91	3.46	0.15	1.91	2.36	0.50
White & Hispanic or Latino	670	5.52	18.71	4.48	3.58	13.56	2.47	2.24	8.04	0.98
White & Other	728	3.85	17.86	4.10	0.82	8.03	1.30	0.14	5.05	0.59
Black or African American & Hispanic or Latino	107	6.54	22.82	4.92	9.35	38.30	7.94	7.48	34.43	6.27
Black or African American & Other	171	0.58	2.36	1.05	0.58	2.44	0.10	0.00	0.00	0.00
Hispanic or Latino & Other	150	5.33	29.40	7.41	1.33	13.93	3.72	2.00	15.36	3.42
Pair Gender										
Male, Male	4,371	3.41	11.36	3.00	2.17	9.55	1.56	0.96	4.90	0.69
Female, Female	4,538	2.09	7.75	1.41	1.19	4.20	0.56	0.37	2.23	0.16
Male, Female	11,067	2.55	12.49	3.43	2.03	7.80	1.39	1.21	5.13	0.74
Household Size										
Two	4,731	0.78	4.35	0.96	0.95	4.63	0.68	0.51	3.11	0.29
Three	5,492	2.75	15.70	5.26	1.75	9.41	1.96	1.29	7.18	1.19
Four or More	9,753	3.47	12.81	2.83	2.39	7.90	1.22	1.00	3.94	0.51

Table J.3 2011 NSDUH Respondent Pair-Level Proportions of Extreme Values and Outwinsors (continued)

Domain	*n*	Before res.pr.ps[1] (SDUWT*PRWT11*...*PRWT13) % Unweighted	Before res.pr.ps[1] (SDUWT*PRWT11*...*PRWT13) % Weighted[2]	Before res.pr.ps[1] (SDUWT*PRWT11*...*PRWT13) % Outwinsor[3]	After res.pr.ps[1] (SDUWT*PRWT11*...*PRWT14) % Unweighted	After res.pr.ps[1] (SDUWT*PRWT11*...*PRWT14) % Weighted[2]	After res.pr.ps[1] (SDUWT*PRWT11*...*PRWT14) % Outwinsor[3]	Final Weight: After res.pr.ev[1] (SDUWT*PRWT11*...*PRWT15) % Unweighted	Final Weight: After res.pr.ev[1] (SDUWT*PRWT11*...*PRWT15) % Weighted[2]	Final Weight: After res.pr.ev[1] (SDUWT*PRWT11*...*PRWT15) % Outwinsor[3]
Census Region										
Northeast	3,634	1.93	8.32	1.84	1.27	5.21	1.06	1.18	4.16	0.60
South	6,493	2.09	9.07	2.53	1.54	4.23	0.87	1.02	2.42	0.35
Midwest	5,506	3.23	14.45	3.38	2.03	9.98	1.56	0.82	6.62	0.75
West	4,343	3.27	14.54	4.15	2.67	11.57	1.76	0.90	6.18	0.94
Quarter										
Quarter 1	4,662	3.17	19.56	5.95	1.97	11.69	2.26	0.99	8.63	1.43
Quarter 2	5,340	1.97	9.21	2.41	1.54	5.98	0.89	0.75	3.56	0.43
Quarter 3	5,199	2.48	8.05	1.84	1.67	5.46	1.00	0.88	3.56	0.37
Quarter 4	4,775	3.02	9.13	1.83	2.37	6.87	0.97	1.28	2.60	0.31
% Hispanic or Latino in Segment										
50-100%	1,624	2.83	12.62	2.27	2.34	8.64	1.00	1.11	3.57	0.44
10-50%	3,504	4.28	13.73	3.08	2.51	8.55	1.44	1.34	5.77	0.55
<10%	14,848	2.22	10.45	3.08	1.67	6.90	1.26	0.86	4.32	0.69
% Black or African American in Segment										
50-100%	1,668	2.58	9.72	2.93	3.06	9.98	2.15	2.04	7.61	1.37
10-50%	3,127	2.59	10.93	2.02	2.21	8.45	1.10	1.44	4.67	0.42
<10%	15,181	2.65	11.73	3.21	1.67	6.99	1.22	0.75	4.23	0.60
% Owner-Occupied DUs[1] in Segment										
50-100%	15,244	2.40	11.53	3.12	1.65	7.39	1.32	0.87	4.64	0.65
10-50%	3,728	4.24	12.13	2.64	3.08	8.39	1.17	1.39	4.34	0.55
<10%	1,004	0.20	0.40	0.11	0.80	2.68	0.35	0.80	3.02	0.25
Combined Median Rent/Housing Value										
1st Quintile	3,589	1.48	5.68	1.00	1.95	8.79	1.40	1.11	6.61	0.69
2nd Quintile	4,211	2.11	7.32	1.60	1.61	5.72	0.80	0.78	1.93	0.25
3rd Quintile	4,600	2.00	9.76	2.50	1.54	8.84	1.62	0.89	6.08	0.78
4th Quintile	4,155	3.97	13.83	3.01	2.33	6.45	1.15	0.99	3.71	0.44
5th Quintile	3,421	3.71	18.25	6.05	1.99	7.80	1.40	1.11	4.87	0.97
Population Density										
Large MSA[1]	8,111	3.24	15.26	4.15	2.23	8.97	1.65	1.37	6.06	0.90
Medium to Small MSA[1]	10,252	2.41	7.56	1.78	1.74	6.02	0.91	0.75	3.04	0.34
Non-MSA,[1] Urban	458	0.66	0.69	0.01	0.44	1.80	0.04	0.00	0.00	0.00
Non-MSA,[1] Rural	1,155	1.13	3.41	0.79	1.13	4.24	0.46	0.43	1.56	0.12
Group Quarters										
Group	288	8.68	21.48	5.20	6.25	20.90	7.62	5.56	22.37	5.83
Non-Group	19,688	2.54	11.40	2.98	1.81	7.42	1.25	0.90	4.50	0.61

Table J.3 2011 NSDUH Respondent Pair-Level Proportions of Extreme Values and Outwinsors (continued)

