

2014 NATIONAL SURVEY ON DRUG USE AND HEALTH

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

DISCLAIMER

SAMHSA provides links to other Internet sites as a service to its users and is not responsible for the availability or content of these external sites. SAMHSA, its employees, and contractors do not endorse, warrant, or guarantee the products, services, or information described or offered at these other Internet sites. Any reference to a commercial product, process, or service is not an endorsement or recommendation by SAMHSA, its employees, or contractors. For documents available from this server, the U.S. Government does not warrant or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed.

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

March 2016

This page intentionally left blank

2014 NATIONAL SURVEY ON DRUG USE AND HEALTH: QUESTIONNAIRE DWELLING UNIT-LEVEL AND PERSON PAIR- LEVEL SAMPLING WEIGHT CALIBRATION

Prepared for the 2014 Methodological Resource Book (Section 12)

Contract No. HHSS283201300001C
RTI Project No. 0213757.004.106.006
Deliverable No. 49

RTI Authors:

Matthew Westlake
Patrick Chen
Harper Gordek

SAMHSA Authors:

Matthew Williams
Art Hughes

RTI Project Director:

David Hunter

SAMHSA Project Officer:

Peter Tice

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

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

Prepared by RTI International, Research Triangle Park, North Carolina

March 2016

Recommended Citation: Center for Behavioral Health Statistics and Quality. (2016). *2014 National Survey on Drug Use and Health: Methodological Resource Book (Section 12, Questionnaire Dwelling Unit-Level and Person Pair-Level Sampling Weight Calibration)*. Substance Abuse and Mental Health Services Administration, Rockville, MD.

This page intentionally left blank

Preface and Acknowledgments

This report documents the method of weight calibration used for producing the final set of questionnaire dwelling unit and pair weights for the National Survey on Drug Use and Health (NSDUH) data from 2014. The weighting team faced several challenges in this task and was able to address them by resorting to innovative modifications of certain basic statistical ideas, which are listed below.

- Under Brewer's method, high weights may occur because of small pair selection probabilities. In any calibration exercise, some treatment of extreme value (ev) in weights is needed, but there is a danger of introducing too much bias by over-treatment. In the generalized exponential model (GEM), which is described in detail in the NSDUH Methodological Resource Book person-level sampling weight calibration report (Center for Behavioral Health Statistics and Quality, 2016b), ev control is built in, but one needs to define suitable ev domains so that not too many evs are defined. If too many design variables are used to define ev domains, then each domain will be very sparse and will not be of much use in defining thresholds for ev. As in past surveys, a hierarchy of domains was defined using pair age (each pair member being in one of the three categories: 12 to 25, 26 to 49, and 50+) and number of people aged 12 to 25 in the household, state, and clusters of states (see Section 5.2 for details).
- Control of evs in weights helps reduce instability of estimates to some extent, but there is a need for methods that do not introduce much bias. Following the famous suggestion of Hajek (1971) in his comments on Basu's fabled example of circus elephants, we performed ratio adjustment (a form of poststratification) to estimated totals obtained from the household data on the number of people belonging to the pair domain of interest. This was implemented in a multivariate manner to get one set of final weights.
- In the absence of a suitable source of poststratification controls for the person pair-level weights and the household-level weights, the inherent two-phase nature of the survey design was capitalized upon to estimate these controls from the first phase of the large screener sample. The first-phase sample weight was poststratified to person-level U.S. Census Bureau counts to get more efficient estimated counts for pair and household data.
- The problem of multiplicities complicated the issue of providing one set of final weights. When dealing with person-level parameters involving drug-related behaviors among members of the same household, it is possible for an individual to manifest himself or herself in the pair sample through different pairs. To avoid overcounting, the pair weights have to be divided by multiplicity factors, which tend to be domain specific. For this reason, multiplicity factors for a key set of pair analysis domains also are produced along with a set of final calibrated pair weights.
- Missing items in the respondent questionnaire led to imputation for deriving pair relationships, multiplicity factors, and household counts for Hajek adjustments.

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

This report was prepared for the Substance Abuse and Mental Health Services Administration (SAMHSA), Center for Behavioral Health Statistics and Quality, by RTI International (a registered trademark and a trade name of Research Triangle Institute). Contributors to this report at RTI include Debbie Bond, Claudia Clark, and Margaret Smith. The authors are grateful to Eunice Park-Lee of SAMHSA for her useful comments and suggestions.

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

Chapter	Page
List of Terms and Abbreviations	xi
1. Introduction.....	1
2. Questionnaire Dwelling Unit and Pair Selection Probabilities.....	7
2.1 Pair Selection Probability	9
2.1.1 Case I: DUs with $S \geq 2$	9
2.1.2 Case II: DUs with $S < 2$	10
2.2 Questionnaire Dwelling Unit Selection Probability	11
3. Brief Description of the Generalized Exponential Model	13
4. Predictor Variables for the Questionnaire Dwelling Unit and Pair Weight Calibration via the Generalized Exponential Model.....	15
4.1 Questionnaire Dwelling Unit Weight Calibration	15
4.2 Pair Weight Calibration	16
5. Definition of Extreme Weights.....	21
5.1 Questionnaire Dwelling Unit Extreme Weight Definition	21
5.2 Person Pair Extreme Weight Definition	22
6. Weight Calibration at Questionnaire Dwelling Unit and Pair Levels.....	25
6.1 Phase I SDU-Level Weight Components	30
6.2 QDU Weight Components.....	31
6.2.1 QDU Weight Component #12: Inverse of Selection Probability of at Least One Person in the Dwelling Unit	31
6.2.2 QDU Weight Component #13: Selected QDU Poststratification to SDU-Based Control Totals	31
6.2.3 QDU Weight Component #14: Respondent QDU Nonresponse Adjustment.....	32
6.2.4 QDU Weight Component #15: Respondent QDU Poststratification to SDU-Based Control Totals.....	32
6.2.5 QDU Weight Component #16: Respondent QDU Extreme Value Adjustment.....	32
6.3 Pair-Level Weight Components.....	32
6.3.1 Pair Weight Component #12: Inverse of Selection Probability of a Person Pair in the Dwelling Unit	32
6.3.2 Pair Weight Component #13: Selected Pair Poststratification to SDU-Based Control Totals	32
6.3.3 Pair Weight Component #14: Respondent Pair Nonresponse Adjustment.....	33
6.3.4 Pair Weight Component #15: Respondent Pair Poststratification to SDU-Based Control Totals	33
6.3.5 Pair Weight Component #16: Respondent Pair Extreme Weight Adjustment.....	34

Table of Contents (continued)

Chapter	Page
7. Evaluation of Calibration Weights.....	35
7.1 Response Rates	35
7.2 Proportions of Extreme Values and Outwinsors.....	35
7.3 Slippage Rates.....	36
7.4 Weight Adjustment Summary Statistics	36
7.5 Sensitivity Analysis of Drug Use Estimates	37
References.....	49

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

List of Tables

Table	Page
1.1 2010–2014 NSDUH Sample Sizes	3
1.2 Pair Domains.....	4
2.1 Building Blocks of the QDU and Person Pair Samples: Dwelling Units and People in the 2010–2014 NSDUHs	8
6.1 Sample Size, by Model Group at QDU and Pair Levels.....	30
7.1 Estimates of Totals and SEs for Domains of Interest Based on QDU Sample: 2014.....	38
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: 2014	39
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: 2014.....	40
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: 2014	41
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: 2014	42
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: 2014.....	43
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: 2014	44
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: 2014.....	45
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: 2014	46
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: 2014	47
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: 2014	48

List of Tables (continued)

Table	Page
C.1a	2014 QDU Weight GEM Modeling Summary (Model Group 1: Northeast) C-11
C.1b	2014 Distribution of Weight Adjustment Factors and Weight Products (Model Group 1: Northeast) C-12
C.2a	2014 QDU Weight GEM Modeling Summary (Model Group 2: Midwest)..... C-19
C.2b	2014 Distribution of Weight Adjustment Factors and Weight Products (Model Group 2: Midwest)..... C-20
C.3a	2014 QDU Weight GEM Modeling Summary (Model Group 3: South) C-27
C.3b	2014 Distribution of Weight Adjustment Factors and Weight Products (Model Group 3: South)..... C-28
C.4a	2014 QDU Weight GEM Modeling Summary (Model Group 4: West) C-35
C.4b	2014 Distribution of Weight Adjustment Factors and Weight Products (Model Group 4: West)..... C-36
D.1	2014 NSDUH QDU-Level Response Rates D-1
E.1	2014 NSDUH Selected QDU-Level Proportions of Extreme Values and Outwinsors E-1
E.2	2014 NSDUH Respondent QDU-Level Proportions of Extreme Values and Outwinsors E-3
F.1	2014 NSDUH QDU-Level Slippage Rates F-1
G.1	2014 NSDUH Selected QDU-Level Weight Summary Statistics G-1
G.2	2014 NSDUH Respondent QDU-Level Weight Summary Statistics G-3
H.1a	2014 Pair Weight GEM Modeling Summary (Model Group 1: Northeast and South)..... H-7
H.1b	2014 Distribution of Weight Adjustment Factors and Weight Products (Model Group 1: Northeast and South) H-8
H.2a	2014 Pair Weight GEM Modeling Summary (Model Group 2: Midwest and West).... H-17
H.2b	2014 Distribution of Weight Adjustment Factors and Weight Products (Model Group 2: Midwest and West)..... H-18
I.1	2014 NSDUH Person Pair-Level Response Rates..... I-1
J.1	2014 NSDUH Selected Pair-Level Proportions of Extreme Values and Outwinsors J-1
J.2	2014 NSDUH Respondent Pair-Level Proportions of Extreme Values and Outwinsors J-3
J.3	2014 NSDUH Respondent Pair-Level Proportions of Extreme Values and Outwinsors J-5
K.1	2014 NSDUH Respondent Pair-Level Slippage Rates K-1
L.1	2014 NSDUH Selected Pair-Level Weight Summary Statistics..... L-1
L.2	2014 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.nr) L-4
L.3	2014 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.ps and res.pr.ev) L-6

List of Exhibits

Exhibit	Page
1.1 QDU and Pair Sampling Weight Calibration Steps.....	5
4.1 Definitions of Levels for QDU-Level Calibration Modeling Variables.....	17
4.2 Definitions of Levels for Pair-Level Calibration Modeling Variables.....	18
6.1 Summary of 2014 NSDUH QDU Sample Weight Components.....	27
6.2 Summary of 2014 NSDUH Person Pair Sample Weight Components.....	28
6.3 U.S. Census Bureau Regions/Model Groups.....	29
C.1 Definitions of Levels for QDU-Level Calibration Modeling Variables.....	C-2
C.2 Covariates for 2014 NSDUH Questionnaire Dwelling Unit Weights.....	C-8
C.1.1 Covariates for 2014 NSDUH Questionnaire Dwelling Unit Weights (sel.qdu.ps) Model Group 1: Northeast.....	C-14
C.1.2 Covariates for 2014 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 1: Northeast.....	C-15
C.1.3 Covariates for 2014 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 1: Northeast.....	C-16
C.2.1 Covariates for 2014 NSDUH Questionnaire Dwelling Unit Weights (sel.qdu.ps) Model Group 2: Midwest.....	C-22
C.2.2 Covariates for 2014 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 2: Midwest.....	C-23
C.2.3 Covariates for 2014 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 2: Midwest.....	C-24
C.3.1 Covariates for 2014 NSDUH Questionnaire Dwelling Unit Weights (sel.qdu.ps) Model Group 3: South.....	C-30
C.3.2 Covariates for 2014 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 3: South.....	C-31
C.3.3 Covariates for 2014 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 3: South.....	C-32
C.4.1 Covariates for 2014 NSDUH Questionnaire Dwelling Unit Weights (sel.qdu.ps) Model Group 4: West.....	C-38
C.4.2 Covariates for 2014 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr) Model Group 4: West.....	C-39
C.4.3 Covariates for 2014 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps) Model Group 4: West.....	C-40
H.1 Definitions of Levels for Pair-Level Calibration Modeling Variables.....	H-2
H.2 Covariates for 2014 NSDUH Pair Weights.....	H-4
H.1.1 Covariates for 2014 NSDUH Pair Weights (sel.pr.ps) Model Group 1: Northeast and South.....	H-10
H.1.2 Covariates for 2014 NSDUH Pair Weights (res.pr.nr) Model Group 1: Northeast and South.....	H-11
H.1.3 Covariates for 2014 NSDUH Pair Weights (res.pr.ps) Model Group 1: Northeast and South.....	H-12
H.1.4 Covariates for 2014 NSDUH Pair Weights (res.pr.ev) Model Group 1: Northeast and South.....	H-13

List of Exhibits (continued)

Exhibit	Page
H.2.1 Covariates for 2014 NSDUH Pair Weights (sel.pr.ps) Model Group 2: Midwest and West.....	H-20
H.2.2 Covariates for 2014 NSDUH Pair Weights (res.pr.nr) Model Group 2: Midwest and West.....	H-21
H.2.3 Covariates for 2014 NSDUH Pair Weights (res.pr.ps) Model Group 2: Midwest and West.....	H-22
H.2.4 Covariates for 2014 NSDUH Pair Weights (res.pr.ev) Model Group 2: Midwest and West.....	H-23

List of Terms and Abbreviations

DU	Dwelling unit.
Ev	Extreme value. See Sections 5.1 and 5.2 for more detail.
GEM	Generalized exponential model. See Chapter 3 for more detail.
Household-level person count	The number of pairs associated with a given domain in a given household. These counts are used as control totals in the poststratification step. See Chapter 11 in the editing and imputation report (Center for Behavioral Health Statistics and Quality, 2016a) for details on how these counts are created, and Chapter 4 for details on their use in poststratification.
IQR	Interquartile range.
Multiplicity factor	The number of pairs associated with a given respondent in a given domain. See Center for Behavioral Health Statistics and Quality (2016a) for more detail.
Nr	Nonresponse.
Outwisor	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 because people from the same household often exhibit the same or similar characteristics and behavioral patterns. The intra-class correlation found among members of the same household leads to a clustering effect on the variance of estimates resulting in less precise estimates compared with estimates of the same sample size from a simple random sample. Selecting only one person per household avoids this clustering effect on the variance. The "one person per household" sampling approach, however, precludes the opportunity to gather information about the relationships among household members. In the National Survey on Drug Use and Health (NSDUH),² we allow for a richer analytic capability of a survey designed to ensure a positive pairwise probability of selection among all eligible household members in each sample household. Achieving positive probabilities for all pairs within sampled households permits unbiased estimation of the within-dwelling-unit component of variance. Besides providing efficient data collection, this sampling method also facilitates the study of the relationships of social behaviors among members of the same household. This report documents the methodology and development of calibrated weights for the second objective, the study of behavioral relationships among people residing in the same household. The report also describes the development of questionnaire dwelling unit (QDU) weights, which are of independent interest for studying household-level characteristics and also are needed for producing household count estimates of the number of people belonging to pair relationship domains for use as poststratification controls for pair weights.

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

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

² This report presents information from the 2014 National Survey on Drug Use and Health (NSDUH). Prior to 2002, the survey was called the National Household Survey on Drug Abuse (NHSDA).

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

Under a stratified design with states serving as the primary strata and state sampling (SS) regions serving as the secondary strata, census tracts, census block groups, segments within census block groups, and dwelling units (DUs) within segments were each selected using probability proportional to size sampling. Also new in the 2014–2017 design, is the incorporation of census block groups at the second stage of selection to potentially reduce sampling variance and facilitate moving to an address-based sampling design in the future, if desired. NSDUH is sometimes referred to as a two-phase sample where the first phase consisted of a large number of screener dwelling units (SDUs, about 200,000) selected to ensure that various age groups (five in all: 12 to 17, 18 to 25, 26 to 34, 35 to 49, and 50+) of eligible individuals were represented adequately in the second phase. In the 2014–2017 NSDUH design, added focus (greater sample) was placed on the 26 or older age group to improve estimates of drug use and related health measures for this population. Unlike the 2005–2013 NSDUHs, which allocated state sample equally across the age categories 12 to 17, 18 to 25 and 26 or older, in 2014, the sample was allocated with 25 percent for 12 to 17, 25 percent for 18 to 25, 15 percent for 26 to 34, 20 percent for 35 to 49, and 15 percent for 50 or older. Information collected from SDUs also provided estimates of population controls (as in two-phase sampling) for calibration at levels (such as pair and QDU) for which suitable U.S. Census Bureau–based controls were not available. The second phase consisted of the selection of zero, one, or two people from each selected SDU using a modification of Brewer's method such that prescribed sampling rates for the five age groups in each state were achieved with high selection rates for youths (12 to 17) and young adults (18 to 25). [Table 1.1](#) shows the eligible number of selected and responding SDUs, QDUs, pairs, and people for each of the 5 years (2010–2014). 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 2010–2014 NSDUH Sample Sizes

Sample Unit		2010	2011	2012	2013	2014
SDU	Selected	166,532	179,293	178,586	190,067	154,533
	Completed	147,010	156,048	153,873	160,325	127,605
QDU	Selected	58,702	61,441	60,621	61,634	64,796
	Completed	48,113	50,133	48,850	48,896	49,672
Pair	Selected	26,295	27,095	27,035	27,108	26,844
	Completed	19,691	19,976	19,459	18,942	18,229
Person	Selected	84,997	88,536	87,656	88,742	91,640
	Completed	67,804	70,109	68,309	67,838	67,901

Note: The 2010 sample sizes reflect the removal of falsified cases found in Pennsylvania and Maryland. The 2011–2014 samples were not affected. For additional information, see Section B.3.5 in Appendix B of the 2011, 2012, and 2013 national findings reports, and Section B.3.5 in the 2014 methodological summary and definitions (Center for Behavioral Health Statistics and Quality, 2012, 2013, 2014, 2015b).