Domain	*n*	Before res.pr.ps[1] (SDUWT*PRWT11*...*PRWT13)			After res.pr.ps[1] (SDUWT*PRWT11*...*PRWT14)			Final Weight: After res.pr.ev[1] (SDUWT*PRWT11*...*PRWT15)		
		% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]	% Unweighted	% Weighted[2]	% Outwinsor[3]
Pair Relationship Domain[4]										
Parent-Child (12-14)	2,049	1.27	4.86	0.92	1.32	3.62	0.34	0.49	1.17	0.10
Parent-Child (12-17)	3,842	1.02	4.13	0.83	1.07	3.61	0.50	0.44	1.90	0.22
Parent-Child (12-20)	4,582	1.42	5.55	1.09	1.48	5.11	0.72	0.57	2.42	0.30
Sibling (12-14)-Sibling (15-17)	2,350	0.30	2.75	0.43	0.34	1.84	0.09	0.09	0.43	0.07
Sibling (12-17)-Sibling (18-25)	2,618	1.60	7.32	1.64	0.84	3.54	0.37	0.19	0.62	0.13
Spouse-Spouse/Partner-Partner	3,693	2.14	12.90	4.25	3.66	11.38	2.35	2.63	8.84	1.36
Spouse-Spouse/Partner-Partner with Children (Younger Than 18)	1,730	2.08	16.86	7.58	5.72	16.74	4.27	4.62	14.52	2.50

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

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

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

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

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

Table K.1 2011 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	19,976	218,003,569	218,003,569	218,003,569	0.00	-0.00
Pair Age Group						
12-17, 12-17	3,885	7,336,152	7,337,936	7,337,936	-0.02	-0.00
12-17, 18-25	2,921	8,284,373	8,204,268	8,204,268	0.98	-0.00
12-17, 26-34	685	5,291,417	5,265,885	5,265,885	0.48	-0.00
12-17, 35-49	2,918	29,673,332	29,825,396	29,825,396	-0.51	0.00
12-17, 50+	565	11,335,386	11,322,054	11,322,054	0.12	-0.00
18-25, 18-25	4,015	12,520,114	12,605,261	12,605,261	-0.68	-0.00
18-25, 26-34	768	7,117,229	7,175,704	7,175,704	-0.81	0.00
18-25, 35-49	994	17,416,814	17,516,184	17,516,184	-0.57	0.00
18-25, 50+	666	18,019,261	18,056,464	18,056,464	-0.21	0.00
26-34, 26-34	599	10,510,424	10,257,001	10,257,001	2.47	0.00
26-34, 35-49	314	8,200,464	8,446,195	8,446,195	-2.91	0.00
26-34, 50+	192	10,989,825	10,395,093	10,395,093	5.72	0.00
35-49, 35-49	474	18,961,313	18,731,573	18,731,573	1.23	0.00
35-49, 50+	271	15,540,919	16,138,008	16,138,008	-3.70	0.00
50+, 50+	709	36,806,547	36,726,547	36,726,547	0.22	0.00
Pair Race/Ethnicity						
Hispanic or Latino	3,112	38,141,357	37,310,806	37,310,806	2.23	-0.00
Black or African American	2,262	24,230,009	24,222,858	24,222,858	0.03	0.00
White	11,400	118,424,127	122,063,516	122,063,516	-2.98	-0.00
Other	1,167	15,554,955	15,939,034	15,939,034	-2.41	0.00
White & Black or African American	209	2,201,641	1,863,495	1,863,495	18.15	0.00
White & Hispanic or Latino	670	8,306,504	8,039,359	8,039,359	3.32	-0.00
White & Other	728	7,176,106	5,498,325	5,498,325	30.51	0.00
Black or African American & Hispanic or Latino	107	1,228,193	1,382,646	1,382,646	-11.17	0.00
Black or African American & Other	171	1,102,535	714,736	714,736	54.26	0.00
Hispanic or Latino & Other	150	1,638,142	968,794	968,794	69.09	-0.00
Pair Gender						
Male, Male	4,371	39,596,145	39,723,670	39,723,670	-0.32	-0.00
Female, Female	4,538	39,080,666	39,113,567	39,113,567	-0.08	-0.00
Male, Female	11,067	139,326,759	139,166,333	139,166,333	0.12	0.00
Pair Relationship Domain[3,4,5]						
Parent-Child (12-14)*	2,049	11,459,749	12,496,945	12,496,945	-8.30	0.00
Parent-Child (12-17)*	3,842	23,686,808	25,085,551	25,085,551	-5.58	-0.00
Parent-Child (15-17)*	1,793	12,227,059	12,588,606	12,588,606	-2.87	-0.00
Parent-Child (12-20)*	4,582	32,348,864	34,477,404	34,477,404	-6.17	-0.00
Parent*-Child (12-14)	2,049	18,082,767	19,248,556	19,248,556	-6.06	0.00
Parent*-Child (12-17)	3,842	30,581,098	32,127,985	32,127,985	-4.81	-0.00
Parent*-Child (15-17)	1,793	19,189,086	19,565,436	19,251,674	-0.33	1.63
Parent*-Child (12-20)	4,582	38,435,099	40,040,580	40,040,580	-4.01	-0.00
Sibling (12-14)-Sibling (15-17)*	2,350	3,909,157	4,127,666	4,127,666	-5.29	-0.00
Sibling (12-17)-Sibling (18-25)*	2,618	6,156,561	6,060,569	6,060,569	1.58	0.00
Spouse-Spouse/Partner-Partner	3,693	65,566,071	70,606,641	70,606,641	-7.14	0.00
Spouse-Spouse/Partner-Partner with Children (Younger Than 18)	1,730	22,157,509	28,976,302	28,976,302	-23.53	0.00

Table K.1 2011 NSDUH Respondent Pair-Level Slippage Rates (continued)