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

In estimation for pair domains, two major problems arise: one is that of multiplicities because, for a given domain defined by the pair relationship, when the parameter of interest is at the person level, several pairs in the household could be associated with the same person. For example, analysts are interested in an outcome at the person level, the proportion of children who use drugs and whose parents report talking to them about drugs, where the focus is on the child in a parent-child pair. Several parent-child pairs in the household could be associated with the same child. If the household has two parents, the selected child has two inclusion possibilities (one with each parent) in the set of all such parent-child pairs (Center for Behavioral Health Statistics and Quality [CBHSQ], 2016a). The other problem is that of extreme weights that may arise due to small selection probabilities for certain pair age groups, which may lead to unstable estimates. Each of these issues is discussed in turn.

If several pairs in the household are associated with the same person, it is necessary to use the average measure of behavior relationships for each member, which gives rise to multiplicities. Thus, the pair weights need to be divided by the person-level multiplicity factors for each domain of interest, and, therefore, multiplicity factors need to be produced along with the final set of calibrated weights. Because it is not straightforward to create these multiplicities,

analyses would have to be necessarily limited to pair relationships where the multiplicities were produced a priori. It was anticipated that analyses of interest would be limited to 14 pair domains, listed in Table 1.2. Because no multiplicity was necessary for the spouse-spouse/partner-partner pair relationships (by definition, each pair member could have only one partner or one spouse), multiplicity factors were produced for only 12 of these domains. Note that a single pair relationship might have two domains associated with it, because the parameter of interest might be associated with only one member of the pair (the "focus" member), and the multiplicity would differ depending upon which pair member was the focus member.

Table 1.2 Pair Domains

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

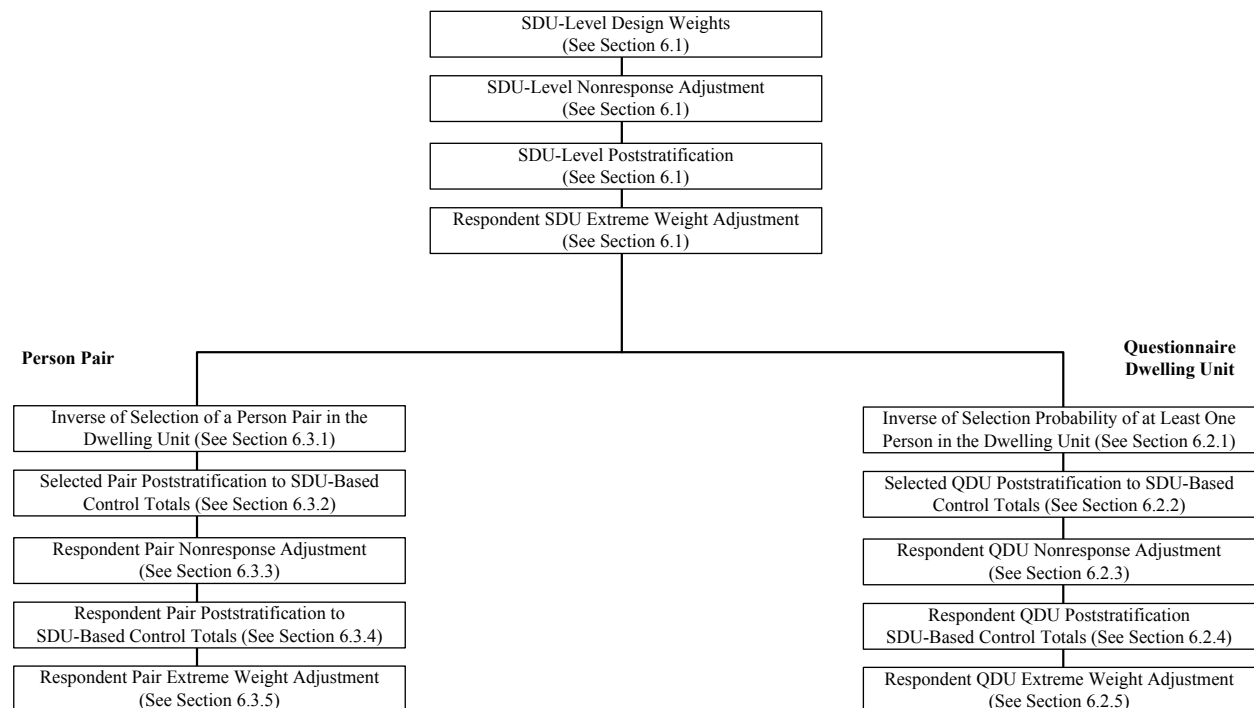
Some of the multiplicities, including counts of all possible pairs in a household for a given domain, were used for poststratification. Details are provided in Chapter 4. Prior editions of this report contained a chapter on editing and imputation of pair relationships, multiplicity factors, and household-level 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 NSDUH Methodological Resource Book editing and imputation report (CBHSQ, 2016a). Special consideration is required for analysis of pair-level data, and details can be found in *How To Prepare and Analyze Pair Data in the National Survey on Drug Use and Health* (CBHSQ, in press). See Appendix M for a draft excerpt from this manual on inferential population and multiplicities.

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

poststratification then was followed by a repeat poststratification to further control the extreme weights by imposing separate bound restrictions on the initially identified extreme weights.

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

Exhibit 1.1 QDU and Pair Sampling Weight Calibration Steps



In dealing with extreme weights, it is assumed that they arise due to design (due to an imperfect frame, assignment of very small selection probabilities to some units, or a big weight adjustment factor after calibration) so that they make the sample representative of the population and, hence, do not introduce bias. The only problem is that they may lead to highly unstable estimates similar to the problem of Basu's circus elephants³ (Hajek, 1971). So, we need to

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

perform some treatment (such as winsorization⁴) within suitably defined extreme weight domains such that these domains contain units possibly from different strata but with similar sample selection probabilities to avoid the occurrence of extreme weights due to a mix of different designs. The domains must be large enough (e.g., at least size 30) to be able to define evs according to the domain-specific weight distribution. Any ev treatment to increase precision of estimates would introduce some bias. However, this bias can be reduced considerably if the ev treatment is performed under calibration controls. This is what the built-in ev control in GEM tries to accomplish.

It follows that the definition of extreme weight domains should depend on factors that affect the selection probabilities of units in the sample, such as state- and age-specific sampling rates, segment selection probabilities, pair age-specific selection probabilities, and household composition. If one tries to define extreme weight domains by taking account of all these factors via cross-classification, it will lead to too many domains with insufficient observations. That is why it is difficult to define suitable extreme weight domains for pair data. In the case of person-level weights, it was less difficult, because state by age group suitably captured the extreme weight domain requirements. The definition of extreme weight domains for pair-level weighting used in the 2014 survey was the same as the one used in the 1999–2013 surveys. The domains were defined as the cross-classification of state, pair age,⁵ and number of people aged 12 to 25 in a household. In particular, the pair age was defined by the age groups of each pair member according to the age categories of 12 to 25, 26 to 49, and 50 or older (resulting in six pair age categories), and the number of people aged 12 to 25 were categorically defined as zero, one, and two or more. For more details, see Chapter 5.

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 2002–2013 National Surveys on Drug Use and Health (NSDUHs),⁶ the 2014 NSDUH had a two-phase design and used a computer-assisted interviewing method. There were five stages of selection: census tracts, census block groups, segments within census block groups, dwelling units (DUs) within segments, and people within dwelling units. Any two survey-eligible people had some nonzero chance of being selected and, when both were selected, they formed a within-household pair. This design feature is of interest to NSDUH researchers because, for example, it allows analysts to examine how the drug use propensity of an individual (in a family) relates to the drug use propensity of other members residing in the same dwelling unit (Center for Behavioral Health Statistics and Quality, 2015a).

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

Starting with the 2002 NSDUH and continuing through 2014, the pair-sampling algorithm was modified to increase the number of pairs selected in the sample. DUs with the sum of person selection probabilities greater than or equal to 2 were treated the same as in previous survey years. However, DUs where the sum of person-level selection probabilities was less than 2 received a slightly different treatment that increased the chance for selecting a pair of real people. Section 2.1 describes the selection process for both types of DUs.

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

⁶ This report presents information from the 2014 National Survey on Drug Use and Health (NSDUH). Prior to 2002, the survey was called the National Household Survey on Drug Abuse (NHSDA).

Table 2.1 Building Blocks of the QDU and Person Pair Samples: Dwelling Units and People in the 2010–2014 NSDUHs

Domain	2010			2011			2012			2013			2014		
	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	166,532	147,010	88.28	179,293	156,048	87.04	178,586	153,873	86.16	190,067	160,325	84.35	154,533	127,605	82.57
QDUs															
Total QDUs	58,702	48,113	81.96	61,441	50,133	81.60	60,621	48,850	80.58	61,634	48,896	79.33	64,796	49,672	76.66
People															
Total People	84,997	67,804	79.77	88,536	70,109	79.19	87,656	68,309	77.93	88,742	67,838	76.44	91,640	67,901	74.10
12-17	25,908	21,960	84.76	27,911	23,510	84.23	27,147	22,473	82.78	27,630	22,494	81.41	21,392	17,007	79.50
18-25	28,164	22,793	80.93	28,589	22,876	80.02	28,639	22,529	78.67	28,921	22,214	76.81	21,726	16,449	75.71
26-34	8,545	6,780	79.34	8,323	6,543	78.61	8,304	6,484	78.08	8,210	6,310	76.86	14,004	10,252	73.21
35-49	12,979	9,668	74.49	12,220	9,149	74.87	12,364	9,076	73.41	12,566	9,058	72.08	19,065	13,590	71.28
50+	9,401	6,603	70.24	11,493	8,031	69.88	11,202	7,747	69.16	11,415	7,762	68.00	15,453	10,603	68.61
Non-Pairs⁴															
Total Non-Pairs	120,715	28,435	N/A	128,953	30,157	N/A	126,838	29,391	N/A	133,217	29,954	N/A	90,443	31,443	N/A
0,0	88,308	N/A	N/A	94,607	N/A	N/A	93,252	N/A	N/A	98,691	N/A	N/A	62,809	N/A	N/A
Total Singletons	32,407	28,435	87.74	34,346	30,157	87.80	33,586	29,391	87.51	34,526	29,954	86.76	37,952	31,443	82.85
0, 12-17	8,595	7,906	91.98	9,402	8,651	92.01	9,017	8,277	91.79	9,420	8,574	91.02	4,850	4,704	96.99
0, 18-25	10,093	9,270	91.85	10,306	9,497	92.15	10,325	9,461	91.63	10,535	9,475	89.94	7,250	6,647	91.68
0, 26-34	3,914	3,418	87.33	3,930	3,386	86.16	3,856	3,327	86.28	3,914	3,367	86.02	7,460	6,034	80.88
0, 35-49	4,659	3,843	82.49	4,431	3,704	83.59	4,368	3,645	83.45	4,506	3,736	82.91	8,074	6,450	79.89
0, 50+	5,146	3,998	77.69	6,277	4,919	78.37	6,020	4,681	77.76	6,151	4,802	78.07	10,318	7,608	73.74
Pairs⁵															
Total Pairs ⁵	26,295	19,691	74.88	27,095	19,976	73.73	27,035	19,459	71.98	27,108	18,942	69.88	26,844	18,229	67.91
12-17, 12-17	4,160	3,525	84.74	4,649	3,885	83.57	4,507	3,668	81.38	4,535	3,609	79.58	3,070	2,407	78.40
12-17, 18-25	3,444	2,718	78.92	3,756	2,921	77.77	3,627	2,759	76.07	3,662	2,754	75.20	2,443	1,832	74.99
12-17, 26-34	922	752	81.56	834	685	82.13	825	658	79.76	811	621	76.57	1,297	941	72.55
12-17, 35-49	3,948	3,044	77.10	3,855	2,918	75.69	3,813	2,812	73.75	3,834	2,756	71.88	5,530	3,940	71.25
12-17, 50+	679	493	72.61	766	565	73.76	851	631	74.15	833	571	68.55	1,132	776	68.55
18-25, 18-25	5,502	4,165	75.70	5,476	4,015	73.32	5,476	3,901	71.24	5,478	3,795	69.28	3,743	2,585	69.06
18-25, 26-34	1,140	851	74.65	1,049	768	73.21	1,079	794	73.59	1,034	690	66.73	1,378	870	63.13
18-25, 35-49	1,639	1,098	66.99	1,469	994	67.67	1,582	1,053	66.56	1,561	990	63.42	1,906	1,180	61.91
18-25, 50+	844	537	63.63	1,057	666	63.01	1,074	660	61.45	1,173	715	60.95	1,263	750	59.38
26-34, 26-34	903	621	68.77	858	599	69.81	880	604	68.64	822	581	70.68	1,356	865	63.79
26-34, 35-49	512	354	69.14	492	314	63.82	469	320	68.23	489	307	62.78	737	442	59.97
26-34, 50+	251	162	64.54	302	192	63.58	315	177	56.19	318	163	51.26	420	235	55.95
35-49, 35-49	886	542	61.17	748	474	63.37	833	487	58.46	857	512	59.74	1,160	658	56.72
35-49, 50+	449	245	54.57	477	271	56.81	466	272	58.37	462	245	53.03	498	262	52.61
50+, 50+	1,016	584	57.48	1,307	709	54.25	1,238	663	53.55	1,239	633	51.09	911	486	53.35

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

Note: The 2010 sample sizes reflect the removal of falsified cases found in Pennsylvania and Maryland. The 2011–2014 samples were not affected. For additional information, see Section B.3.5 in Appendix B of the 2011, 2012, and 2013 national findings reports, and Section B.3.5 of the 2014 methodological summary and definitions (Center for Behavioral Health Statistics and Quality, 2012, 2013, 2014, 2015b).

¹ Selected pairs are based on the screener age.

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

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

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

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

2.1 Pair Selection Probability

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

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

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

by setting K at

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

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

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

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

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

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

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

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

for all values of i .

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

2.1.2 Case II: DUs with $S < 2$

If the sum S of person-level selection probabilities was less than 2, the method used in survey years 1991–2001 consisted of dividing $2 - S$ equally among the three dummy people added to the household, and then used Brewer's method to select a pair, as in Case I. However, if the household had two or more people, we preferred a pair of real people to have a greater chance of being selected. To achieve this goal, the individual selection probabilities, $P_{h(i)}$, were scaled upward by the factor F_s such that their sum came close to but did not exceed 2 and such that each person selection probability did not exceed the maximum allowed probability of 0.99. Thus, denoting the revised person selection probabilities by $P'_{h(i)}$, the factor F_s is given by

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

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

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

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

by setting K' at

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

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

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

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

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

Note that we now have $\sum_{j \neq i} P'_{h(ij)} = P'_{h(i)}$. To maintain the original person selection

probabilities despite the scale adjustment by F_s , we modified Brewer's method as follows. First, draw a random number, R , from a uniform (0,1) distribution. If $R \leq 1/F_s$, then select a pair using Brewer's method based on formula (2.6). However, if $R > 1/F_s$, then no people are selected from the household. In this way, the probability for selecting a pair (i,j) in household h becomes $P^*_{h(ij)} = P'_{h(ij)}/F_s$, which, in turn, gives the original person selection probabilities, $P_{h(i)}$. Unlike Case I, where a pair of eligible people was always selected, this adjusted selection scheme allows for zero, one, or two people to be selected from a DU.

2.2 Questionnaire Dwelling Unit Selection Probability

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

The QDU selection probability was defined as

$$P_h^* = (1 - P_{h(00)}^*), \quad (2.8)$$

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

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

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

This page intentionally left blank

3. Brief Description of the Generalized Exponential Model

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

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

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

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

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

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

To implement GEM, several steps need to be followed: (1) define and create all the covariates; (2) define the extreme weights; (3) fit the GEM model. The details of practical aspects of GEM implementation can be found in Chapters 4 and 5 of this report and Chapter 4 of the National Survey on Drug Use and Health Methodological Resource Book person-level sampling weight calibration report (Center for Behavioral Health Statistics and Quality, 2016b).

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

We note that unlike the person-level weight calibration, the control totals for the questionnaire dwelling unit (QDU)-level and person pair-level poststratification are not available from the U.S. Census Bureau. A way around this problem is to take advantage of the two-phase nature of the design, in which the screener data provide a large sample containing demographic information that can be used to derive control totals for the QDU-level and person pair-level sampling weight calibrations, as well as for the selected person poststratification adjustment. The stability of control totals from the screener dwelling unit (SDU)-level data can be improved by poststratification of the SDU sample using person-level counts from the census. This was indeed done and is documented in the National Survey on Drug Use and Health Methodological Resource Book person-level sampling weight calibration report (Center for Behavioral Health Statistics and Quality, 2016b).