Domain	*n*	Initial Total (*I*)[1]	Final Total (*F*)[2]	Control Total from SDU (*C*)	(*I* - *C*)/*C*%	(*F* - *C*)/*C*%
Household Size						
Two	4,731	54,610,438	54,610,438	54,610,438	-0.00	-0.00
Three	5,492	56,064,021	56,064,021	56,064,021	-0.00	-0.00
Four or More	9,753	107,329,111	107,329,111	107,329,111	0.00	0.00
Census Region						
Northeast	3,634	40,154,533	40,154,533	40,154,533	-0.00	0.00
South	6,493	77,243,179	77,243,179	77,243,179	0.00	-0.00
Midwest	5,506	44,741,677	44,741,677	44,741,677	0.00	-0.00
West	4,343	55,864,180	55,864,180	55,864,180	-0.00	-0.00
Quarter						
Quarter 1	4,662	53,598,809	53,598,809	53,598,809	0.00	-0.00
Quarter 2	5,340	54,185,729	54,185,729	54,185,729	0.00	0.00
Quarter 3	5,199	56,011,516	56,011,516	56,011,516	0.00	-0.00
Quarter 4	4,775	54,207,516	54,207,516	54,207,516	-0.00	0.00
% Hispanic or Latino in Segment						
50-100%	1,624	24,478,758	24,478,758	24,478,758	0.00	0.00
10-50%	3,504	49,527,769	49,527,769	49,527,769	0.00	0.00
<10%	14,848	143,997,041	143,997,041	143,997,041	0.00	-0.00
% Black or African American in Segment						
50-100%	1,668	17,152,725	17,152,725	17,152,725	0.00	0.00
10-50%	3,127	36,370,312	36,370,312	36,370,312	0.00	0.00
<10%	15,181	164,480,533	164,480,533	164,480,533	-0.00	-0.00
% Owner-Occupied DUs in Segment						
50-100%	15,244	178,084,113	178,084,113	178,084,113	0.00	-0.00
10-50%	3,728	36,157,686	36,157,686	36,157,686	-0.00	0.00
<10%	1,004	3,761,770	3,761,770	3,761,770	0.00	0.00
Combined Median Rent/Housing Value						
1st Quintile	3,589	32,921,442	32,921,442	32,921,442	-0.00	0.00
2nd Quintile	4,211	41,991,775	41,991,775	41,991,775	-0.00	0.00
3rd Quintile	4,600	46,809,185	46,809,185	46,809,185	0.00	-0.00
4th Quintile	4,155	48,676,684	48,676,684	48,676,684	0.00	-0.00
5th Quintile	3,421	47,604,483	47,604,483	47,604,483	0.00	0.00
Population Density						
Large MSA	8,111	117,802,921	117,802,921	117,802,921	0.00	-0.00
Medium to Small MSA	10,252	87,608,568	87,608,568	87,608,568	-0.00	-0.00
Non-MSA, Urban	458	3,424,435	3,424,435	3,424,435	0.00	0.00
Non-MSA, Rural	1,155	9,167,645	9,167,645	9,167,645	0.00	0.00
Group Quarters						
Group	288	838,742	838,742	838,742	-0.00	-0.00
Non-Group	19,688	217,164,828	217,164,828	217,164,828	0.00	0.00

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

[1] WT1*...*WT10*PRWT11*...*PRWT13 (before person pair poststratification).

[2] WT1*...*WT10*PRWT11*...*PRWT14 (after person pair poststratification).

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

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

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

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

Table L.1 2011 NSDUH Selected Pair-Level Weight Summary Statistics

Domain	*n*	SDU-Level Weights[1] (SDUWT: WT1*...*WT10)						Before sel.pr.ps[1] (SDUWT*PRWT11)						After sel.pr.ps[1] (SDUWT*PRWT11*PRWT12)					
		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	27,095	11	417	664	1,070	10,206	1.64	14	1,068	2,745	7,380	6,644,771	43.78	7	975	2,682	7,414	1,047,876	8.83
Pair Age Group																			
12-17, 12-17	4,649	12	378	622	1,009	8,360	1.64	14	586	992	1,834	59,548	2.97	7	511	960	1,948	23,706	2.53
12-17, 18-25	3,756	22	437	687	1,128	8,513	1.62	40	730	1,284	2,326	83,496	3.33	10	711	1,363	2,764	23,984	2.27
12-17, 26-34	834	28	365	625	1,028	10,165	1.80	170	2,064	3,771	6,543	87,625	2.65	49	1,821	3,904	7,346	85,203	3.03
12-17, 35-49	3,855	15	422	640	1,006	8,568	1.57	159	2,942	5,057	8,915	394,893	3.81	127	2,281	4,350	9,311	105,719	2.59
12-17, 50+	766	29	436	652	1,041	5,708	1.58	488	5,317	8,640	14,785	527,868	4.44	413	4,883	8,725	16,426	242,009	2.92
18-25, 18-25	5,476	17	416	710	1,165	10,072	1.69	68	730	1,392	2,485	47,804	2.84	35	579	1,252	2,962	25,254	2.50
18-25, 26-34	1,049	32	430	698	1,156	7,361	1.70	289	2,419	4,106	7,483	252,561	4.17	123	1,584	3,361	7,397	144,541	3.86
18-25, 35-49	1,469	11	433	707	1,153	6,432	1.58	209	3,079	5,995	12,085	273,086	4.04	159	2,738	6,261	13,972	134,386	2.73
18-25, 50+	1,057	32	465	715	1,156	10,206	1.70	532	5,857	9,236	17,042	292,982	3.02	373	5,409	10,768	20,328	226,411	2.50
26-34, 26-34	858	23	405	654	1,066	4,546	1.54	243	4,517	7,393	12,498	247,130	2.79	131	3,883	6,800	13,466	292,033	3.82
26-34, 35-49	492	22	420	630	959	5,006	1.68	269	4,992	8,347	14,220	1,927,023	26.16	335	3,595	8,126	16,883	469,873	6.22
26-34, 50+	302	66	401	672	1,074	3,468	1.52	1,437	9,350	16,795	31,581	798,883	6.28	769	9,231	20,944	42,992	553,790	3.40
35-49, 35-49	748	32	444	652	1,012	5,834	1.56	569	5,867	8,921	14,956	2,681,904	27.09	810	4,297	9,578	21,188	1,047,876	9.34
35-49, 50+	477	54	370	609	1,014	3,735	1.53	1,056	8,013	14,135	27,118	6,644,771	71.21	708	7,841	16,031	41,283	1,046,187	4.44
50+, 50+	1,307	37	408	633	1,028	4,090	1.45	1,145	10,741	16,673	30,478	1,262,757	4.50	753	11,039	18,465	34,161	946,747	3.14
Pair Race/Ethnicity																			
Hispanic or Latino	4,001	11	509	815	1,408	8,568	1.58	14	1,306	3,151	8,358	1,927,023	19.13	7	1,122	3,075	8,423	564,887	7.73
Black or African American	2,912	35	487	728	1,120	10,165	1.65	46	1,117	2,850	7,251	772,062	11.47	10	1,042	2,750	7,186	755,091	9.15
White	16,243	29	412	634	963	7,889	1.50	74	1,039	2,661	7,144	1,275,192	12.32	45	967	2,594	7,020	1,046,187	9.00
Other	1,770	18	195	539	1,098	6,272	2.11	22	705	2,070	6,865	2,681,904	50.59	33	803	2,375	7,298	1,047,876	12.68
White & Black or African American	236	32	494	766	1,123	6,432	1.60	53	1,198	2,760	7,739	110,319	3.95	29	888	2,571	6,626	175,469	5.16
White & Hispanic or Latino	858	32	452	743	1,275	6,058	1.64	73	1,390	3,607	9,126	115,157	3.55	35	1,109	3,599	10,535	194,654	4.36
White & Other	732	22	218	514	885	5,119	1.93	49	1,033	2,797	7,141	122,719	4.16	38	829	2,429	7,654	182,895	4.71
Black or African American & Hispanic or Latino	111	53	654	1,002	1,855	10,206	2.05	105	1,633	4,006	13,891	6,644,771	80.59	40	1,434	4,484	12,531	329,438	8.02
Black or African American & Other	97	36	289	491	842	2,550	1.67	68	1,377	2,839	5,748	48,196	2.80	33	1,244	3,081	7,164	117,177	4.61
Hispanic or Latino & Other	135	17	223	513	1,040	6,172	2.42	62	577	1,799	6,407	253,518	11.31	49	311	1,041	4,570	310,091	17.12