4.1 Questionnaire Dwelling Unit Weight Calibration

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

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

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

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

4.2 Pair Weight Calibration

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

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

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

<p>Age^b 1: 12-17, 2: 18-25, 3: 26-34, 4: 35-49, 5: 50+¹</p> <p>Gender^b 1: Male, 2: Female¹</p> <p>Group Quarter Indicator^a 1: College Dorm, 2: Other Group Quarter, 3: Non-Group Quarter¹</p> <p>Hispanicity^b 1: Hispanic or Latino, 2: Non-Hispanic or Latino¹</p> <p>Household Size^b Continuous Variable Count of Individuals Rostered with DU</p> <p>Household Type (Ages of People 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</p> <p>Percentage of Owner-Occupied Dwelling Units in Segment (% Owner-Occupied)^a 1: 50-100%,¹ 2: 10-<50%, 3: 0-<10%</p> <p>Percentage of Segments That Are Black or African American^a 1: 50-100%, 2: 10-<50%, 3: 0-<10%¹</p> <p>Percentage of Segments That Are Hispanic or Latino^a 1: 50-100%, 2: 10-<50%, 3: 0-<10%¹</p> <p>Population Density^a 1: MSA 1,000,000 or More, 2: MSA Less than 1,000,000, 3: Non-MSA Urban, 4: Non-MSA Rural¹</p> <p>Quarter^{a,b} 1: Quarter 1, 2: Quarter 2, 3: Quarter 3, 4: Quarter 4¹</p> <p>Race (3 Levels)^b 1: White,¹ 2: Black or African American, 3: Other</p> <p>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</p> <p>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</p> <p>Relation to Householder^a 1: Householder or Spouse, 2: Child, 3: Other Relative, 4: Nonrelative¹</p> <p>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 Quintile¹</p> <p>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: Wisconsin¹ 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</p> <p>State/Region^{a,3} Model Group 1: 1: New York, 2: Pennsylvania, 3: Other¹ Model Group 2: 1: Illinois, 2: Michigan, 3: Ohio, 4: Other¹ Model Group 3: 1: Florida, 2: Texas, 3: Other¹ Model Group 4: 1: California, 2: Other¹</p>

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 people in the household.

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 Quarter¹

Household Size

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

Pair Age (15 Levels)

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

Pair Age (6 Levels)

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

Pair Age (3 Levels)

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

Pair Gender

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

Pair Race/Ethnicity (10 Levels)

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

Pair Race/Ethnicity (5 Levels)

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

Pair Race/Ethnicity (4 Levels)

1: Two or More Races Pair or Other and Other, 2: Hispanic or Latino Pair, 3: Black or African-American Pair, 4: White Pair¹

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 Quintile¹

Population Density

1: MSA 1,000,000 or More, 2: MSA Less than 1,000,000, 3: Non-MSA Urban, 4: Non-MSA Rural¹

Quarter

1: Quarter 1, 2: Quarter 2, 3: Quarter 3, 4: Quarter 4¹

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

States³

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

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

Pair Relationship Associated with Multiplicity

1: Parent-Child (12-14)*

2: Parent-Child (12-17)*

3: Parent-Child (12-20)*

4: Parent*-Child (12-14)

5: Parent*-Child (12-17)

6: Parent*-Child (12-20)

7: Sibling (12-14)-Sibling (15-17)*

8: Sibling (12-17)-Sibling (18-25)*

9: Spouse-Spouse/Partner-Partner

10: Spouse-Spouse/Partner-Partner with Children (Younger than 18)

DU = dwelling unit; MSA = metropolitan statistical area.

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

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

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

* The pair member focused on.

This page intentionally left blank

5. Definition of Extreme Weights

An important feature of the generalized exponential model (GEM) is the built-in provision of extreme value (ev) treatment. Sampling weights are often classified as extreme (high or low) if they fall outside the interval, median $\pm 3 \times$ interquartile range (IQR). The interval is set for prespecified domains defined usually by design variables corresponding to deep stratification.⁸ Similar to previous National Surveys on Drug Use and Health (NSDUHs), for the GEM modeling used in the 2014 NSDUH, a more conservative (narrower) interval was defined, median $\pm 2.5 \times$ IQR. The narrower interval better prevents the adjusted weights from crossing the standard interval boundaries by treating weights near but not outside the commonly used boundaries (i.e., those that have the most potential to become extreme) as extreme as well.

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

5.1 Questionnaire Dwelling Unit Extreme Weight Definition

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

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

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

⁸ Deep stratification refers to the stratification that was used in the sample design. In the case of the 2014 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 2014 survey. This DU gave rise to a selected pair of respondents, one aged 59 and the other aged 36. The selection probability in this DU was 0.19374 for a respondent aged 50 or older and 0.33542 for a respondent aged 35 to 49. Using the formula (2.6) in Chapter 2, the pair selection probability was computed to be 0.00174. Therefore, the inverse of the selection probability, the pair-level design weight, was 574.07. Thus, a small pair selection probability can create a high initial weight, which is the product of the screener dwelling unit weight and the person pair design-based weight.

As mentioned in the introduction, it turns out to be difficult to select suitable domains for defining extreme weights for pair-level data. However, as was done for the 1999–2013 surveys, the extreme weight definition was based on the following hierarchy of domains:

1. Pair age group⁹ (with three age categories, 12 to 25, 26 to 49, and 50+) by number (0, 1, 2+) of people aged 12 to 25 in the household;
2. State cluster (with five levels [explained below]) by pair age group by number (0, 1, 2+) of people aged 12 to 25 in the household;
3. State cluster (with three levels [explained below]) by pair age group by number (0, 1, 2+) of people aged 12 to 25 in the household; and
4. State by pair age group by number of people aged 12 to 25 (0, 1, 2+) in the household.

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

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

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

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

This page intentionally left blank

6. Weight Calibration at Questionnaire Dwelling Unit and Pair Levels

The 2014 National Survey on Drug Use and Health (NSDUH) was based on probability sampling so that valid inferences can be made from survey findings about the target population. Probability sampling refers to sampling in which every unit on the frame is given a known, nonzero probability for inclusion in the survey. This is required for unbiased estimation of the population total. The assumption of nonzero inclusion probability for every pair of units in the frame also is required for unbiased variance estimation. The 2014–2017 NSDUH sample design plans slightly modify the 2005–2013 approach, such that the basic sampling plan involved five stages of selection across two phases of design: within Phase I, (1) the selection of census tracts and (2) census blocks within each state sampling (SS) region, (3) the selection of subareas or segments (comprising U.S. Census Bureau blocks) within SS regions; (4) the selection of dwelling units (DUs) within these subareas; and, finally, within Phase II, (5) the selection of eligible individuals within DUs. Specific details of the sample design and selection procedures for the sample and changes to the design for this year can be found in the 2014 NSDUH Methodological Resource Book (MRB) sample design report (Center for Behavioral Health Statistics and Quality [CBHSQ], 2015a).

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

The final sample weights for Phase I screener dwelling units (SDU) and Phase II QDU, person, and pair levels for the 2014 samples consisted of products of several factors, each representing either a probability of selection at some particular stage or some form of ev, nr, or poststratification (ps) calibration adjustment. In the following sections, we describe the QDU and pair weight components in greater detail. In summary, the first 11 factors were defined for all SDUs and reflected the fully adjusted SDU sample weight. The remaining components branched to reflect QDU and pair selection probabilities, as well as additional adjustments for ev, nr, and ps. Note that the final QDU and pair weights for the 2014 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 2010–2014 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 2002–2014 survey pair data, the built-in ev control feature was used for each adjustment step.

Exhibit 6.1 Summary of 2014 NSDUH QDU Sample Weight Components

Phase I Screener Dwelling Unit Level

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

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

Phase II Questionnaire Dwelling Unit Level

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

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

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

* These adjustments use the generalized exponential model (GEM), which also involves pre- and post-processing in addition to running the GEM macro. See [Exhibit 4.1](#) in the NSDUH Methodological Resource Book person-level sampling weight calibration report (Center for Behavioral Health Statistics and Quality, 2016b). 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 2014 NSDUH Person Pair Sample Weight Components

Phase I Screener Dwelling Unit Level

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

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

Phase II Person Pair Level

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

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

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

* These adjustments use the generalized exponential model (GEM), which also involves pre- and post-processing in addition to running the GEM macro. See [Exhibit 4.1](#) in the NSDUH Methodological Resource Book person-level sampling weight calibration report (Center for Behavioral Health Statistics and Quality, 2016b). 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	<p>Northeast (9 States) Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont</p>
2	<p>Midwest (12 States) Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, Wisconsin</p>
3	<p>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</p>
4	<p>West (13 States) Alaska, Arizona, California, Colorado, Idaho, Hawaii, Montana, Nevada, New Mexico, Oregon, Utah, Washington, Wyoming</p>
Pair	
1	<p>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</p>
2	<p>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</p>

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

Model Group	2010		2011		2012		2013		2014	
	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs	Selected QDUs	Completed QDUs
QDU										
Northeast	11,627	9,339	11,997	9,456	12,616	9,917	12,791	9,954	12,950	9,664
South	17,880	14,857	19,690	16,487	18,345	15,019	18,766	15,073	21,448	16,680
Midwest	16,670	13,686	17,045	13,752	16,984	13,687	17,207	13,519	15,276	11,618
West	12,525	10,231	12,709	10,438	12,676	10,227	12,870	10,350	15,122	11,710
Total	58,702	48,113	61,441	50,133	60,621	48,850	61,634	48,896	64,796	49,672
Model Group	2010		2011		2012		2013		2014	
	Selected Pairs	Completed Pairs	Selected Pairs	Completed Pairs	Selected Pairs	Completed Pairs	Selected Pairs	Completed Pairs	Selected Pairs	Completed Pairs
Pair										
Northeast + South	12,872	9,590	13,686	10,127	13,619	9,723	13,535	9,416	13,969	9,436
Midwest + West	13,423	10,101	13,409	9,849	13,416	9,736	13,573	9,526	12,875	8,793
Total	26,295	19,691	27,095	19,976	27,035	19,459	27,108	18,942	26,844	18,229

QDU = questionnaire dwelling unit.

Note: The 2010 sample sizes reflect the removal of falsified cases found in Pennsylvania and Maryland. The 2011–2014 samples were not affected. For additional information, see Section B.3.5 in Appendix B of the 2011, 2012, and 2013 national findings reports, and Section B.3.5 of the 2014 methodological summary and definitions (Center for Behavioral Health Statistics and Quality, 2012, 2013, 2014, 2015b).

6.1 Phase I SDU-Level Weight Components

A total of 11 weight components for the SDU level correspond to selection probabilities and nr, ps, and ev adjustment factors. Note that this differs from previous National Household Surveys on Drug Abuse and NSDUHs in that beginning in 2014, a new design-based component was incorporated at the beginning of the process so that corresponding weight component numbers are incremented by one when compared to previous survey years with an otherwise similar weighting scheme. The first eight components in the Phase I sample weights reflect the probability of selecting the DUs. These components were derived from (1) the probability of selecting the census tract and (2) census block within each SS region, (3) the probability of selecting the geographic segment within each SS region, (4) a quarter segment weight adjustment, (5) a subsegmentation inflation factor, (6) the probability of selecting a DU from within each counted and listed sampled segment, (7) the probability of inclusion of added DUs, and (8) DU percent release adjustment. The three remaining weight components, #9 through #11, are GEM calibration adjustments accounting for (9) DU nonresponse at the screening level, (10) DU poststratification to census controls, and (11) DU-level ev adjustment, although in 2014, ev adjustment at this stage was deemed unnecessary, and thus Weight Component #11 was set to one for all respondent DUs. The person-level, QDU-level, and person pair-level weights use the product of the above 11 weight components as the common initial weight before further adjustments. For more detailed information on Weight Components #1 through #3 and #5 through #8, refer to the 2014 NSDUH MRB sample design report (CBHSQ, 2015a), and for more detail on Weight Components #4 and #9 through #11, see the 2014 NSDUH MRB person-level sampling weight calibration report (CBHSQ, 2016b).

Note that from 2008 to 2010, there was an occasional second subsegmentation step when the initial partitioning of segments was insufficient because of out-of-date census counts or the segment was still too large to list after the original subsegmentation. This second partitioning was not accounted for in the weighting over these survey years. A comparison was done to evaluate the effect of this omission, and it was determined that the missing second subsegmenting factor in the analysis weight had minimal impact on estimates. Therefore, weights for these years were not re-created with a correcting factor. Additional detail can be found in CBHSQ (2015a).

New for the 2014 process was the inclusion of Weight Component #2, which is an adjustment to provide for census block sampling. This step was added to allow for possible transitioning to an address-based sampling design in the future. Additional changes to sample allocation and survey design are discussed in detail in CBHSQ (2015a).

6.2 QDU Weight Components

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

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

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

This poststratification factor adjusts the weights for selected QDUs to the SDU-based control totals. The SDU-based control totals are obtained by using the calibrated SDU weights. This adjustment step provides more stable controls for the subsequent nonresponse adjustment (Weight Component #14). [Exhibit 4.1](#) lists the initially proposed variables for GEM modeling. The predictor variables are either 0/1 indicators or counting variables representing the number of people who fall into a given demographic domain. The counting variables are derived from the screener demographic information. It may be noted that during screening, the only required demographic information was the age of each person rostered. Thus, other demographic information necessary for weight calibration, such as race/ethnicity and gender, may be missing for certain rostered eligible people, and so imputation was done to replace these missing data. For more details on the imputation of screener demographic information, see CBHSQ (2016b).

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

6.2.3 QDU Weight Component #14: Respondent QDU Nonresponse Adjustment

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

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

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

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

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

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

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

6.3 Pair-Level Weight Components

[Exhibit 4.2](#) lists the initially proposed predictor variables for the following adjustment steps via GEM.

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

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

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

Similar to QDU Weight Component #13, this step was motivated by the consideration that the larger sample of all possible pairs provides more stable control totals for the respondent pair nonresponse adjustment. The weights of selected pairs were poststratified to the control totals that derived from calibrated SDU weights of all possible pairs. The pair-level demographic

variables for all selected pairs, such as pair age group, pair race/ethnicity, etc., were derived from screener demographic information.

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

6.3.3 Pair Weight Component #14: Respondent Pair Nonresponse Adjustment

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

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

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

This final poststratification utilized the same set of initially proposed predictor variables as previous adjustment steps. In addition, 10 pair relationship domain-level indicator variables were added to the set of covariates. The control totals for GEM calibration were derived from the SDU sample of all possible pairs of eligible people, as was done for Weight Component #13. The calibration control totals for these 10 domains used household-level person counts and the final QDU weights. As mentioned in the introduction, use of these household-level count totals for pair relationship domains in GEM calibration provided Hajek-type weight adjustment in the interest of obtaining more stable estimates. In setting up calibration covariates, multiplicity factors were needed. These factors, as discussed in the introduction, are used in constructing estimates for person-level parameters based on pair-related drug behavior. The factors depend on the pair domains of interest. For a selected set of pair domains, multiplicity factors are provided along with the pair-level analysis weights. See Chapter 11 in the NSDUH MRB editing and imputation report (CBHSQ, 2016a) for more detail on the creation of and imputation of missing values in the pair relationship, multiplicity, and household-level person counts. See Chapter 4 for more detail on the use of multiplicities and household-level person counts in poststratification.

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

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

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

We checked the extreme weight proportions for the weights up to Weight Component #15, using the extreme weight domains (see Section 5.2). 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 #16. 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 2014 National Survey on Drug Use and Health for various national domains. This table also shows the weighted response rates. Most domains reflect the overall 73.86 percent response rate, with most rates relatively close to 80 percent, although the highest response rate is 93.67 percent, from the Household Type Age 12-17, 18-25 category. The lowest response rate came from fifth quintile of segment-combined median rent and housing value, with 69.68 percent.

Table I.1 in Appendix I displays the final selected and responding pair-level sample sizes from the 2014 survey for various national domains. Because of the nature of the pair data, the response rates were lower in all domains examined than at the QDU level, with an overall response rate of 61.34 percent. The response rates range from a low of 44.99 percent in the Pair Race/Ethnicity Other category to a high of 78.91 percent from the Pair Age 12-17, 12-17 category. This extreme range of response rates is probably due to a combination of small sample sizes and response burden as a result of selection of pairs within households among various domains. Like at the QDU level, the top response rates are among the younger respondents (as measured by household type for the QDU data and pair age for the pair data). This pattern may be related to the relatively high response rates in the group level of the variable group quarters because it includes college dormitories.

7.2 Proportions of Extreme Values and Outwinsors

During the stages of modeling adjustments (i.e., nonresponse [nr] and poststratification [ps]), one major issue of concern when deciding the adequacy of a particular model was the extent of the resulting proportions of extreme value (ev) and outwinsor weights (see Sections 5.1 and 5.2 for these definitions). For each weight adjustment step, these proportions are computed before and after the step for various domains. Prior to adjustment, the product of all weight components is used to compute proportions of evs and outwinsors, while after the adjustment, the product includes the new adjustment factor. If the proportion of evs and outwinsors is deemed high (normally 3% of unweighted, 15% weighted, and 5% of outwinsor), a separate ev treatment step after ps could be performed. Although this threshold was not met in the 2014 data, this step has been implemented for pair-level weighting to reduce final ev and outwinsor proportions and to maintain year-to-year consistency. This was done for the pair-level weights. Details of this step are explained in Section 6.3.5. A separate ev treatment step was deemed unnecessary for the QDU-level weights.

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

7.3 Slippage Rates

The slippage rate for a given domain is defined as the relative percentage difference between the sampling weights and the external control totals, both before and after ps. The control totals for QDU and person pair ps are derived from the screener dwelling unit weights, which were poststratified to U.S. Census Bureau population estimates (Center for Behavioral Health Statistics and Quality, 2016b). 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 2014 QDU domain household size of two, indicates a sample size of 20,354 with a total design-based weight of 55,444,416 and a census total of 55,458,172 with an initial slippage rate of -0.45 percent, which would imply a common weight adjustment approximately equal to 1.004538, if this were the only calibration control. Similarly, looking at pair data in Table K.1, the pair domain category of pair age 12-17, 35-49 has a sample size of 3,940, a design-based weight of 29,326,152, and a census total of 29,000,052, showing an initial slippage of 1.12 percent. The resultant required adjustment would be approximately equal to 0.98888, if this were the only control. However, in the generalized exponential model (GEM), all controls are simultaneously satisfied under a complex algorithm that allows for different adjustment factors for different units.

7.4 Weight Adjustment Summary Statistics

Tables G.1, G.2, and L.1 through L.3 display summary statistics on the product of weight components before and after all stages of adjustment for the QDU and person pair, respectively. The summary statistics include sample size (n), minimum (min), maximum (max), median (med), 25th percentile (Q1), 75th percentile (Q3), and the unequal weighting effect (UWE). Note that in Tables L.2 and L.3 the sample size for pair age group, pair race/ethnicity, and pair gender are slightly different. This is because those variables were defined using screening demographic information in the nonresponse adjustment of respondent pairs, while in the poststratification of

respondent pairs, they were defined from questionnaire demographic information. Because UWE is directly affected by weight adjustment factors and extreme weights, these values—along with the percentage of extreme weights as noted in Section 7.2—were used as guidelines for determining model adequacy.

7.5 Sensitivity Analysis of Drug Use Estimates

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

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

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

Table 7.1 Estimates of Totals and SEs for Domains of Interest Based on QDU Sample: 2014

Domain	<i>n</i>	Baseline (B)¹	Final (F)²	(B-F)/F% (Estimate)	(B-F)/F% (SE)
<i>Households with Family Income</i>					
\$0 - < \$10,000	4,396	9,684,167 (265,404)	9,681,809 (265,544)	0.02	-0.05
\$10,000 - < \$20,000	6,194	16,553,233 (351,729)	16,546,433 (351,767)	0.04	-0.01
\$20,000 - < \$30,000	5,456	13,890,744 (332,237)	13,889,186 (332,396)	0.01	-0.05
\$30,000 - < \$40,000	5,187	12,734,419 (268,572)	12,742,644 (268,945)	-0.06	-0.14
\$40,000 - < \$50,000	5,043	12,526,131 (260,058)	12,520,849 (260,010)	0.04	0.02
\$50,000 - < \$75,000	7,995	19,547,028 (339,117)	19,550,015 (339,622)	-0.02	-0.15
\$75,000 - < \$100,000	5,578	13,161,048 (274,740)	13,159,600 (275,059)	0.01	-0.12
\$100,000+	9,823	23,686,959 (480,515)	23,693,191 (482,080)	-0.03	-0.32
<i>Households with Number of Youths (< 18)</i>					
0	23,323	79,790,076 (865,479)	79,769,737 (865,023)	0.03	0.05
1	10,627	17,932,907 (245,316)	17,946,161 (246,272)	-0.07	-0.39
2	9,343	15,238,092 (227,319)	15,246,460 (227,855)	-0.05	-0.23
3	4,240	6,163,244 (126,576)	6,163,269 (126,738)	-0.00	-0.13
4+	2,139	2,659,410 (71,978)	2,658,100 (72,183)	0.05	-0.29
<i>Households with Insurance Coverage</i>					
Yes	43,487	108,195,696 (942,443)	108,202,113 (942,713)	-0.01	-0.03
No	6,185	13,588,031 (278,042)	13,581,615 (277,858)	0.05	0.07
<i>Households with Number of Older Adults (65+)</i>					
0	42,670	89,667,769 (761,889)	89,665,111 (761,724)	0.00	0.02
1	4,525	20,974,869 (469,646)	20,975,466 (469,701)	-0.00	-0.01
2	2,420	10,968,373 (320,255)	10,969,987 (320,645)	-0.01	-0.12
3+	57	172,717 (29,009)	173,164 (29,103)	-0.26	-0.32

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: 2014

Drug	Mother User	<i>n</i>	Baseline¹	Final²
Alcohol				
Lifetime	Yes	2,781	31.23 (1.49)	31.48 (1.67)
	No	385	14.15 (2.80)	14.45 (2.72)
	Overall	3,166	29.10 (1.35)	29.44 (1.50)
Past Year	Yes	2,303	28.46 (1.62)	28.82 (1.85)
	No	863	13.99 (1.81)	14.07 (1.76)
	Overall	3,166	24.52 (1.29)	24.92 (1.46)
Past Month	Yes	1,745	14.24 (1.38)	13.99 (1.39)
	No	1,421	6.03 (0.95)	5.76 (0.90)
	Overall	3,166	10.61 (0.91)	10.32 (0.90)
Cigarettes				
Lifetime	Yes	2,029	16.28 (1.35)	16.10 (1.33)
	No	1,137	5.99 (1.17)	6.02 (1.23)
	Overall	3,166	12.44 (0.96)	12.33 (0.97)
Past Year	Yes	894	18.49 (2.27)	18.10 (2.22)
	No	2,272	5.23 (0.66)	5.30 (0.69)
	Overall	3,166	8.57 (0.81)	8.52 (0.81)
Past Month	Yes	785	11.36 (2.06)	10.77 (1.94)
	No	2,381	2.98 (0.46)	3.07 (0.49)
	Overall	3,166	4.87 (0.60)	4.81 (0.59)

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

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

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

Table 7.2b 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: 2014

Drug	Father User	n	Baseline¹	Final²
Alcohol				
Lifetime	Yes	1,872	25.57 (1.69)	26.27 (1.78)
	No	140	12.08 (3.96)	11.89 (3.80)
	Overall	2,012	24.52 (1.57)	25.16 (1.66)
Past Year	Yes	1,549	21.08 (1.64)	21.97 (1.79)
	No	463	15.35 (2.79)	15.56 (2.82)
	Overall	2,012	19.69 (1.41)	20.43 (1.52)
Past Month	Yes	1,293	12.35 (1.38)	12.59 (1.41)
	No	719	7.35 (1.49)	7.42 (1.50)
	Overall	2,012	10.54 (1.05)	10.72 (1.07)
Cigarettes				
Lifetime	Yes	1,466	13.41 (1.36)	13.46 (1.37)
	No	546	5.14 (1.22)	5.28 (1.23)
	Overall	2,012	11.02 (1.04)	11.13 (1.06)
Past Year	Yes	532	13.18 (2.24)	13.17 (2.22)
	No	1,480	5.69 (0.91)	5.73 (0.93)
	Overall	2,012	7.42 (0.88)	7.42 (0.89)
Past Month	Yes	462	8.18 (1.83)	7.99 (1.74)
	No	1,550	3.40 (0.73)	3.49 (0.76)
	Overall	2,012	4.37 (0.69)	4.39 (0.70)

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: 2014

Drug	Mother User	<i>n</i>	Baseline¹	Final²
Any Illicit				
Lifetime	Yes	1,593	30.35 (1.98)	29.99 (1.98)
	No	1,573	16.64 (1.61)	17.69 (2.17)
	Overall	3,166	23.48 (1.29)	23.84 (1.46)
Past Year	Yes	346	27.48 (3.45)	27.29 (3.44)
	No	2,820	15.90 (1.16)	16.52 (1.41)
	Overall	3,166	16.96 (1.11)	17.51 (1.33)
Past Month	Yes	202	14.14 (3.51)	13.25 (3.33)
	No	2,964	8.33 (0.82)	8.34 (0.83)
	Overall	3,166	8.65 (0.80)	8.61 (0.80)
Marijuana				
Lifetime	Yes	1,416	24.60 (2.08)	24.60 (2.09)
	No	1,750	11.79 (1.47)	12.99 (2.01)
	Overall	3,166	17.53 (1.23)	18.19 (1.43)
Past Year	Yes	245	25.31 (4.31)	24.98 (4.28)
	No	2,921	12.57 (1.06)	13.27 (1.34)
	Overall	3,166	13.41 (1.04)	14.04 (1.28)
Past Month	Yes	153	15.17 (4.30)	14.26 (4.08)
	No	3,013	6.40 (0.76)	6.47 (0.78)
	Overall	3,166	6.78 (0.75)	6.80 (0.76)

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: 2014

Drug	Father User	n	Baseline¹	Final²
Any Illicit				
Lifetime	Yes	1,182	21.58 (1.92)	22.34 (2.06)
	No	830	16.38 (1.99)	16.56 (2.02)
	Overall	2,012	19.41 (1.41)	19.95 (1.48)
Past Year	Yes	264	22.47 (3.97)	23.30 (4.21)
	No	1,748	12.36 (1.29)	12.92 (1.41)
	Overall	2,012	13.54 (1.22)	14.15 (1.33)
Past Month	Yes	175	15.88 (4.36)	17.98 (5.18)
	No	1,837	7.72 (1.11)	7.76 (1.10)
	Overall	2,012	8.31 (1.07)	8.50 (1.09)
Marijuana				
Lifetime	Yes	1,082	15.52 (1.73)	16.47 (1.96)
	No	930	10.25 (1.47)	10.47 (1.49)
	Overall	2,012	13.05 (1.16)	13.69 (1.28)
Past Year	Yes	212	20.81 (4.51)	20.38 (4.40)
	No	1,800	9.00 (1.07)	9.64 (1.24)
	Overall	2,012	10.05 (1.03)	10.59 (1.16)
Past Month	Yes	142	14.82 (5.06)	14.93 (5.10)
	No	1,870	5.81 (0.91)	5.96 (0.93)
	Overall	2,012	6.31 (0.89)	6.45 (0.91)

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: 2014

Drug	Parent Talked about Dangers with Child	<i>n</i>	Baseline¹	Final²
Alcohol				
Lifetime	Yes	2,888	26.41 (1.46)	26.91 (1.58)
	No	2,210	29.41 (1.56)	29.63 (1.57)
	Overall	5,098	27.74 (1.09)	28.11 (1.15)
Past Year	Yes	2,888	21.70 (1.37)	22.28 (1.49)
	No	2,210	24.38 (1.45)	24.61 (1.47)
	Overall	5,098	22.89 (1.03)	23.31 (1.09)
Past Month	Yes	2,888	10.44 (0.99)	10.28 (1.01)
	No	2,210	10.88 (1.16)	10.84 (1.17)
	Overall	5,098	10.64 (0.77)	10.53 (0.78)
Cigarettes				
Lifetime	Yes	2,888	11.98 (1.02)	11.95 (1.04)
	No	2,210	12.35 (1.08)	12.39 (1.08)
	Overall	5,098	12.15 (0.78)	12.14 (0.79)
Past Year	Yes	2,888	8.85 (0.95)	8.79 (0.96)
	No	2,210	7.48 (0.88)	7.44 (0.86)
	Overall	5,098	8.24 (0.67)	8.19 (0.67)
Past Month	Yes	2,888	5.72 (0.82)	5.65 (0.83)
	No	2,210	4.01 (0.71)	4.03 (0.70)
	Overall	5,098	4.96 (0.55)	4.94 (0.56)

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: 2014

Drug	Parent Talked about Dangers with Child	<i>n</i>	Baseline¹	Final²
Any Illicit				
Lifetime	Yes	2,888	22.21 (1.43)	22.81 (1.57)
	No	2,210	23.65 (1.57)	23.74 (1.59)
	Overall	5,098	22.85 (1.12)	23.22 (1.18)
Past Year	Yes	2,888	16.75 (1.30)	17.54 (1.46)
	No	2,210	15.03 (1.13)	15.07 (1.14)
	Overall	5,098	15.98 (0.91)	16.45 (0.99)
Past Month	Yes	2,888	9.74 (1.07)	9.81 (1.07)
	No	2,210	7.61 (0.92)	7.59 (0.91)
	Overall	5,098	8.80 (0.73)	8.83 (0.73)
Marijuana				
Lifetime	Yes	2,888	16.50 (1.34)	17.31 (1.51)
	No	2,210	16.93 (1.48)	17.28 (1.51)
	Overall	5,098	16.69 (1.04)	17.30 (1.12)
Past Year	Yes	2,888	13.63 (1.22)	14.39 (1.40)
	No	2,210	11.48 (1.06)	11.60 (1.07)
	Overall	5,098	12.67 (0.85)	13.16 (0.93)
Past Month	Yes	2,888	7.45 (0.98)	7.55 (0.98)
	No	2,210	6.27 (0.88)	6.26 (0.88)
	Overall	5,098	6.93 (0.68)	6.98 (0.68)

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: 2014

Drug	Mother Talked about Dangers with Child	<i>n</i>	Baseline¹	Final²
Alcohol				
Lifetime	0 times	239	20.43 (4.28)	21.12 (4.14)
	1-2 times	472	20.21 (2.89)	19.92 (3.07)
	A few times	862	29.76 (2.66)	30.18 (2.74)
	Many times	1,468	32.58 (2.01)	33.10 (2.33)
	Overall	3,041	28.88 (1.38)	29.25 (1.55)
Past Year	0 times	239	16.16 (4.12)	16.37 (3.90)
	1-2 times	472	17.84 (2.79)	17.76 (2.97)
	A few times	862	23.03 (2.33)	23.24 (2.37)
	Many times	1,468	28.67 (1.95)	29.42 (2.32)
	Overall	3,041	24.42 (1.33)	24.86 (1.50)
Past Month	0 times	239	4.59 (1.82)	5.10 (2.06)
	1-2 times	472	8.18 (1.90)	7.89 (1.92)
	A few times	862	9.22 (1.69)	9.36 (1.70)
	Many times	1,468	12.55 (1.42)	11.96 (1.39)
	Overall	3,041	10.31 (0.92)	10.07 (0.90)
Cigarettes				
Lifetime	0 times	239	8.48 (2.61)	9.02 (2.68)
	1-2 times	472	8.58 (2.32)	8.60 (2.42)
	A few times	862	12.57 (1.89)	12.68 (1.86)
	Many times	1,468	14.22 (1.41)	13.80 (1.41)
	Overall	3,041	12.42 (0.98)	12.28 (0.99)
Past Year	0 times	239	6.19 (2.46)	6.58 (2.52)
	1-2 times	472	6.69 (2.25)	6.74 (2.34)
	A few times	862	8.31 (1.57)	8.54 (1.58)
	Many times	1,468	9.66 (1.15)	9.32 (1.13)
	Overall	3,041	8.54 (0.83)	8.48 (0.83)
Past Month	0 times	239	2.24 (0.97)	2.44 (1.07)
	1-2 times	472	5.22 (1.92)	5.38 (2.01)
	A few times	862	5.03 (1.42)	5.14 (1.42)
	Many times	1,468	5.10 (0.76)	4.83 (0.69)
	Overall	3,041	4.87 (0.62)	4.82 (0.61)

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: 2014

Drug	Father Talked about Dangers with Child	<i>n</i>	Baseline¹	Final²
Alcohol				
Lifetime	0 times	226	21.31 (4.06)	21.08 (4.10)
	1-2 times	405	17.70 (2.71)	18.93 (3.39)
	A few times	601	26.46 (2.84)	27.12 (2.90)
	Many times	652	26.92 (2.86)	27.42 (2.89)
	Overall	1,884	24.00 (1.62)	24.63 (1.71)
Past Year	0 times	226	12.40 (2.80)	12.21 (2.73)
	1-2 times	405	13.77 (2.58)	15.37 (3.34)
	A few times	601	22.49 (2.69)	23.22 (2.77)
	Many times	652	21.86 (2.60)	22.30 (2.65)
	Overall	1,884	19.14 (1.45)	19.88 (1.57)
Past Month	0 times	226	4.13 (1.57)	4.13 (1.61)
	1-2 times	405	5.77 (1.53)	5.76 (1.53)
	A few times	601	9.92 (1.86)	10.10 (1.91)
	Many times	652	15.32 (2.36)	15.65 (2.41)
	Overall	1,884	10.24 (1.08)	10.40 (1.10)
Cigarettes				
Lifetime	0 times	226	5.77 (1.93)	5.84 (1.96)
	1-2 times	405	4.70 (1.06)	4.58 (1.02)
	A few times	601	10.84 (2.05)	11.20 (2.11)
	Many times	652	15.48 (2.19)	15.56 (2.20)
	Overall	1,884	10.49 (1.04)	10.60 (1.06)
Past Year	0 times	226	3.99 (1.60)	3.98 (1.60)
	1-2 times	405	3.00 (0.95)	2.88 (0.90)
	A few times	601	6.68 (1.58)	6.74 (1.61)
	Many times	652	11.64 (2.02)	11.67 (2.03)
	Overall	1,884	7.28 (0.90)	7.27 (0.91)
Past Month	0 times	226	1.63 (1.02)	1.63 (1.01)
	1-2 times	405	2.55 (0.91)	2.43 (0.86)
	A few times	601	3.47 (1.12)	3.53 (1.16)
	Many times	652	7.11 (1.66)	7.22 (1.69)
	Overall	1,884	4.35 (0.72)	4.37 (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 7.7a 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: 2014