Table L.1 2011 NSDUH Selected Pair-Level Weight Summary Statistics (continued)

Domain	*n*	SDU-Level Weights[1] (SDUWT: WT1*...*WT10)						Before sel.pr.ps[1] (SDUWT*PRWT11)						After sel.pr.ps[1] (SDUWT*PRWT11*PRWT12)					
		Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
Pair Gender																			
Male, Male	6,020	31	431	679	1,106	8,450	1.64	46	956	2,282	6,145	394,893	5.29	10	887	2,304	5,950	284,226	5.62
Female, Female	5,895	11	414	668	1,049	10,206	1.62	29	1,051	2,445	6,142	848,933	9.43	25	936	2,419	6,337	718,711	7.12
Male, Female	15,180	12	413	655	1,065	10,165	1.64	14	1,143	3,099	8,473	6,644,771	54.49	7	1,032	3,024	8,700	1,047,876	9.54
Household Size																			
Two	6,903	17	397	645	999	7,361	1.59	67	1,344	4,055	10,439	95,523	2.63	29	832	2,837	9,520	167,728	3.52
Three	7,438	11	418	655	1,042	8,360	1.63	22	1,146	2,935	6,533	6,644,771	135.89	15	1,138	2,904	7,116	1,047,876	8.84
Four or More	12,754	12	426	683	1,122	10,206	1.66	14	947	2,213	6,563	1,927,023	17.42	7	982	2,498	6,831	1,046,187	11.33
Census Region																			
Northeast	5,254	28	335	653	901	8,047	1.56	46	1,050	2,670	7,213	772,062	6.95	10	921	2,583	7,176	755,091	6.67
South	8,432	11	500	847	1,246	6,937	1.51	14	1,359	3,287	8,755	6,644,771	71.58	7	1,127	3,129	9,025	1,046,187	8.41
Midwest	7,525	21	449	567	737	6,263	1.38	29	881	2,446	6,180	1,275,192	21.89	26	913	2,404	5,717	946,747	10.01
West	5,884	18	243	674	1,467	10,206	1.87	22	993	2,602	8,250	2,681,904	28.64	36	936	2,696	8,254	1,047,876	9.12
Quarter																			
Quarter1	6,282	18	437	726	1,171	8,450	1.63	26	1,121	2,833	7,897	2,681,904	27.63	15	988	2,799	7,716	1,047,876	12.18
Quarter2	7,314	11	408	629	1,001	8,047	1.57	14	1,051	2,667	7,293	741,770	6.84	7	952	2,552	6,931	553,790	6.72
Quarter3	7,038	19	399	620	1,055	7,653	1.64	22	1,013	2,567	7,108	6,644,771	105.36	23	938	2,628	7,340	654,096	8.33
Quarter4	6,461	21	456	674	1,077	10,206	1.68	29	1,085	2,913	7,355	1,927,023	17.37	29	1,029	2,809	7,580	718,711	7.73
% Hispanic or Latino in Segment																			
50-100%	2,128	22	617	1,033	1,508	8,568	1.43	40	1,768	4,303	10,408	599,658	8.00	36	1,464	3,894	10,884	564,887	7.13
10-50%	4,792	11	570	862	1,428	10,206	1.56	14	1,467	3,562	9,512	1,927,023	13.78	7	1,419	3,763	10,207	469,873	5.29
<10%	20,175	18	354	605	930	7,653	1.62	26	958	2,450	6,717	6,644,771	62.94	10	881	2,389	6,454	1,047,876	10.49
% Black or African American in Segment																			
50-100%	2,101	11	446	708	1,064	10,206	1.82	14	1,122	2,880	7,439	1,927,023	33.94	7	986	2,766	7,340	553,790	8.21
10-50%	4,075	34	513	764	1,242	8,513	1.58	52	1,294	3,053	7,832	6,644,771	134.28	40	1,195	3,130	8,189	564,887	5.70
<10%	20,919	18	389	642	1,040	8,568	1.62	22	1,030	2,661	7,286	2,681,904	17.55	10	943	2,586	7,219	1,047,876	9.66
% Owner-Occupied DUs[1] in Segment																			
50-100%	20,773	18	416	649	1,050	10,206	1.62	26	1,087	2,881	7,637	6,644,771	50.19	19	1,071	2,939	8,086	1,047,876	8.77
10-50%	4,996	11	415	710	1,114	8,568	1.66	14	1,008	2,385	6,609	741,770	9.12	7	927	2,401	6,502	553,790	7.78
<10%	1,326	19	441	726	1,273	7,889	1.66	22	1,077	2,244	5,971	364,198	6.36	10	405	908	2,279	88,635	6.14
Combined Median Rent/Housing Value																			
1st Quintile	4,650	18	331	603	962	8,568	1.70	26	985	2,436	6,527	741,770	9.34	23	871	2,337	6,488	654,096	8.90
2nd Quintile	5,587	27	389	635	989	10,206	1.72	36	979	2,560	7,108	6,644,771	125.56	22	909	2,521	6,853	564,887	9.08
3rd Quintile	6,126	11	364	624	1,006	8,450	1.69	14	989	2,535	6,936	1,262,757	14.04	7	900	2,495	7,080	946,747	8.66
4th Quintile	5,804	19	447	703	1,094	6,432	1.60	22	1,135	2,868	7,605	285,905	4.95	10	1,069	2,940	8,043	455,857	5.19
5th Quintile	4,928	32	526	797	1,246	7,653	1.48	53	1,338	3,272	9,171	2,681,904	27.79	36	1,187	3,231	8,979	1,047,876	11.60