Drug	Mother Talked about Dangers with Child	<i>n</i>	Baseline¹	Final²
Any Illicit				
Lifetime	0 times	239	16.05 (3.35)	16.87 (3.46)
	1-2 times	472	18.54 (3.12)	18.36 (3.25)
	A few times	862	21.93 (2.34)	22.15 (2.39)
	Many times	1,468	26.69 (2.01)	27.19 (2.37)
	Overall	3,041	23.25 (1.32)	23.60 (1.49)
Past Year	0 times	239	11.13 (3.03)	11.87 (3.16)
	1-2 times	472	12.09 (2.32)	12.00 (2.43)
	A few times	862	15.99 (2.08)	16.00 (2.07)
	Many times	1,468	19.28 (1.73)	20.27 (2.24)
	Overall	3,041	16.60 (1.13)	17.13 (1.35)
Past Month	0 times	239	4.55 (1.65)	4.92 (1.80)
	1-2 times	472	6.67 (2.01)	7.00 (2.12)
	A few times	862	8.84 (1.66)	8.98 (1.67)
	Many times	1,468	9.06 (1.10)	8.70 (1.08)
	Overall	3,041	8.26 (0.80)	8.20 (0.80)
Marijuana				
Lifetime	0 times	239	8.96 (2.66)	9.59 (2.76)
	1-2 times	472	11.73 (2.82)	11.74 (2.91)
	A few times	862	17.47 (2.36)	17.74 (2.42)
	Many times	1,468	20.49 (1.88)	21.57 (2.34)
	Overall	3,041	17.35 (1.26)	18.02 (1.47)
Past Year	0 times	239	7.16 (2.52)	7.55 (2.59)
	1-2 times	472	8.35 (2.15)	8.42 (2.24)
	A few times	862	13.34 (1.96)	13.50 (1.97)
	Many times	1,468	15.67 (1.62)	16.80 (2.20)
	Overall	3,041	13.20 (1.05)	13.84 (1.31)
Past Month	0 times	239	1.54 (0.60)	1.65 (0.67)
	1-2 times	472	4.50 (1.91)	4.68 (1.99)
	A few times	862	7.74 (1.64)	7.86 (1.64)
	Many times	1,468	7.04 (0.98)	6.94 (1.00)
	Overall	3,041	6.38 (0.75)	6.41 (0.76)

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: 2014

Drug	Father Talked about Dangers with Child	<i>n</i>	Baseline¹	Final²
Any Illicit				
Lifetime	0 times	226	15.04 (3.74)	15.33 (3.87)
	1-2 times	405	16.59 (3.16)	17.65 (3.71)
	A few times	601	15.87 (2.35)	16.69 (2.48)
	Many times	652	23.72 (2.56)	23.65 (2.58)
	Overall	1,884	18.74 (1.45)	19.24 (1.54)
Past Year	0 times	226	7.35 (2.63)	7.06 (2.48)
	1-2 times	405	11.35 (2.66)	12.75 (3.39)
	A few times	601	11.56 (2.20)	12.18 (2.32)
	Many times	652	17.03 (2.21)	17.28 (2.27)
	Overall	1,884	13.01 (1.26)	13.60 (1.38)
Past Month	0 times	226	2.00 (0.87)	2.02 (0.88)
	1-2 times	405	6.37 (2.19)	6.17 (2.09)
	A few times	601	6.95 (1.90)	7.58 (2.03)
	Many times	652	11.48 (2.01)	11.68 (2.06)
	Overall	1,884	7.90 (1.08)	8.12 (1.10)
Marijuana				
Lifetime	0 times	226	6.78 (2.21)	6.49 (2.06)
	1-2 times	405	10.64 (2.38)	12.20 (3.24)
	A few times	601	11.56 (2.15)	12.30 (2.29)
	Many times	652	16.09 (2.17)	16.24 (2.20)
	Overall	1,884	12.45 (1.17)	13.07 (1.30)
Past Year	0 times	226	4.05 (1.75)	3.76 (1.56)
	1-2 times	405	8.05 (1.92)	9.62 (2.98)
	A few times	601	8.75 (1.99)	9.23 (2.10)
	Many times	652	13.00 (1.95)	13.10 (1.99)
	Overall	1,884	9.60 (1.04)	10.13 (1.19)
Past Month	0 times	226	1.44 (0.67)	1.46 (0.67)
	1-2 times	405	3.49 (1.06)	3.45 (1.04)
	A few times	601	5.42 (1.74)	5.91 (1.85)
	Many times	652	9.35 (1.84)	9.41 (1.87)
	Overall	1,884	5.94 (0.89)	6.10 (0.91)

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.

Center for Behavioral Health Statistics and Quality. (2013). *Results from the 2012 National Survey on Drug Use and Health: Summary of national findings* (HHS Publication No. SMA 13-4795, NSDUH Series H-46). Rockville, MD: Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (2014). *Results from the 2013 National Survey on Drug Use and Health: Summary of national findings* (HHS Publication No. SMA 14-4863, NSDUH Series H-48). Rockville, MD: Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (2015a). *2014 National Survey on Drug Use and Health: Methodological resource book (Section 2, Sample design report)*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (2015b). *Results from the 2014 National Survey on Drug Use and Health: Methodological summary and definitions*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (2016a). *2014 National Survey on Drug Use and Health: Methodological resource book (Section 10, Editing and imputation report)*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (2016b). *2014 National Survey on Drug Use and Health: Methodological resource book (Section 11, Person-level sampling weight calibration)*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

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

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

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

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

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

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

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

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

Laufenberg, J., Kroutil, L., Frechtel, P., Carpenter, L., Edwards, S., Ewing, G., et al. (2014). Editing and imputation report. In *2012 National Survey on Drug Use and Health: Methodological resource book* (Sections 10/11, prepared for the Substance Abuse and Mental Health Services Administration, Contract No. HHSS283201000003C, Deliverable No. 41, RTI/0212800.001.107.006.007). 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. Retrieved from <http://www.amstat.org/sections/srms/proceedings/>

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

Appendix A: Technical Details about the Generalized Exponential Model

This page intentionally left blank

Appendix A: Technical Details about the Generalized Exponential Model

A.1 Distance Function

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

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

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, \dots, x_p) . Then, the calibration constraints for the above minimization problem are

$$\sum_{k \in s} x_k d_k a_k = T_x. \quad (\text{A.1.2})$$

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

$$a_k(\lambda) = \frac{\ell_k (u_k - c_k) + u_k (c_k - \ell_k) \exp\{A_k x'_k \lambda\}}{(u_k - c_k) + (c_k - \ell_k) \exp\{A_k x'_k \lambda\}}. \quad (\text{A.1.3})$$

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

$$\ell_k < a_k < u_k, \quad k = 1, \dots, n. \quad (\text{A.1.4})$$

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

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

and $a_k(\lambda) = \exp(x'_k \lambda)$.

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

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

By choosing the user-specified parameters ℓ_k , c_k , and u_k appropriately, the unified GEM formula (A.1.3) can be justified for all three types of adjustment: extreme value treatment, nonresponse, and poststratification. For extreme value treatment via winsorization, denote the winsorized weights by $\{b_k\}$, where $b_k = d_k$ if d_k is not an extreme weight, and

$b_k = \text{med}\{d_k\} \pm 3 * \text{IQR}$ if d_k is an extreme weight, where IQR denotes the interquartile range, and the median and quartiles for the weights are defined with respect to a suitable design-based stratum.

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

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

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

A.3 Newton-Raphson Steps

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

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

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

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

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

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

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

A.4 Scaled Constrained Exponential Model

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

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

This page intentionally left blank

Appendix B: Derivation of Poststratification Control Totals

This page intentionally left blank

Appendix B: Derivation of Poststratification Control Totals

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

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

B.1 Derivation of QDU-Level Poststratification Controls

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

B.1.1 Person Level

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

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

The breakdown received from the census did not match the levels of the domains that we wanted to control. To account for this, we collapsed levels. From this altered data, we created datasets with model group-specific control totals. Observations in these datasets corresponded to a breakdown by quarter (4), Hispanicity (2), race (5), gender (2), age (11), and number of states¹ in the model group (number of states varied according to which census region was represented in the model group).

B.1.1.2 Adjusting SDU Data to the Control Totals

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

B.1.2 QDU Level

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

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

¹ The District of Columbia is included among states.

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

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

B.2 Derivation of Person Pair-Level Poststratification Controls

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

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

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

On the other hand, for the 10 selected pair domains based on relationship, the control totals for the poststratification adjustments were obtained from the questionnaire roster. This involved calibrating the pair weights to the number of people in households belonging to each domain of interest. These controls were obtained from the larger sample of singles and pairs (i.e., one or two people selected from dwelling units) and were calculated at the QDU (household) level. The pair weights were adjusted by the appropriate multiplicity. See Chapter 11 in the NSDUH Methodological Resource Book editing and imputation report (Center for Behavioral Health Statistics and Quality, 2016a) for details on the multiplicity counts and household-level control totals, which are referred to as household-level person counts.

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

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

This page intentionally left blank

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

This page intentionally left blank

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

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

For 2014, 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

<p>Age^b 1: 12-17, 2: 18-25, 3: 26-34, 4: 35-49, 5: 50+¹</p> <p>Gender^b 1: Male, 2: Female¹</p> <p>Group Quarter Indicator^a 1: College Dorm, 2: Other Group Quarter, 3: Non-Group Quarter¹</p> <p>Hispanicity^b 1: Hispanic or Latino, 2: Non-Hispanic or Latino¹</p> <p>Household Size^b Continuous Variable Count of Individuals Rostered with DU</p> <p>Household Type (Ages of People 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</p> <p>Percentage of Owner-Occupied Dwelling Units in Segment (% Owner-Occupied)^a 1: 50-100%,¹ 2: 10-<50%, 3: 0-<10%</p> <p>Percentage of Segments That Are Black or African American^a 1: 50-100%, 2: 10-<50%, 3: 0-<10%¹</p> <p>Percentage of Segments That Are Hispanic or Latino^a 1: 50-100%, 2: 10-<50%, 3: 0-<10%¹</p> <p>Population Density^a 1: MSA 1,000,000 or More, 2: MSA Less than 1,000,000, 3: Non-MSA Urban, 4: Non-MSA Rural¹</p> <p>Quarter^{a,b} 1: Quarter 1, 2: Quarter 2, 3: Quarter 3, 4: Quarter 4¹</p> <p>Race (3 Levels)^b 1: White¹, 2: Black or African American, 3: Other</p> <p>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</p> <p>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</p> <p>Relation to Householder^a 1: Householder or Spouse, 2: Child, 3: Other Relative, 4: Nonrelative¹</p> <p>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 Quintile¹</p> <p>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: Wisconsin¹ 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</p> <p>State/Region^{a,3} Model Group 1: 1: New York, 2: Pennsylvania, 3: Other¹ Model Group 2: 1: Illinois, 2: Michigan, 3: Ohio, 4: Other¹ Model Group 3: 1: Florida, 2: Texas, 3: Other¹ Model Group 4: 1: California, 2: Other¹</p>
--

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

State (for the model group in question, in this case, Model Group 9)

Model Group 9: 1: Alaska, 2: Hawaii, 3: Oregon, 4: Washington,¹ 5: California

Age

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

Race (3 Levels)

1: White,¹ 2: Black or African American, 3: Other

3. Construct the cross-classification table.

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

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

Shading indicates the reference-level set.

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

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

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

Shading indicates the reference-level set.

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

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

Shading indicates the reference-level set.

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

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

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

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

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

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

Shading indicates the reference-level set.

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

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

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

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

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

Shading indicates the reference-level set.

Exhibit C.2 Covariates for 2014 NSDUH Questionnaire Dwelling Unit Weights

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

This page intentionally left blank

Table C.1a 2014 QDU Weight GEM Modeling Summary (Model Group 1: Northeast)

Modeling Step ¹	Extreme Weight Proportions			UWE ²	# Covariates ³	Bounds ⁴	
	% Unweighted	% Weighted	% Outwisor			Nominal	Realized
<i>sel.qdu.ps</i>	1.89	4.29	1.21	2.2876	243	(0.38, 1.20)	(0.38, 1.20)
	1.73	3.71	0.54	2.2028	242	(0.54, 2.12)	(0.54, 2.12)
						(0.90, 2.41)	(0.90, 2.41)
<i>res.qdu.nr</i>	1.78	4.19	0.69	2.2620	243	(1.00, 2.30)	(1.00, 2.30)
	1.16	2.96	0.45	2.3959	241	(1.00, 3.01)	(1.00, 3.00)
						(1.30, 1.58)	(1.30, 1.58)
<i>res.qdu.ps</i>	1.16	2.96	0.45	2.3959	243	(0.97, 1.10)	(0.98, 1.10)
	1.26	3.22	0.20	2.3984	242	(0.65, 3.45)	(0.66, 3.44)
						(0.90, 1.01)	(0.90, 0.90)

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 2014 Distribution of Weight Adjustment Factors and Weight Products (Model Group 1: Northeast)

	SDU Weight	QDU Design Weight		sel.qdu.ps ¹		res.qdu.nr ¹		res.qdu.ps ¹	
	1-11	duwght12	1-12	duwght13	1-13	duwght14	1-14	duwght15	1-15
Minimum	33	1.00	33	0.29	44	0.68	50	0.57	46
1%	88	1.00	106	0.63	108	1.00	127	0.83	126
5%	124	1.00	177	0.77	175	1.05	210	0.96	209
10%	186	1.00	262	0.84	254	1.10	300	0.98	297
25%	303	1.00	610	0.92	582	1.19	677	0.99	667
Median	910	1.34	1,184	1.00	1,191	1.30	1,486	1.00	1,482
75%	1,126	2.11	1,931	1.09	1,957	1.46	2,622	1.01	2,636
90%	1,426	4.34	3,655	1.18	3,708	1.62	5,069	1.02	5,067
95%	1,712	6.41	5,079	1.29	5,163	1.74	7,587	1.03	7,634
99%	2,692	8.89	8,946	1.68	9,510	2.03	13,856	1.10	13,760
Maximum	8,241	10.57	53,116	2.66	21,217	3.00	30,966	3.44	30,852
<i>n</i>	12,950	-	12,950	-	12,950	-	9,664	-	9,664
Mean	854	2.06	1,680	1.02	1,695	1.34	2,271	1.00	2,271
Max/Mean	10	-	31.61	-	12.52	-	13.63	-	14

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

Out of 243 proposed effects, 242 were kept in the model, with the exception of combining the fourth quartile of segment-combined median rent and housing values of 10-50 and 50-100 levels of percentage of segments that are Black or African American.

Respondent Questionnaire Dwelling Unit-Level Nonresponse

Out of 243 proposed effects, 241 were kept in the model, with the exception of State \times Race, which combined American Indian or Alaska Native and Asian for Connecticut, and combining the fourth quartile of segment-combined median rent and housing values of 10-50 and 50-100 levels of percentage of segments that are Black or African American.

Respondent Questionnaire Dwelling Unit-Level Poststratification

This step used exactly the same variables as in the selected questionnaire dwelling unit-level poststratification step.