Table L.1 2011 NSDUH Selected Pair-Level Weight Summary Statistics (continued)

Domain	*n*	SDU-Level Weights[1] (SDUWT: WT1*...*WT10)						Before sel.pr.ps[1] (SDUWT*PRWT11)						After sel.pr.ps[1] (SDUWT*PRWT11*PRWT12)					
		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]	11,353	11	588	838	1,321	10,206	1.48	14	1,480	3,700	9,732	2,681,904	18.91	7	1,409	3,637	10,158	1,047,876	8.16
Medium to Small MSA[1]	13,593	19	302	564	855	7,361	1.69	22	882	2,214	6,010	6,644,771	89.73	10	827	2,204	5,910	755,091	7.76
Non-MSA,[1] Urban	592	18	242	539	835	3,992	1.69	26	763	1,910	5,935	311,787	10.35	40	646	1,538	4,862	654,096	24.52
Non-MSA,[1] Rural	1,557	29	202	447	801	6,263	1.89	46	696	1,952	5,772	1,094,909	24.53	33	590	1,685	4,942	564,845	12.02
Group Quarters																			
Group	349	41	396	677	930	4,274	1.51	67	851	1,216	2,130	26,319	3.01	29	500	1,336	2,462	36,869	3.20
Non-Group	26,746	11	417	663	1,071	10,206	1.64	14	1,077	2,789	7,457	6,644,771	43.56	7	984	2,719	7,492	1,047,876	8.78

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

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

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

Table L.2 2011 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.nr)

Domain	*n*	Before res.pr.nr[1] (SDUWT*PRWT11*PRWT12) Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	After res.pr.nr[1] (SDUWT*PRWT11*...*PRWT13) Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
Total	19,976	7	915	2,433	6,516	1,047,876	9.64	7	1,074	2,990	8,707	2,063,702	11.26
Pair Age Group													
12-17, 12-17	3,886	7	515	969	1,979	23,706	2.54	7	581	1,145	2,329	31,948	2.50
12-17, 18-25	2,934	10	695	1,374	2,805	23,984	2.28	17	851	1,691	3,581	35,668	2.31
12-17, 26-34	674	49	1,748	3,817	6,891	85,203	3.27	49	1,957	4,321	8,488	106,722	3.23
12-17, 35-49	2,938	127	2,259	4,361	9,393	105,719	2.54	132	2,624	5,421	12,031	123,243	2.75
12-17, 50+	561	413	4,759	8,426	15,808	242,009	3.06	413	5,924	11,356	24,487	197,145	2.58
18-25, 18-25	4,084	35	568	1,241	2,996	25,254	2.53	35	646	1,474	3,916	48,483	2.76
18-25, 26-34	718	150	1,630	3,220	6,954	144,541	3.74	173	1,892	4,049	10,761	186,604	4.07
18-25, 35-49	995	159	2,663	6,254	13,972	134,386	2.78	175	3,367	8,719	21,394	190,504	2.84
18-25, 50+	663	373	5,057	10,578	20,469	226,411	2.56	373	7,319	15,893	35,192	198,867	2.31
26-34, 26-34	583	131	3,855	6,851	13,098	292,033	3.79	254	4,347	8,169	17,611	557,722	5.19
26-34, 35-49	305	335	3,416	8,152	16,893	469,873	6.64	499	4,738	10,876	22,330	783,738	6.79
26-34, 50+	179	769	9,143	21,495	45,517	553,790	3.85	769	11,204	31,508	72,125	647,476	3.15
35-49, 35-49	477	810	4,236	9,119	20,989	1,047,876	9.61	870	5,726	13,718	29,173	2,063,702	11.12
35-49, 50+	273	708	7,543	15,598	35,721	1,046,187	6.00	711	9,699	23,131	67,165	1,471,845	4.90
50+, 50+	706	753	10,072	17,471	31,656	946,747	3.52	1,133	17,848	34,278	65,380	1,395,732	2.85
Pair Race/Ethnicity													
Hispanic or Latino	3,052	7	1,052	2,875	7,684	469,873	7.80	7	1,203	3,440	9,700	783,738	9.25
Black or African American	2,343	10	973	2,495	6,378	553,790	7.51	15	1,090	3,024	7,903	647,476	8.28
White	11,729	45	904	2,357	6,232	1,046,187	10.08	46	1,071	2,894	8,335	1,471,845	10.90
Other	1,204	33	682	1,865	5,877	1,047,876	22.89	33	868	2,513	9,763	2,063,702	25.50
White & Black or African American	182	29	873	2,367	6,273	175,469	5.90	36	942	2,758	8,510	174,136	5.06
White & Hispanic or Latino	639	35	1,088	3,330	9,877	126,829	3.43	35	1,194	4,038	12,553	326,585	4.85
White & Other	553	38	799	2,252	5,982	103,654	4.54	38	836	2,643	7,794	279,333	6.98
Black or African American & Hispanic or Latino	90	40	1,253	3,877	10,906	329,438	10.46	41	1,520	4,206	11,661	352,451	8.77
Black or African American & Other	79	33	1,455	3,081	7,759	54,486	2.79	34	1,524	3,848	8,783	109,030	4.22
Hispanic or Latino & Other	105	49	284	966	3,632	79,751	8.15	49	424	1,864	7,874	215,827	8.79

Table L.2 2011 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.nr) (continued)