**Exhibit C.1.1 Covariates for 2014 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	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	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	7	Coll. (1,4) & (2,4); conv.
% Hispanicity × % Owner-Occupied	3 × 3	4	4	All levels present.
% Hispanicity × Rent/Housing	3 × 5	8	8	All levels present.
% Owner-Occupied × Rent/Housing	3 × 5	8	8	All levels present.
Three-Factor Effects		50	50	
Race (3 Levels) × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Hispanicity	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Race (3 Levels)	3 × 5 × 3	16	16	All levels present.
State/Region × Hispanicity × Gender	3 × 2 × 2	2	2	All levels present.
State/Region × Race (3 Levels) × Hispanicity	3 × 3 × 2	4	4	All levels present.
State/Region × Race (3 Levels) × Gender	3 × 3 × 2	4	4	All levels present.
Total		243	242	

**Exhibit C.1.2 Covariates for 2014 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	131	
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. (1,3) & (1,4); 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	7	Coll. (1,4) & (2,4); conv.
% Hispanicity × % Owner-Occupied	3 × 3	4	4	All levels present.
% Hispanicity × Rent/Housing	3 × 5	8	8	All levels present.
% Owner-Occupied × Rent/Housing	3 × 5	8	8	All levels present.
Three-Factor Effects		50	50	
Race (3 Levels) × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Hispanicity	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Race (3 Levels)	3 × 5 × 3	16	16	All levels present.
State/Region × Hispanicity × Gender	3 × 2 × 2	2	2	All levels present.
State/Region × Race (3 Levels) × Hispanicity	3 × 3 × 2	4	4	All levels present.
State/Region × Race (3 Levels) × Gender	3 × 3 × 2	4	4	All levels present.
Total		243	241	

**Exhibit C.1.3 Covariates for 2014 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps)
Model Group 1: Northeast**

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

Appendix C.2: Model Group 2: Midwest

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

This page intentionally left blank

Table C.2a 2014 QDU Weight GEM Modeling Summary (Model Group 2: Midwest)

Modeling Step ¹	Extreme Weight Proportions			UWE ²	# Covariates ³	Bounds ⁴	
	% Unweighted	% Weighted	% Outwisor			Nominal	Realized
<i>sel.qdu.ps</i>	1.04	1.72	0.40	1.9441	300	(0.56, 1.10)	(0.56, 1.10)
	0.83	1.11	0.17	1.9076	299	(0.53, 4.89)	(0.54, 4.89)
						(0.90, 1.15)	(0.90, 1.15)
<i>res.qdu.nr</i>	0.83	0.99	0.17	1.9440	300	(1.00, 1.70)	(1.00, 1.70)
	0.68	0.88	0.13	2.0348	299	(1.00, 4.18)	(1.00, 4.12)
						(1.30, 1.41)	(1.30, 1.41)
<i>res.qdu.ps</i>	0.68	0.88	0.13	2.0348	300	(0.99, 1.60)	(1.00, 1.60)
	0.73	0.91	0.10	2.0343	299	(0.96, 1.09)	(0.97, 1.08)
						(0.98, 1.02)	(0.98, 1.01)

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

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

² Unequal weighting effect (UWE) defined as $1 + [(n - 1)/n] * CV^2$, 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 2014 Distribution of Weight Adjustment Factors and Weight Products (Model Group 2: Midwest)

	SDU Weight	QDU Design Weight		sel.qdu.ps ¹		res.qdu.nr ¹		res.qdu.ps ¹	
	1-11	duwght12	1-12	duwght13	1-13	duwght14	1-14	duwght15	1-15
Minimum	33	1.00	33	0.28	47	0.38	61	0.66	63
1%	118	1.00	136	0.71	136	1.01	160	0.99	160
5%	153	1.00	227	0.82	229	1.06	255	1.00	256
10%	210	1.00	356	0.87	352	1.11	417	1.00	417
25%	634	1.00	814	0.93	797	1.19	964	1.00	964
Median	917	1.32	1,269	0.99	1,260	1.30	1,563	1.00	1,565
75%	1,169	2.22	2,076	1.05	2,071	1.42	2,684	1.00	2,683
90%	1,419	3.89	3,756	1.13	3,706	1.53	5,145	1.00	5,142
95%	1,606	6.13	5,289	1.20	5,182	1.61	7,331	1.01	7,329
99%	2,090	7.55	8,955	1.46	8,645	1.90	11,918	1.02	11,925
Maximum	7,734	8.40	45,450	4.89	15,588	4.12	18,709	1.23	18,739
<i>n</i>	15,276	-	15,276	-	15,276	-	11,618	-	11,618
Mean	904	4.23	1,771	1.00	1,762	1.32	2,316	1.00	2,316
Max/Mean	9	-	25.66	-	8.85	-	8.08	-	8

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

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

Model Group 2 Overview

Selected Questionnaire Dwelling Unit-Level Poststratification

Out of 300 proposed effects, 299 were kept in the model, with the exception of combining the fourth quartile of segment-combined median rent and housing values of 10-50 and 50-100 levels of percentage of segments that are Black or African American.

Respondent Questionnaire Dwelling Unit-Level Nonresponse

This step used exactly the same variables as in the selected questionnaire dwelling unit-level poststratification step.

Respondent Questionnaire Dwelling Unit-Level Poststratification

This step used exactly the same variables as in the selected questionnaire dwelling unit-level poststratification step.

**Exhibit C.2.1 Covariates for 2014 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	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	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	7	Coll. (1,4) & (2,4); conv.
% 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	299	

**Exhibit C.2.2 Covariates for 2014 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	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	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	7	Coll. (1,4) & (2,4); conv.
% Hispanicity × % Owner-Occupied	3 × 3	4	4	All levels present.
% Hispanicity × Rent/Housing	3 × 5	8	8	All levels present.
% Owner-Occupied × Rent/Housing	3 × 5	8	8	All levels present.
Three-Factor Effects		71	71	
Race (3 Levels) × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Gender	4 × 5 × 2	12	12	All levels present.
State/Region × Age × Hispanicity	4 × 5 × 2	12	12	All levels present.
State/Region × Age × Race (3 Levels)	4 × 5 × 3	24	24	All levels present.
State/Region × Hispanicity × Gender	4 × 2 × 2	3	3	All levels present.
State/Region × Race (3 Levels) × Hispanicity	4 × 3 × 2	6	6	All levels present.
State/Region × Race (3 Levels) × Gender	4 × 3 × 2	6	6	All levels present.
Total		300	299	

**Exhibit C.2.3 Covariates for 2014 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	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	7	Coll. (1,4) & (2,4); conv.
% Hispanicity × % Owner-Occupied	3 × 3	4	4	All levels present.
% Hispanicity × Rent/Housing	3 × 5	8	8	All levels present.
% Owner-Occupied × Rent/Housing	3 × 5	8	8	All levels present.
Three-Factor Effects		71	71	
Race (3 Levels) × Age × Gender	3 × 5 × 2	8	8	All levels present.
State/Region × Age × Gender	4 × 5 × 2	12	12	All levels present.
State/Region × Age × Hispanicity	4 × 5 × 2	12	12	All levels present.
State/Region × Age × Race (3 Levels)	4 × 5 × 3	24	24	All levels present.
State/Region × Hispanicity × Gender	4 × 2 × 2	3	3	All levels present.
State/Region × Race (3 Levels) × Hispanicity	4 × 3 × 2	6	6	All levels present.
State/Region × Race (3 Levels) × Gender	4 × 3 × 2	6	6	All levels present.
Total		300	299	

Appendix C.3: Model Group 3: South

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

This page intentionally left blank

Table C.3a 2014 QDU Weight GEM Modeling Summary (Model Group 3: South)

Modeling Step ¹	Extreme Weight Proportions			UWE ²	# Covariates ³	Bounds ⁴	
	% Unweighted	% Weighted	% Outwisor			Nominal	Realized
<i>sel.qdu.ps</i>	1.06	1.84	0.43	1.8383	339	(0.75, 2.87)	(0.76, 2.87)
	0.89	1.42	0.25	1.8447	339	(0.44, 2.58)	(0.46, 2.56)
						(0.90, 1.73)	(0.90, 1.73)
<i>res.qdu.nr</i>	0.88	1.59	0.27	1.8382	339	(1.00, 1.50)	(1.00, 1.50)
	0.66	1.07	0.14	2.0186	339	(1.00, 4.65)	(1.00, 4.63)
						(1.30, 1.49)	(1.30, 1.49)
<i>res.qdu.ps</i>	0.66	1.07	0.14	2.0186	339	(0.20, 1.10)	(0.96, 1.10)
	0.59	0.87	0.04	2.0173	339	(0.20, 5.00)	(0.71, 1.56)
						(0.90, 5.00)	(0.90, 1.13)

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 2014 Distribution of Weight Adjustment Factors and Weight Products (Model Group 3: South)

	SDU Weight	QDU Design Weight		sel.qdu.ps ¹		res.qdu.nr ¹		res.qdu.ps ¹	
	1-11	duwght12	1-12	duwght13	1-13	duwght14	1-14	duwght15	1-15
Minimum	14	1.00	14	0.46	11	0.54	12	0.50	12
1%	88	1.00	110	0.75	105	1.00	117	0.96	117
5%	142	1.00	262	0.85	267	1.05	326	0.99	325
10%	287	1.00	458	0.89	466	1.08	579	0.99	579
25%	766	1.00	975	0.94	970	1.16	1,139	1.00	1,141
Median	1,098	1.30	1,549	1.00	1,555	1.26	1,883	1.00	1,881
75%	1,467	2.23	2,565	1.07	2,581	1.39	3,235	1.00	3,223
90%	1,863	4.03	4,469	1.13	4,552	1.52	6,003	1.01	5,997
95%	2,100	5.18	6,071	1.18	6,045	1.61	8,550	1.01	8,544
99%	2,919	8.02	9,718	1.31	10,023	1.86	13,713	1.04	13,687
Maximum	7,583	9.83	29,230	2.81	22,788	4.63	32,038	1.56	32,049
<i>n</i>	21,448	-	21,448	-	21,448	-	16,680	-	16,680
Mean	1,126	1.99	2,120	1.01	2,136	1.29	2,747	1.00	2,747
Max/Mean	7	-	14	-	11	-	12	-	12

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

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

Model Group 3 Overview

Selected Questionnaire Dwelling Unit-Level Poststratification

All 339 proposed effects were kept in the model.

Respondent Questionnaire Dwelling Unit-Level Nonresponse

All 339 proposed effects were kept in the model.

Respondent Questionnaire Dwelling Unit-Level Poststratification

All 339 proposed effects were kept in the model.

**Exhibit C.3.1 Covariates for 2014 NSDUH Questionnaire Dwelling Unit Weights (sel.qdu.ps)
Model Group 3: South**

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

**Exhibit C.3.2 Covariates for 2014 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr)
Model Group 3: South**

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

**Exhibit C.3.3 Covariates for 2014 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.ps)
Model Group 3: South**

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

Appendix C.4: Model Group 4: West

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

This page intentionally left blank

Table C.4a 2014 QDU Weight GEM Modeling Summary (Model Group 4: West)

Modeling Step ¹	Extreme Weight Proportions			UWE ²	# Covariates ³	Bounds ⁴	
	% Unweighted	% Weighted	% Outwisor			Nominal	Realized
<i>sel.qdu.ps</i>	1.43	2.87	0.71	2.1595	270	(0.74, 1.10)	(0.74, 1.10)
	1.01	2.15	0.29	2.1254	266	(0.58, 3.70)	(0.58, 3.69)
						(0.93, 3.12)	(0.95, 3.12)
<i>res.qdu.nr</i>	1.08	2.32	0.36	2.1650	270	(1.00, 1.70)	(1.00, 1.70)
	0.86	1.38	0.14	2.3200	266	(1.00, 4.40)	(1.00, 4.35)
						(1.30, 5.00)	(1.30, 1.30)
<i>res.qdu.ps</i>	0.86	1.38	0.14	2.3200	270	(0.20, 1.20)	(0.96, 1.20)
	0.84	1.23	0.07	2.3192	266	(0.20, 5.00)	(0.57, 3.03)
						(0.90, 5.00)	(0.92, 1.00)

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

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

² Unequal weighting effect (UWE) defined as $1 + [(n - 1)/n] * CV^2$, 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 2014 Distribution of Weight Adjustment Factors and Weight Products (Model Group 4: West)

	SDU Weight	QDU Design Weight		sel.qdu.ps ¹		res.qdu.nr ¹		res.qdu.ps ¹	
	1-11	duwght12	1-12	duwght13	1-13	duwght14	1-14	duwght15	1-15
Minimum	34	1.00	34	0.36	22	0.46	23	0.47	20
1%	92	1.00	101	0.72	100	1.01	120	0.96	118
5%	120	1.00	156	0.83	156	1.05	190	0.99	191
10%	146	1.00	221	0.87	220	1.08	265	0.99	266
25%	326	1.00	486	0.93	485	1.16	596	1.00	596
Median	855	1.26	1,305	0.99	1,296	1.26	1,541	1.00	1,543
75%	1,582	2.03	2,348	1.06	2,342	1.38	2,902	1.00	2,900
90%	2,003	3.57	4,042	1.13	3,968	1.52	5,299	1.01	5,318
95%	2,317	4.79	5,764	1.19	5,630	1.61	7,916	1.01	7,922
99%	3,300	7.98	9,519	1.40	9,247	1.86	12,915	1.04	12,946
Maximum	9,667	9.89	35,464	3.69	24,560	4.35	30,055	3.03	28,180
<i>n</i>	15,122	-	15,122	-	15,122	-	11,710	-	11,710
Mean	1,021	1.85	1,813	1.00	1,793	1.29	2,315	1.00	2,315
Max/Mean	9	-	20	-	14	-	13	-	12

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

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

Model Group 4 Overview

Selected Questionnaire Dwelling Unit-Level Poststratification

Out of 270 proposed effects, 266 were kept in the model. All main effects were maintained in full. Two-factor effects were modified for percent Black or African American \times Rent/Housing, combining 50-100 percent and 10-<50 percent for the first, second, and fourth quintiles. Also combined were \times 50-100 percent and 10-<50 percent Black or African American \times 0-<10 percent Owner-Occupied.

Respondent Questionnaire Dwelling Unit-Level Nonresponse

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

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 2014 NSDUH Questionnaire Dwelling Unit Weights (sel.qdu.ps)
Model Group 4: West**

Variables	Levels	Proposed	Final	Comments
One-Factor Effects		68	68	
Intercept	1	1	1	All levels present.
Group Quarter	3	2	2	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Household Type	7	6	6	All levels present.
Household Size	1	1	1	All levels present.
Rent/Housing	5	4	4	All levels present.
Population Density	4	3	3	All levels present.
% Black or African American	3	2	2	All levels present.
% Hispanic or Latino	35	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
State (Count)	13	12	12	All levels present.
State (Binary)	13	12	12	All levels present.
Quarter (Count)	4	3	3	All levels present.
Quarter (Binary)	4	3	3	All levels present.
Age Group	5	4	4	All levels present.
Race	5	4	4	All levels present.
Hispanicity	2	1	1	All levels present.
Gender	2	1	1	All levels present.
Two-Factor Effects		173	169	
Age × Race (3 Levels)	5 × 3	8	8	All levels present.
Age × Hispanicity	5 × 2	4	4	All levels present.
Age × Gender	5 × 2	4	4	All levels present.
Race (3 Levels) × Hispanicity	3 × 2	2	2	All levels present.
Race (3 Levels) × Gender	3 × 2	2	2	All levels present.
Hispanicity × Gender	2 × 2	1	1	All levels present.
State × Age	13 × 5	48	48	All levels present.
State × Race	13 × 5	48	48	All levels present.
State × Gender	13 × 2	12	12	All levels present.
State × Hispanicity	13 × 2	12	12	All levels present.
% Black or African American × % Owner-Occupied	3 × 3	4	3	Coll. (1,3) & (2,3); sing.
% Black or African American × Rent/Housing	3 × 5	8	5	Coll. (1,1) & (2,1), (1,2) & (2,2), (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.4.2 Covariates for 2014 NSDUH Questionnaire Dwelling Unit Weights (res.qdu.nr)
Model Group 4: West**

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

**Exhibit C.4.3 Covariates for 2014 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**

This page intentionally left blank

Table D.1 2014 NSDUH QDU-Level Response Rates

Domain	Selected QDU	Respondent QDU	% Interview Response Rate¹
Total	64,796	49,672	73.86
<i>Census Region</i>			
Northeast	12,950	9,664	71.62
South	21,448	16,680	74.64
Midwest	15,276	11,618	73.58
West	15,122	11,710	74.65
<i>Quarter</i>			
Quarter 1	14,980	11,476	73.99
Quarter 2	17,779	13,626	73.72
Quarter 3	16,060	12,472	75.37
Quarter 4	15,977	12,098	72.39
<i>Household Type</i>			
12-17, 18-25, 26+	4,862	4,067	83.53
12-17, 18-25	49	45	93.67
12-17, 26+	14,252	11,695	81.92
18-25, 26+	10,923	8,428	76.73
12-17	30	22	80.19
18-25	5,076	4,161	81.73
26+	29,604	21,254	70.90
<i>Race/Ethnicity of Householder</i>			
Hispanic or Latino White	8,539	6,820	77.32
Hispanic or Latino Black or African American	182	160	89.64
Hispanic or Latino Other	505	417	81.88
Non-Hispanic or Latino White	43,044	32,326	72.22
Non-Hispanic or Latino Black or African American	7,324	6,018	79.79
Non-Hispanic or Latino Other	5,202	3,931	70.99
<i>% Hispanic or Latino in Segment</i>			
50-100%	5,053	4,058	77.04
10-<50%	16,177	12,418	74.39
<10%	43,566	33,196	73.23
<i>% Black or African American in Segment</i>			
50-100%	4,613	3,766	78.72
10-<50%	12,617	9,909	76.16
<10%	47,566	35,997	72.69
<i>% Owner-Occupied DUs in Segment</i>			
50-100%	47,513	35,967	72.79
10-<50%	13,599	10,720	76.13
<10%	3,684	2,985	78.83
<i>Combined Median Rent/Housing Value</i>			
1st Quintile	10,327	8,251	77.32
2nd Quintile	13,751	10,863	76.84
3rd Quintile	15,495	11,939	73.99
4th Quintile	13,951	10,487	72.42
5th Quintile	11,272	8,132	69.68
<i>Population Density</i>			
Large MSA	28,666	21,539	72.52
Medium to Small MSA	30,940	24,114	75.54
Non-MSA, Urban	1,609	1,246	73.64
Non-MSA, Rural	3,581	2,773	74.80
<i>Group Quarters</i>			
Group	676	604	85.62
Non-Group	64,120	49,068	73.78
<i>Household Size</i>			
One	8,836	6,661	73.50
Two	27,393	20,354	71.79
Three	15,764	12,242	75.66
Four or More	12,803	10,415	80.30

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.