Domain	*n*	Before res.pr.nr[1] (SDUWT*PRWT11*PRWT12)						After res.pr.nr[1] (SDUWT*PRWT11*...*PRWT13)					
		Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
Pair Gender													
Male, Male	4,381	10	845	2,179	5,494	226,411	5.13	15	1,009	2,665	7,265	442,186	6.82
Female, Female	4,549	25	899	2,311	5,888	268,624	5.22	25	1,042	2,776	7,361	341,977	6.22
Male, Female	11,046	7	953	2,628	7,381	1,047,876	11.20	7	1,116	3,263	10,005	2,063,702	12.68
Household Size													
Two	4,731	29	722	2,186	7,405	103,654	3.76	35	777	2,404	9,886	326,585	5.70
Three	5,492	15	1,017	2,657	6,394	1,047,876	9.64	15	1,187	3,235	8,621	2,063,702	13.54
Four or More	9,753	7	955	2,382	6,287	1,046,187	11.87	7	1,173	3,028	8,368	1,471,845	13.07
Census Region													
Northeast	3,634	10	815	2,184	5,991	242,009	4.97	17	969	2,772	8,715	487,944	6.77
South	6,493	7	1,067	2,843	7,786	1,046,187	9.72	7	1,205	3,393	9,799	1,471,845	10.39
Midwest	5,506	26	862	2,189	5,135	946,747	10.66	27	1,053	2,676	6,976	1,395,732	11.78
West	4,343	36	874	2,491	7,517	1,047,876	9.95	36	1,051	3,085	9,905	2,063,702	13.85
Quarter													
Quarter1	4,662	15	923	2,489	6,771	1,047,876	15.31	15	1,080	3,035	9,051	2,063,702	18.74
Quarter2	5,340	7	905	2,344	6,199	553,790	7.34	7	1,034	2,771	8,190	956,398	9.65
Quarter3	5,199	26	901	2,455	6,522	654,096	8.20	27	1,079	3,074	8,797	847,428	7.81
Quarter4	4,775	29	936	2,477	6,589	469,873	7.06	33	1,107	3,083	8,948	783,738	8.49
% Hispanic or Latino in Segment													
50-100%	1,624	52	1,391	3,572	10,341	409,960	6.84	52	1,490	4,234	12,781	565,978	8.01
10-50%	3,504	7	1,334	3,517	8,986	469,873	5.08	7	1,540	4,373	12,010	783,738	6.84
<10%	14,848	10	825	2,146	5,718	1,047,876	11.94	15	973	2,613	7,577	2,063,702	13.64
% Black or African American in Segment													
50-100%	1,668	7	929	2,417	6,529	553,790	9.98	7	1,022	2,902	7,743	783,738	11.04
10-50%	3,127	40	1,113	2,930	7,379	329,438	5.00	41	1,227	3,461	9,391	565,978	6.86
<10%	15,181	10	885	2,349	6,339	1,047,876	10.81	17	1,048	2,908	8,660	2,063,702	12.31
% Owner-Occupied DUs[1] in Segment													
50-100%	15,244	19	992	2,659	7,106	1,047,876	9.55	27	1,186	3,295	9,478	2,063,702	11.16
10-50%	3,728	7	893	2,230	5,772	553,790	8.62	7	1,014	2,679	7,829	647,476	9.75
<10%	1,004	10	389	833	2,120	75,569	6.31	15	407	924	2,481	151,143	9.95
Combined Median Rent/Housing Value													
1st Quintile	3,589	33	827	2,175	5,944	654,096	10.45	33	930	2,455	7,039	847,428	11.30
2nd Quintile	4,211	23	883	2,398	6,200	541,824	9.49	23	1,016	2,800	7,982	956,398	11.76
3rd Quintile	4,600	7	843	2,302	6,387	946,747	9.74	7	987	2,738	8,494	1,395,732	10.24
4th Quintile	4,155	10	1,007	2,653	6,825	268,624	4.66	17	1,260	3,492	9,416	487,944	6.34
5th Quintile	3,421	36	1,035	2,802	7,507	1,047,876	13.53	36	1,298	3,595	10,943	2,063,702	15.03

Table L.2 2011 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.nr) (continued)

Domain	*n*	Before res.pr.nr[1] (SDUWT*PRWT11*PRWT12)						After res.pr.nr[1] (SDUWT*PRWT11*...*PRWT13)					
		Min	Q1[2]	Med	Q3[2]	Max	UWE[3]	Min	Q1[2]	Med	Q3[2]	Max	UWE[3]
Population Density													
Large MSA[1]	8,111	7	1,286	3,284	8,741	1,047,876	9.44	7	1,538	4,109	11,803	2,063,702	10.77
Medium to Small MSA[1]	10,252	10	793	2,052	5,349	553,790	7.68	17	930	2,451	7,005	647,476	8.42
Non-MSA,[1] Urban	458	40	643	1,492	4,602	654,096	31.63	40	745	1,675	5,301	847,428	30.88
Non-MSA,[1] Rural	1,155	33	562	1,591	4,594	334,524	7.89	33	608	1,854	5,757	956,398	17.34
Group Quarters													
Group	288	29	467	1,232	2,114	36,869	3.47	36	563	1,263	2,832	47,921	3.82
Non-Group	19,688	7	926	2,467	6,582	1,047,876	9.58	7	1,086	3,047	8,807	2,063,702	11.18

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

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

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

Table L.3 2011 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.ps and res.pr.ev)