This page intentionally left blank

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

This page intentionally left blank

Table E.1 2014 NSDUH Selected QDU-Level Proportions of Extreme Values and Outwinors

Domain	n	SDU-Level Weights ¹ (SDUWT: WT1*...*WT11)			Before sel.qdu.ps ¹ (SDUWT*DUWT12)			After sel.qdu.ps ¹ (SDUWT*DUWT12*DUWT13)		
		% Unweighted	% Weighted ²	% Outwinor ³	% Unweighted	% Weighted ²	% Outwinor ³	% Unweighted	% Weighted ²	% Outwinor ³
Total	64,796	2.03	4.43	1.15	1.31	2.48	0.62	1.07	1.93	0.30
<i>Census Region</i>										
Northeast	12,950	2.89	7.35	2.06	1.89	4.29	1.21	1.73	3.71	0.54
South	21,448	1.92	3.87	1.00	1.06	1.84	0.43	0.89	1.42	0.25
Midwest	15,276	1.65	2.85	0.72	1.04	1.72	0.40	0.83	1.11	0.17
West	15,122	1.84	4.61	1.11	1.43	2.87	0.71	1.01	2.15	0.29
<i>Quarter</i>										
Quarter 1	14,980	2.34	4.67	1.39	1.61	2.80	0.77	1.31	2.18	0.35
Quarter 2	17,779	1.63	3.77	0.89	1.00	2.31	0.54	0.89	1.66	0.27
Quarter 3	16,060	2.23	5.09	1.23	1.44	2.58	0.61	1.10	1.91	0.25
Quarter 4	15,977	1.99	4.18	1.07	1.24	2.25	0.58	1.01	1.97	0.31
<i>Household Type</i>										
12-17, 18-25, 26+	4,862	2.43	6.82	1.82	2.43	6.82	1.82	2.04	5.62	1.14
12-17, 18-25	49	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12-17, 26+	14,252	1.86	4.38	1.16	1.86	4.42	1.15	1.66	3.59	0.50
18-25, 26+	10,923	2.13	4.94	1.39	1.87	4.66	1.19	1.48	3.86	0.57
12-17	30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18-25	5,076	2.68	4.45	1.03	2.07	3.86	0.94	1.50	3.35	0.67
26+	29,604	1.91	3.80	0.95	0.52	1.44	0.35	0.41	1.00	0.14
<i>Race/Ethnicity of Householder</i>										
Hispanic or Latino White	8,539	2.33	3.69	0.89	1.36	2.44	0.55	1.15	2.02	0.38
Hispanic or Latino Black or African American	182	56.04	78.09	34.23	41.21	58.71	23.75	42.86	47.65	9.70
Hispanic or Latino Other	505	26.93	52.83	17.21	17.82	34.51	10.47	14.06	30.24	4.26
Non-Hispanic or Latino White	43,044	0.85	1.75	0.36	0.58	1.06	0.21	0.39	0.66	0.08
Non-Hispanic or Latino Black or African American	7,324	3.62	6.82	1.66	2.39	3.42	0.76	2.05	2.89	0.47
Non-Hispanic or Latino Other	5,202	4.77	9.34	1.62	2.75	4.89	0.92	2.50	4.46	0.57

Table E.1 2014 NSDUH Selected QDU-Level Proportions of Extreme Values and Outwinors (continued)

Domain	n	SDU-Level Weights ¹ (SDUWT: WT1*...*WT11)			Before sel.qdu.ps ¹ (SDUWT*DUWT12)			After sel.qdu.ps ¹ (SDUWT*DUWT12*DUWT13)		
		% Unweighted	% Weighted ²	% Outwinor ³	% Unweighted	% Weighted ²	% Outwinor ³	% Unweighted	% Weighted ²	% Outwinor ³
<i>% Hispanic or Latino in Segment</i>										
50-100%	5,053	1.76	4.31	1.37	1.33	2.95	1.05	1.41	2.68	0.51
10-<50%	16,177	2.70	5.90	1.79	1.76	3.62	1.00	1.60	2.93	0.47
<10%	43,566	1.81	3.71	0.79	1.14	1.92	0.40	0.84	1.39	0.19
<i>% Black or African American in Segment</i>										
50-100%	4,613	4.36	10.14	3.31	2.77	5.26	1.44	2.45	4.49	0.82
10-<50%	12,617	2.36	4.88	1.22	1.41	2.58	0.66	1.17	2.07	0.30
<10%	47,566	1.72	3.66	0.89	1.14	2.16	0.53	0.91	1.62	0.24
<i>% Owner-Occupied DUs in Segment</i>										
50-100%	47,513	1.61	3.48	0.85	1.11	1.96	0.43	0.86	1.52	0.22
10-<50%	13,599	2.85	5.74	1.36	1.63	3.13	0.85	1.38	2.40	0.37
<10%	3,684	4.42	10.28	3.67	2.63	6.47	2.10	2.69	5.18	0.98
<i>Combined Median Rent/Housing Value</i>										
1 st Quintile	10,327	1.58	3.35	0.82	0.99	1.58	0.38	0.78	1.45	0.25
2 nd Quintile	13,751	2.06	4.65	1.21	1.40	3.01	0.77	1.16	2.02	0.27
3 rd Quintile	15,495	1.80	3.75	1.01	1.08	1.92	0.54	0.98	1.85	0.29
4 th Quintile	13,951	2.61	5.61	1.56	1.63	3.00	0.81	1.20	2.28	0.42
5 th Quintile	11,272	2.01	4.40	0.99	1.40	2.67	0.53	1.19	1.89	0.22
<i>Population Density</i>										
Large MSA ¹	28,666	2.44	5.34	1.45	1.68	3.12	0.79	1.44	2.49	0.39
Medium to Small MSA ¹	30,940	1.80	3.44	0.80	1.10	1.91	0.46	0.84	1.38	0.19
Non-MSA, ¹ Urban	1,609	1.37	2.64	0.58	0.68	1.04	0.30	0.56	0.58	0.09
Non-MSA, ¹ Rural	3,581	1.01	2.21	0.57	0.42	0.63	0.17	0.36	0.71	0.15
<i>Group Quarters</i>										
Group	676	6.36	7.34	1.09	2.81	5.93	2.07	2.07	4.10	0.70
Non-Group	64,120	1.99	4.40	1.15	1.29	2.46	0.61	1.06	1.91	0.29
<i>Household Size</i>										
One	8,836	1.54	3.44	1.02	0.46	1.34	0.45	0.33	0.63	0.10
Two	27,393	1.95	3.82	1.00	1.02	1.90	0.42	0.69	1.50	0.24
Three	15,764	1.98	4.51	1.18	1.58	3.79	0.98	1.47	3.24	0.47
Four or More	12,803	2.61	6.08	1.47	2.17	5.60	1.32	1.90	4.77	0.73

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

Table E.2 2014 NSDUH Respondent QDU-Level Proportions of Extreme Values and Outwinors

Domain	n	Before res.qdu.nr ¹ (SDUWT*DUWT12*DUWT13)			After res.qdu.nr ¹ (SDUWT*DUWT12*...*DUWT14)			Final Weight: After res.qdu.ps ¹ (SDUWT*DUWT12*...*DUWT15)		
		% Unweighted	% Weighted ²	% Outwinor ³	% Unweighted	% Weighted ²	% Outwinor ³	% Unweighted	% Weighted ²	% Outwinor ³
Total	49,672	1.09	2.08	0.34	0.81	1.44	0.19	0.81	1.38	0.09
<i>Census Region</i>										
Northeast	9,664	1.78	4.19	0.69	1.16	2.96	0.45	1.26	3.22	0.20
South	16,680	0.88	1.59	0.27	0.66	1.07	0.14	0.59	0.87	0.04
Midwest	11,618	0.83	0.99	0.17	0.68	0.88	0.13	0.73	0.91	0.10
West	11,710	1.08	2.32	0.36	0.86	1.38	0.14	0.84	1.23	0.07
<i>Quarter</i>										
Quarter 1	11,476	1.32	2.32	0.40	0.98	1.55	0.20	0.93	1.42	0.08
Quarter 2	13,626	0.87	1.89	0.33	0.68	1.39	0.16	0.67	1.22	0.07
Quarter 3	12,472	1.23	2.03	0.31	0.91	1.50	0.25	0.95	1.54	0.12
Quarter 4	12,098	0.98	2.08	0.33	0.69	1.31	0.16	0.72	1.34	0.08
<i>Household Type</i>										
12-17, 18-25, 26+	4,067	1.97	5.56	1.14	1.60	4.53	0.72	1.57	4.09	0.33
12-17, 18-25	45	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12-17, 26+	11,695	1.71	3.83	0.54	1.16	2.80	0.42	1.23	2.72	0.20
18-25, 26+	8,428	1.58	4.15	0.63	1.01	2.35	0.31	1.01	2.32	0.13
12-17	22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18-25	4,161	1.61	3.63	0.65	1.63	3.35	0.47	1.49	2.84	0.15
26+	21,254	0.29	0.99	0.17	0.23	0.74	0.08	0.23	0.74	0.04
<i>Race/Ethnicity of Householder</i>										
Hispanic or Latino White	6,820	1.20	2.35	0.44	1.10	1.70	0.30	1.09	1.57	0.13
Hispanic or Latino Black or African American	160	40.00	52.39	10.92	16.88	26.18	3.12	19.38	28.37	2.16
Hispanic or Latino Other	417	15.35	31.33	4.98	6.00	14.05	2.47	8.15	16.65	0.87
Non-Hispanic or Latino White	32,326	0.36	0.58	0.06	0.33	0.52	0.05	0.28	0.40	0.02
Non-Hispanic or Latino Black or African American	6,018	1.93	2.88	0.49	1.16	1.82	0.18	1.16	1.73	0.11
Non-Hispanic or Latino Other	3,931	2.57	4.59	0.74	2.47	5.39	0.84	2.62	5.56	0.36
<i>% Hispanic or Latino in Segment</i>										
50-100%	4,058	1.55	3.40	0.63	0.81	1.85	0.27	0.96	2.12	0.15
10-<50%	12,418	1.60	3.14	0.52	0.98	1.77	0.23	1.10	1.83	0.11
<10%	33,196	0.84	1.43	0.23	0.74	1.24	0.17	0.69	1.09	0.07

Table E.2 2014 NSDUH Respondent QDU-Level Proportions of Extreme Values and Outwinsors (continued)

Domain	n	Before res.qdu.nr ¹ (SDUWT*DUWT12*DUWT13)			After res.qdu.nr ¹ (SDUWT*DUWT12*...*DUWT14)			Final Weight: After res.qdu.ps ¹ (SDUWT*DUWT12*...*DUWT15)		
		% Unweighted	% Weighted ²	% Outwinsor ³	% Unweighted	% Weighted ²	% Outwinsor ³	% Unweighted	% Weighted ²	% Outwinsor ³
<i>% Black or African American in Segment</i>										
50-100%	3,766	2.18	3.87	0.71	1.14	2.03	0.35	1.19	2.06	0.10
10-<50%	9,909	1.20	2.42	0.41	0.78	1.63	0.26	0.91	1.76	0.13
<10%	35,997	0.95	1.78	0.28	0.78	1.32	0.16	0.75	1.20	0.08
<i>% Owner-Occupied DUs in Segment</i>										
50-100%	35,967	0.85	1.52	0.25	0.68	1.20	0.16	0.68	1.13	0.08
10-<50%	10,720	1.47	2.89	0.46	1.00	1.85	0.23	1.04	1.86	0.11
<10%	2,985	2.61	5.63	0.97	1.68	2.90	0.41	1.68	2.68	0.18
<i>Combined Median Rent/Housing Value</i>										
1 st Quintile	8,251	0.78	1.38	0.23	0.70	1.09	0.16	0.75	1.08	0.07
2 nd Quintile	10,863	1.24	2.26	0.32	0.58	0.92	0.13	0.61	1.11	0.08
3 rd Quintile	11,939	1.01	2.00	0.36	0.76	1.31	0.17	0.78	1.33	0.09
4 th Quintile	10,487	1.22	2.39	0.45	0.91	1.59	0.23	0.96	1.50	0.10
5 th Quintile	8,132	1.16	2.18	0.30	1.17	2.23	0.27	1.01	1.83	0.09
<i>Population Density</i>										
Large MSA ¹	21,539	1.45	2.69	0.45	1.03	1.87	0.26	1.04	1.81	0.11
Medium to Small MSA ¹	24,114	0.88	1.51	0.24	0.68	1.01	0.13	0.67	0.96	0.07
Non-MSA, ¹ Urban	1,246	0.48	0.68	0.12	0.48	0.75	0.16	0.56	0.76	0.08
Non-MSA, ¹ Rural	2,773	0.43	0.76	0.18	0.36	0.29	0.03	0.40	0.36	0.03
<i>Group Quarters</i>										
Group	604	2.98	8.55	1.27	0.99	5.27	0.46	0.66	4.55	0.21
Non-Group	49,068	1.07	2.03	0.34	0.81	1.41	0.19	0.82	1.36	0.09
<i>Household Size</i>										
One	6,661	0.36	0.86	0.11	0.41	0.53	0.04	0.38	0.50	0.03
Two	20,354	0.66	1.58	0.32	0.54	1.16	0.14	0.53	1.12	0.07
Three	12,242	1.49	3.32	0.46	0.84	1.76	0.22	0.99	2.13	0.15
Four or More	10,415	1.92	4.79	0.74	1.56	4.11	0.69	1.45	3.41	0.22

¹ DU = dwelling unit, MSA = metropolitan statistical area, NR = nonresponse adjustment, PS = poststratification adjustment, QDU = questionnaire dwelling unit, Res = Respondent, SDU = screener dwelling unit.

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

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

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

This page intentionally left blank

Table F.1 2014 NSDUH QDU-Level Slippage Rates

Domain	<i>n</i>	Initial Total (<i>I</i>)¹	Final Total (<i>F</i>)²	Control from SDU Weights (<i>C</i>)	(<i>I - C</i>)/<i>C</i>%	(<i>F - C</i>)/<i>C</i>%
Total	49,672	121,783,728	121,783,727	121,783,727	0.00	0.00
<i>Census Region</i>						
Northeast	9,664	21,947,536	21,947,536	21,947,536	0.00	0.00
South	16,680	45,816,540	45,816,540	45,816,540	0.00	-0.00
Midwest	11,618	26,911,664	26,911,664	26,911,664	0.00	0.00
West	11,710	27,107,987	27,107,987	27,107,987	0.00	-0.00
<i>Quarter</i>						
Quarter 1	11,476	30,469,672	30,469,672	30,469,672	0.00	-0.00
Quarter 2	13,626	30,170,066	30,170,066	30,170,066	0.00	0.00
Quarter 3	12,472	30,412,151	30,412,151	30,412,151	0.00	0.00
Quarter 4	12,098	30,731,838	30,731,838	30,731,838	0.00	0.00
<i>Household Type</i>						
12-17, 18-25, 26+	4,067	5,249,395	5,249,395	5,249,395	-0.00	-0.00
12-17, 18-25	45	44,690	44,690	44,690	0.00	0.00
12-17, 26+	11,695	13,464,785	13,464,785	13,464,785	0.00	0.00
18-25, 26+	8,428	14,182,098	14,182,098	14,182,098	0.00	0.00
12-17	22	25,737	25,737	25,737	0.00	0.00
18-25	4,161	5,736,867	5,736,867	5,736,867	0.00	0.00
26+	21,254	83,080,156	83,080,156	83,080,156	0.00	0.00
<i>Race/Ethnicity of Householder</i>						
Hispanic or Latino White	6,820	14,348,908	14,348,908	14,348,908	0.00	0.00
Hispanic or Latino Black or African American	160	766,451	766,451	766,451	0.00	0.00
Hispanic or Latino Other	417	1,156,851	1,156,851	1,156,851	0.00	0.00
Non-Hispanic or Latino White	32,326	82,370,880	82,370,880	82,370,880	0.00	-0.00
Non-Hispanic or Latino Black or African American	6,018	14,881,661	14,881,661	14,881,661	0.00	-0.00
Non-Hispanic or Latino Other	3,931	8,258,977	8,258,977	8,258,977	0.00	0.00
<i>% Hispanic or Latino in Segment</i>						
50-100%	4,058	9,959,879	9,959,879	9,959,879	-0.00	0.00
10-<50%	12,418	33,994,116	33,994,116	33,994,116	0.00	0.00
<10%	33,196	77,829,733	77,829,733	77,829,733	0.00	0.00
<i>% Black or African American in Segment</i>						
50-100%	3,766	9,053,321	9,053,321	9,053,321	0.00	0.00
10-<50%	9,909	25,507,748	25,507,748	25,507,748	0.00	0.00
<10%	35,997	87,222,659	87,222,658	87,222,658	0.00	0.00
<i>% Owner-Occupied DUs in Segment</i>						
50-100%	35,967	88,579,914	88,579,914	88,579,914	0.00	0.00
10-<50%	10,720	25,889,909	25,889,909	25,889,909	0.00	0.00
<10%	2,985	7,313,905	7,313,905	7,313,905	0.00	0.00

Table F.1 2014 NSDUH QDU-Level Slippage Rates (continued)