Domain	*n*	Before res.pr.ps[1] (SDUWT*PRWT11*...*PRWT13)						After res.pr.ps[1] (SDUWT*PRWT11*...*PRWT14)						Final Weight: After res.pr.ev[1] (SDUWT*PRWT11*...*PRWT15)					
		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	19,976	7	1,074	2,990	8,707	2,063,702	11.26	8	1,007	2,886	8,532	1,232,038	10.10	8	1,003	2,885	8,547	1,203,872	10.03
Pair Age Group																			
12-17, 12-17	3,885	7	581	1,143	2,323	31,948	2.50	8	527	1,116	2,321	29,428	2.63	8	525	1,110	2,324	27,045	2.62
12-17, 18-25	2,921	17	858	1,723	3,588	35,735	2.36	17	777	1,673	3,632	28,659	2.37	18	771	1,667	3,631	25,152	2.35
12-17, 26-34	685	49	1,912	4,184	8,285	106,722	3.34	27	1,688	3,920	8,511	123,384	3.63	26	1,654	3,995	8,517	123,301	3.64
12-17, 35-49	2,918	132	2,627	5,411	12,047	123,243	2.74	75	2,501	5,376	12,070	131,479	2.78	71	2,479	5,343	12,068	132,246	2.78
12-17, 50+	565	413	5,924	11,078	23,736	197,145	2.61	259	5,009	9,542	23,391	175,429	2.84	245	4,855	9,537	23,495	177,360	2.85
18-25, 18-25	4,015	35	644	1,471	3,921	53,589	2.87	23	552	1,436	4,129	48,874	2.84	23	553	1,449	4,166	33,823	2.74
18-25, 26-34	768	124	1,708	3,640	9,603	184,008	4.11	134	1,576	3,910	10,243	201,276	4.03	132	1,585	3,943	10,132	209,828	3.99
18-25, 35-49	994	175	3,411	8,683	21,389	190,504	2.85	119	2,873	8,387	21,334	164,337	2.95	112	2,882	8,537	21,772	154,663	2.92
18-25, 50+	666	373	7,319	15,750	35,060	198,867	2.32	258	5,872	14,584	35,646	188,037	2.45	249	5,854	14,618	35,643	174,420	2.44
26-34, 26-34	599	254	4,085	7,878	17,458	557,722	5.29	141	2,834	7,165	15,737	716,475	7.02	138	2,778	7,291	15,456	743,129	7.23
26-34, 35-49	314	499	4,625	10,421	21,644	783,738	7.00	348	5,321	12,303	26,227	763,994	5.88	322	5,395	11,668	24,933	761,308	5.89
26-34, 50+	192	769	10,858	29,126	70,139	647,476	3.18	476	7,072	22,797	63,951	677,133	3.33	450	6,979	23,223	65,061	661,602	3.30
35-49, 35-49	474	870	5,627	13,639	32,351	2,063,702	10.89	536	5,525	12,295	28,450	1,232,038	9.04	499	5,629	11,804	28,986	1,203,872	9.06
35-49, 50+	271	711	9,440	22,321	65,673	1,471,845	5.12	733	8,729	22,191	64,953	954,969	3.94	703	8,451	22,625	65,452	966,269	3.96
50+, 50+	709	1,133	17,848	34,252	64,801	1,395,732	2.85	1,224	16,100	32,689	68,230	1,116,206	2.56	1,163	16,628	32,899	68,117	1,022,170	2.47
Pair Race/Ethnicity																			
Hispanic or Latino	3,112	7	1,212	3,397	9,737	783,738	9.35	8	1,110	3,206	9,640	763,994	9.11	8	1,101	3,220	9,748	761,308	9.14
Black or African American	2,262	15	1,090	3,076	8,032	647,476	8.42	19	1,040	2,969	7,710	453,464	8.62	18	1,044	2,972	7,791	450,516	8.67
White	11,400	41	1,072	2,877	8,329	1,471,845	11.00	26	1,053	2,928	8,412	1,141,198	10.35	25	1,051	2,929	8,457	1,135,896	10.29
Other	1,167	33	894	2,564	9,852	2,063,702	25.83	26	714	2,419	9,892	1,232,038	14.15	26	693	2,405	10,052	1,203,872	13.74
White & Black or African American	209	36	942	2,666	8,260	174,136	5.24	17	625	1,975	6,230	208,133	6.09	18	616	1,959	6,005	220,723	6.40
White & Hispanic or Latino	670	35	1,139	4,022	12,082	326,585	4.87	19	1,028	3,677	11,134	223,383	4.87	18	1,048	3,747	11,059	232,563	4.81
White & Other	728	38	855	2,714	8,178	279,333	6.90	14	535	1,832	6,152	303,947	8.40	14	528	1,821	6,060	280,856	8.27
Black or African American & Hispanic or Latino	107	41	1,280	3,881	8,910	172,322	5.30	30	932	2,340	7,726	235,278	7.69	29	967	2,363	7,547	202,653	7.29
Black or African American & Other	171	51	1,026	3,380	6,649	109,030	4.37	17	536	1,796	4,762	70,258	4.08	17	532	1,806	4,907	70,889	4.10
Hispanic or Latino & Other	150	49	428	2,515	8,783	215,827	6.72	27	270	1,379	4,720	131,293	6.54	26	276	1,389	4,698	128,504	6.44

Table L.3 2011 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.ps and res.pr.ev) (continued)

Domain	*n*	Before res.pr.ps[1] (SDUWT*PRWT11*...*PRWT13)						After res.pr.ps[1] (SDUWT*PRWT11*...*PRWT14)						Final Weight: After res.pr.ev[1] (SDUWT*PRWT11*...*PRWT15)					
		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	4,371	15	1,008	2,663	7,242	442,186	6.84	17	940	2,540	7,328	677,133	8.06	18	936	2,566	7,358	661,602	8.01
Female, Female	4,538	25	1,041	2,774	7,379	341,977	6.22	26	954	2,641	7,464	388,280	6.62	26	940	2,639	7,466	386,992	6.61
Male, Female	11,067	7	1,121	3,265	9,989	2,063,702	12.68	8	1,062	3,120	9,656	1,232,038	10.79	8	1,061	3,116	9,706	1,203,872	10.72
Household Size																			
Two	4,731	35	777	2,404	9,886	326,585	5.70	17	651	2,198	9,055	303,947	5.95	18	639	2,197	9,025	277,414	5.94
Three	5,492	15	1,187	3,235	8,621	2,063,702	13.54	14	1,097	3,196	8,513	1,232,038	9.62	14	1,093	3,190	8,546	1,203,872	9.51
Four or More	9,753	7	1,173	3,028	8,368	1,471,845	13.07	8	1,133	2,983	8,277	1,141,198	12.50	8	1,133	2,988	8,280	1,135,896	12.43
Census Region																			
Northeast	3,634	17	969	2,772	8,715	487,944	6.77	14	829	2,506	8,587	453,464	6.96	14	829	2,524	8,654	450,516	6.91
South	6,493	7	1,205	3,393	9,799	1,471,845	10.39	8	1,081	3,060	9,414	1,141,198	9.88	8	1,080	3,064	9,454	1,135,896	9.92
Midwest	5,506	27	1,053	2,676	6,976	1,395,732	11.78	26	1,058	2,777	6,881	1,116,206	10.53	25	1,051	2,764	6,868	1,022,170	10.20
West	4,343	36	1,051	3,085	9,905	2,063,702	13.85	18	988	3,022	9,842	1,232,038	11.17	16	977	3,017	9,852	1,203,872	11.12
Quarter																			
Quarter1	4,662	15	1,080	3,035	9,051	2,063,702	18.74	19	1,052	3,055	9,123	1,232,038	11.92	19	1,059	3,083	9,173	1,203,872	11.50
Quarter2	5,340	7	1,034	2,771	8,190	956,398	9.65	8	966	2,673	7,928	860,558	9.56	8	959	2,666	7,913	860,845	9.59
Quarter3	5,199	27	1,079	3,074	8,797	847,428	7.81	14	983	2,923	8,357	1,141,198	9.48	14	977	2,907	8,341	1,135,896	9.53
Quarter4	4,775	33	1,107	3,083	8,948	783,738	8.49	17	1,027	2,900	8,656	763,994	9.27	18	1,028	2,910	8,652	761,308	9.36
% Hispanic or Latino in Segment																			
50-100%	1,624	52	1,490	4,234	12,781	565,978	8.01	24	1,295	3,881	12,807	644,306	8.16	23	1,276	3,887	12,765	651,163	8.15
10-50%	3,504	7	1,540	4,373	12,010	783,738	6.84	8	1,461	4,175	11,983	763,994	7.37	8	1,472	4,159	11,927	761,308	7.38
<10%	14,848	15	973	2,613	7,577	2,063,702	13.64	14	922	2,538	7,496	1,232,038	11.35	14	916	2,543	7,531	1,203,872	11.25
% Black or African American in Segment																			
50-100%	1,668	7	1,022	2,902	7,743	783,738	11.04	8	957	2,911	7,751	763,994	10.31	8	944	2,918	7,979	761,308	10.14
10-50%	3,127	41	1,227	3,461	9,391	565,978	6.86	17	1,091	3,094	8,778	583,288	7.64	17	1,089	3,107	8,896	562,893	7.61
<10%	15,181	17	1,048	2,908	8,660	2,063,702	12.31	14	1,004	2,838	8,492	1,232,038	10.65	14	998	2,844	8,519	1,203,872	10.59
% Owner-Occupied DUs[1] in Segment																			
50-100%	15,244	27	1,186	3,295	9,478	2,063,702	11.16	14	1,119	3,172	9,283	1,232,038	9.83	14	1,115	3,179	9,300	1,203,872	9.75
10-50%	3,728	7	1,014	2,679	7,829	647,476	9.75	8	910	2,564	7,653	644,306	9.70	8	905	2,564	7,578	651,163	9.72
<10%	1,004	15	407	924	2,481	151,143	9.95	18	272	754	2,377	177,815	12.25	16	262	735	2,350	175,829	12.48
Combined Median Rent/Housing Value																			
1st Quintile	3,589	33	930	2,455	7,039	847,428	11.30	14	863	2,385	6,834	1,141,198	13.94	14	861	2,369	6,814	1,135,896	13.83
2nd Quintile	4,211	23	1,016	2,800	7,982	956,398	11.76	17	931	2,651	8,057	860,558	11.51	17	928	2,655	8,052	860,845	11.55
3rd Quintile	4,600	7	987	2,738	8,494	1,395,732	10.24	8	907	2,632	8,105	1,116,206	9.51	8	900	2,633	8,199	1,022,170	9.10
4th Quintile	4,155	17	1,260	3,492	9,416	487,944	6.34	24	1,174	3,434	9,335	453,464	6.48	23	1,161	3,450	9,407	469,525	6.51
5th Quintile	3,421	36	1,298	3,595	10,943	2,063,702	15.03	18	1,230	3,585	10,774	1,232,038	10.22	16	1,236	3,599	10,781	1,203,872	10.28

Table L.3 2011 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.ps and res.pr.ev) (continued)

Domain	*n*	Before res.pr.ps[1] (SDUWT*PRWT11*...*PRWT13)						After res.pr.ps[1] (SDUWT*PRWT11*...*PRWT14)						Final Weight: After res.pr.ev[1] (SDUWT*PRWT11*...*PRWT15)					
		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]	8,111	7	1,538	4,109	11,803	2,063,702	10.77	8	1,428	4,059	11,887	1,232,038	8.45	8	1,431	4,102	11,850	1,203,872	8.33
Medium to Small MSA[1]	10,252	17	930	2,451	7,005	647,476	8.42	14	857	2,354	6,834	716,475	9.32	14	852	2,346	6,874	743,129	9.39
Non-MSA,[1] Urban	458	40	745	1,675	5,301	847,428	30.88	19	575	1,450	4,173	1,141,198	53.33	18	584	1,449	4,172	1,135,896	52.87
Non-MSA,[1] Rural	1,155	33	608	1,854	5,757	956,398	17.34	25	543	1,741	6,013	860,558	16.73	24	553	1,735	6,052	860,845	16.92
Group Quarters																			
Group	288	36	563	1,263	2,832	47,921	3.82	17	372	1,098	2,461	64,472	5.07	18	363	1,097	2,374	64,840	4.92
Non-Group	19,688	7	1,086	3,047	8,807	2,063,702	11.18	8	1,022	2,932	8,660	1,232,038	10.02	8	1,016	2,929	8,647	1,203,872	9.96
Pair Relationship Domain[4]																			
Parent-Child (12-14)	2,049	83	2,403	5,084	11,380	114,526	2.85	77	2,397	5,221	11,896	123,384	2.93	77	2,406	5,181	11,939	123,301	2.95
Parent-Child (12-17)	3,842	49	2,597	5,538	12,397	183,335	3.02	27	2,500	5,644	12,801	175,429	3.16	26	2,481	5,624	12,868	177,360	3.17
Parent-Child (12-20)	4,582	49	2,717	5,984	14,089	198,867	3.11	27	2,610	6,081	14,365	188,037	3.27	26	2,586	6,106	14,336	177,360	3.26
Sibling (12-14)-Sibling (15-17)	2,350	17	591	1,151	2,286	31,948	2.49	18	575	1,179	2,400	26,393	2.52	16	573	1,171	2,392	25,589	2.53
Sibling (12-17)-Sibling (18-25)	2,618	17	872	1,738	3,588	35,735	2.34	31	789	1,683	3,617	28,659	2.34	31	782	1,678	3,621	25,152	2.32
Spouse-Spouse/ Partner-Partner	3,693	38	1,009	4,099	15,571	2,063,702	12.06	23	1,077	4,383	16,336	1,232,038	9.57	24	1,091	4,377	16,480	1,203,872	9.46
Spouse-Spouse/ Partner-Partner with Children (Younger Than 18)	1,730	40	939	2,959	9,001	2,063,702	25.47	48	1,356	4,202	12,421	1,232,038	14.19	50	1,372	4,313	12,438	1,203,872	14.25

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

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

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

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