Domain	<i>n</i>	Initial Total (<i>I</i>)¹	Final Total (<i>F</i>)²	Control from SDU Weights (<i>C</i>)	(<i>I</i> - <i>C</i>)/<i>C</i>%	(<i>F</i> - <i>C</i>)/<i>C</i>%
<i>Combined Median Rent/Housing Value</i>						
1 st Quintile	8,251	17,321,549	17,321,549	17,321,549	0.00	0.00
2 nd Quintile	10,863	25,284,353	25,284,353	25,284,353	0.00	0.00
3 rd Quintile	11,939	28,423,089	28,423,089	28,423,089	0.00	-0.00
4 th Quintile	10,487	26,821,760	26,821,760	26,821,760	0.00	0.00
5 th Quintile	8,132	23,932,977	23,932,977	23,932,977	0.00	0.00
<i>Population Density</i>						
Large MSA	21,539	64,936,424	64,936,424	64,936,424	0.00	0.00
Medium to Small MSA	24,114	49,481,444	49,481,444	49,481,444	0.00	0.00
Non-MSA, Urban	1,246	2,499,094	2,499,094	2,499,094	0.00	-0.00
Non-MSA, Rural	2,773	4,866,766	4,866,766	4,866,766	0.00	-0.00
<i>Group Quarters</i>						
Group	604	826,033	826,033	826,033	0.00	0.00
Non-Group	49,068	120,957,695	120,957,695	120,957,695	0.00	0.00
<i>Household Size</i>						
One	6,661	32,728,215	32,703,742	32,490,024	0.73	0.66
Two	20,354	55,444,416	55,458,172	55,696,048	-0.45	-0.43
Three	12,242	19,157,606	19,196,438	19,278,166	-0.63	-0.42
Four or More	10,415	14,453,490	14,425,376	14,319,489	0.94	0.74

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**

This page intentionally left blank

Table G.1 2014 NSDUH Selected QDU-Level Weight Summary Statistics

Domain	n	SDU-Level Weights ¹ (SDUWT: WT1*...*WT11)						Before sel.qdu.ps ¹ (SDUWT*DUWT12)						After sel.qdu.ps ¹ (SDUWT*DUWT12*DUWT13)					
		Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³
Total	64,796	14	525	967	1,336	9,667	1.41	14	762	1,357	2,282	53,116	2.02	11	746	1,355	2,289	24,560	1.99
<i>Census Region</i>																			
Northeast	12,950	33	303	910	1,126	8,241	1.46	33	610	1,184	1,931	53,116	2.29	44	582	1,191	1,957	21,217	2.20
South	21,448	14	766	1,098	1,467	7,583	1.31	14	975	1,549	2,565	29,230	1.84	11	970	1,555	2,581	22,788	1.84
Midwest	15,276	33	634	917	1,169	7,734	1.26	33	814	1,269	2,076	45,450	1.94	47	797	1,260	2,071	15,588	1.91
West	15,122	34	326	855	1,582	9,667	1.60	34	486	1,305	2,348	35,464	2.16	22	485	1,296	2,342	24,560	2.13
<i>Quarter</i>																			
Quarter 1	14,980	26	605	1,056	1,464	8,241	1.39	26	831	1,471	2,480	45,450	2.00	17	816	1,471	2,477	24,560	1.96
Quarter 2	17,779	19	488	874	1,201	7,734	1.41	34	683	1,223	2,028	30,089	2.04	26	663	1,219	2,037	22,788	2.02
Quarter 3	16,060	14	510	976	1,329	8,241	1.43	14	763	1,365	2,302	40,175	2.00	11	753	1,370	2,314	20,082	1.98
Quarter 4	15,977	16	559	998	1,348	9,667	1.38	16	796	1,396	2,307	53,116	1.99	11	783	1,402	2,326	22,917	1.98
<i>Household Type</i>																			
12-17, 18-25, 26+	4,862	16	590	1,033	1,437	7,583	1.44	16	590	1,033	1,437	7,583	1.44	11	577	1,024	1,446	9,527	1.44
12-17, 18-25	49	34	294	842	1,296	2,638	1.60	34	294	842	1,297	2,639	1.60	22	298	870	1,331	2,930	1.58
12-17, 26+	14,252	14	486	922	1,273	7,182	1.40	14	487	924	1,275	7,385	1.40	20	484	925	1,289	7,865	1.39
18-25, 26+	10,923	16	639	1,067	1,452	6,946	1.40	24	740	1,251	1,683	8,419	1.38	17	726	1,256	1,719	8,329	1.37
12-17	30	108	522	866	1,190	1,738	1.22	109	527	875	1,202	1,756	1.22	105	455	948	1,197	1,745	1.24
18-25	5,076	34	395	917	1,312	7,083	1.48	35	454	1,076	1,549	8,368	1.46	51	449	1,077	1,574	6,832	1.48
26+	29,604	19	528	958	1,308	9,667	1.38	37	1,253	2,212	3,687	53,116	1.71	33	1,216	2,193	3,678	24,560	1.70
<i>Race/Ethnicity of Householder</i>																			
Hispanic or Latino White	8,539	33	699	1,090	1,524	7,182	1.29	38	837	1,381	1,994	17,162	1.71	44	823	1,381	2,026	20,081	1.73
Hispanic or Latino Black or African American	182	31	1,232	2,273	3,831	8,069	1.49	31	1,519	3,161	5,493	53,116	2.85	36	1,674	3,094	5,152	21,523	2.02
Hispanic or Latino Other	505	14	297	826	2,129	8,241	2.13	14	355	1,232	3,198	40,175	3.05	11	396	1,355	3,286	24,560	2.49
Non-Hispanic or Latino White	43,044	44	477	943	1,250	9,667	1.37	45	754	1,336	2,331	29,230	2.02	35	737	1,331	2,326	22,788	2.02
Non-Hispanic or Latino Black or African American	7,324	31	779	1,087	1,437	7,083	1.32	34	955	1,486	2,406	17,231	1.84	22	943	1,505	2,421	15,662	1.84
Non-Hispanic or Latino Other	5,202	25	249	806	1,496	6,516	1.65	25	394	1,166	2,130	25,908	2.11	17	387	1,183	2,183	21,859	2.13
<i>% Hispanic or Latino in Segment</i>																			
50-100%	5,053	16	837	1,306	1,614	8,069	1.26	16	1,049	1,558	2,331	53,116	1.81	11	1,028	1,544	2,291	21,217	1.71
10-<50%	16,177	14	738	1,124	1,592	8,241	1.34	14	952	1,578	2,604	45,450	1.86	17	940	1,586	2,603	24,560	1.81
<10%	43,566	16	369	904	1,202	9,667	1.43	16	641	1,236	2,120	28,435	2.11	11	630	1,238	2,133	22,788	2.11

Table G.1 2014 NSDUH Selected QDU-Level Weight Summary Statistics (continued)

Domain	n	SDU-Level Weights ¹ (SDUWT: WT1*...*WT11)						Before sel.qdu.ps ¹ (SDUWT*DUWT12)						After sel.qdu.ps ¹ (SDUWT*DUWT12*DUWT13)					
		Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³
<i>% Black or African American in Segment</i>																			
50-100%	4,613	14	734	1,043	1,369	6,910	1.40	14	877	1,406	2,324	29,230	1.99	11	867	1,413	2,301	20,082	1.99
10-<50%	12,617	16	779	1,061	1,421	8,241	1.28	16	980	1,508	2,448	40,175	1.82	11	956	1,500	2,436	24,560	1.79
<10%	47,566	33	416	933	1,306	9,667	1.44	33	668	1,298	2,224	53,116	2.08	22	658	1,305	2,239	22,917	2.05
<i>% Owner-Occupied DUs¹ in Segment</i>																			
50-100%	47,513	14	502	955	1,303	8,241	1.40	14	750	1,336	2,268	28,435	1.97	11	738	1,339	2,275	22,788	1.97
10-<50%	13,599	16	573	994	1,383	9,667	1.41	17	771	1,379	2,270	45,450	2.12	17	750	1,381	2,275	24,560	2.06
<10%	3,684	28	682	1,097	1,553	8,069	1.44	50	850	1,490	2,483	53,116	2.16	47	833	1,468	2,448	21,217	2.01
<i>Combined Median Rent/Housing Value</i>																			
1 st Quintile	10,327	33	346	842	1,175	6,640	1.45	33	571	1,179	1,997	16,826	2.07	22	573	1,195	2,016	14,931	2.09
2 nd Quintile	13,751	33	498	930	1,241	7,182	1.41	33	737	1,287	2,103	45,450	2.12	26	733	1,324	2,153	24,560	2.05
3 rd Quintile	15,495	26	523	961	1,349	9,667	1.40	26	764	1,355	2,222	22,135	1.99	17	732	1,332	2,218	22,917	2.00
4 th Quintile	13,951	14	592	1,022	1,419	8,241	1.40	14	810	1,427	2,377	53,116	2.00	11	781	1,399	2,374	21,523	1.96
5 th Quintile	11,272	16	703	1,088	1,466	8,241	1.35	16	912	1,519	2,645	30,089	1.89	11	908	1,533	2,678	20,358	1.87
<i>Population Density</i>																			
Large MSA ¹	28,666	14	890	1,163	1,553	9,667	1.26	14	1,098	1,661	2,694	53,116	1.78	11	1,095	1,673	2,726	22,917	1.76
Medium to Small MSA ¹	30,940	33	320	802	1,146	6,946	1.48	33	525	1,109	1,915	45,450	2.20	17	524	1,112	1,908	24,560	2.17
Non-MSA, ¹ Urban	1,609	64	291	693	1,034	3,958	1.54	82	461	995	1,904	24,477	2.56	77	449	1,000	1,857	17,548	2.37
Non-MSA, ¹ Rural	3,581	37	186	503	977	4,721	1.66	38	310	874	1,668	14,118	2.38	40	305	875	1,650	15,588	2.41
<i>Group Quarters</i>																			
Group	676	31	206	558	1,244	4,693	1.86	37	297	685	1,685	30,089	3.55	54	280	612	1,615	20,358	2.99
Non-Group	64,120	14	534	970	1,336	9,667	1.40	14	770	1,361	2,287	53,116	2.01	11	754	1,359	2,296	24,560	1.99
<i>Household Size</i>																			
One	8,836	41	478	923	1,248	8,069	1.39	90	1,296	2,679	5,194	53,116	1.79	83	1,283	2,678	5,156	24,560	1.75
Two	27,393	17	511	953	1,301	8,241	1.39	18	866	1,606	2,666	22,135	1.65	22	861	1,593	2,678	21,523	1.67
Three	15,764	17	558	987	1,352	9,667	1.40	17	624	1,090	1,587	20,324	1.53	18	615	1,096	1,613	9,527	1.51
Four or More	12,803	14	560	1,024	1,446	8,241	1.44	14	575	1,047	1,499	8,248	1.46	11	561	1,039	1,500	9,624	1.47

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 2014 NSDUH Respondent QDU-Level Weight Summary Statistics

Domain	n	Before res.qdu.nr ¹ (SDUWT*DUWT12*DUWT13)						After res.qdu.nr ¹ (SDUWT*DUWT12*...*DUWT14)						Final Weight: After res.qdu.ps ¹ (SDUWT*DUWT12*...*DUWT15)					
		Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³
Total	49,672	11	723	1,311	2,181	24,560	2.01	12	888	1,661	2,918	32,038	2.16	12	887	1,663	2,917	32,049	2.16
<i>Census Region</i>																			
Northeast	9,664	44	550	1,145	1,861	21,217	2.26	50	677	1,486	2,622	30,966	2.40	46	667	1,482	2,636	30,852	2.40
South	16,680	11	948	1,512	2,452	18,185	1.84	12	1,139	1,883	3,235	32,038	2.02	12	1,141	1,881	3,223	32,049	2.02
Midwest	11,618	47	770	1,216	1,988	13,117	1.94	61	964	1,563	2,684	18,709	2.03	63	964	1,565	2,683	18,739	2.03
West	11,710	22	477	1,241	2,234	24,560	2.17	23	596	1,541	2,902	30,055	2.32	20	596	1,543	2,900	28,180	2.32
<i>Quarter</i>																			
Quarter 1	11,476	17	773	1,418	2,356	24,560	2.00	20	958	1,804	3,136	25,798	2.13	21	957	1,804	3,129	25,818	2.13
Quarter 2	13,626	26	639	1,184	1,952	20,358	2.03	34	791	1,506	2,637	28,664	2.18	33	792	1,506	2,636	28,597	2.18
Quarter 3	12,472	11	738	1,327	2,222	20,082	2.00	12	888	1,653	2,922	32,038	2.14	12	888	1,656	2,920	32,049	2.14
Quarter 4	12,098	11	762	1,356	2,222	22,917	1.98	21	954	1,731	3,027	30,816	2.17	21	953	1,729	3,028	30,850	2.17
<i>Household Type</i>																			
12-17, 18-25, 26+	4,067	11	570	1,017	1,447	9,527	1.45	12	654	1,209	1,734	6,447	1.44	12	656	1,211	1,734	6,168	1.44
12-17, 18-25	45	22	298	870	1,392	2,930	1.58	23	306	914	1,424	3,721	1.65	20	305	913	1,413	3,713	1.65
12-17, 26+	11,695	21	490	921	1,287	7,865	1.39	22	581	1,104	1,579	6,144	1.41	23	577	1,105	1,584	5,938	1.41
18-25, 26+	8,428	17	721	1,244	1,709	8,329	1.38	24	896	1,601	2,253	7,671	1.40	24	893	1,600	2,257	7,861	1.40
12-17	22	206	625	1,010	1,223	1,745	1.18	228	878	1,223	1,538	2,343	1.21	227	878	1,209	1,539	2,315	1.21
18-25	4,161	51	442	1,071	1,574	6,321	1.48	58	538	1,312	1,941	7,080	1.50	59	539	1,311	1,944	6,880	1.49
26+	21,254	33	1,187	2,158	3,609	24,560	1.71	39	1,589	2,961	5,153	32,038	1.75	39	1,585	2,955	5,158	32,049	1.75
<i>Race/Ethnicity of Householder</i>																			
Hispanic or Latino White	6,820	44	809	1,350	1,949	20,081	1.74	51	964	1,643	2,485	21,791	1.89	46	962	1,649	2,495	21,690	1.89
Hispanic or Latino Black or African American	160	36	1,597	3,194	5,268	21,523	2.06	36	1,707	3,214	5,650	28,335	2.12	35	1,707	3,315	6,050	27,216	2.12
Hispanic or Latino Other	417	11	428	1,417	3,271	24,560	2.47	12	524	1,566	3,698	30,055	2.61	12	506	1,574	3,865	28,180	2.53
Non-Hispanic or Latino White	32,326	35	707	1,283	2,210	20,556	2.04	44	895	1,671	3,034	32,038	2.18	33	895	1,673	3,031	32,049	2.18
Non-Hispanic or Latino Black or African American	6,018	22	923	1,464	2,338	15,540	1.86	23	1,049	1,740	2,922	23,668	2.00	20	1,048	1,740	2,923	23,602	2.00
Non-Hispanic or Latino Other	3,931	17	356	1,092	2,045	18,263	2.15	25	453	1,366	2,732	23,053	2.39	25	446	1,354	2,752	24,140	2.40
<i>% Hispanic or Latino in Segment</i>																			
50-100%	4,058	11	982	1,504	2,179	21,217	1.73	12	1,190	1,808	2,814	28,335	1.88	12	1,191	1,807	2,821	27,216	1.88
10-<50%	12,418	17	911	1,532	2,499	24,560	1.84	20	1,097	1,941	3,367	30,055	1.98	20	1,096	1,946	3,373	28,180	1.97
<10%	33,196	11	605	1,202	2,037	20,556	2.13	21	768	1,528	2,738	32,038	2.28	21	768	1,530	2,736	32,049	2.28

Table G.2 2014 NSDUH Respondent QDU-Level Weight Summary Statistics (continued)

Domain	n	Before res.qdu.nr ¹ (SDUWT*DUWT12*DUWT13)						After res.qdu.nr ¹ (SDUWT*DUWT12*...*DUWT14)						Final Weight: After res.qdu.ps ¹ (SDUWT*DUWT12*...*DUWT15)					
		Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³
% Black or African American in Segment																			
50-100%	3,766	11	854	1,371	2,207	20,082	1.99	20	969	1,647	2,819	22,952	2.12	21	964	1,657	2,824	22,708	2.12
10-<50%	9,909	11	931	1,458	2,344	24,560	1.80	12	1,116	1,812	3,075	25,798	1.92	12	1,114	1,812	3,079	26,863	1.92
<10%	35,997	22	629	1,257	2,129	22,917	2.08	23	795	1,613	2,886	32,038	2.24	20	795	1,613	2,883	32,049	2.24
% Owner-Occupied DUs¹ in Segment																			
50-100%	35,967	11	712	1,293	2,166	21,523	1.99	12	887	1,657	2,942	32,038	2.15	12	887	1,659	2,936	32,049	2.15
10-<50%	10,720	17	733	1,332	2,179	24,560	2.09	24	879	1,649	2,842	30,055	2.23	24	877	1,650	2,843	28,180	2.23
<10%	2,985	47	802	1,440	2,395	21,217	2.02	55	926	1,745	2,988	28,335	2.15	54	928	1,745	2,991	27,216	2.15
Combined Median Rent/Housing Value																			
1 st Quintile	8,251	22	563	1,176	1,932	14,931	2.11	23	683	1,435	2,445	18,887	2.23	20	680	1,434	2,443	18,727	2.23
2 nd Quintile	10,863	26	728	1,291	2,088	24,560	2.05	34	869	1,593	2,668	32,038	2.20	35	868	1,597	2,668	32,049	2.20
3 rd Quintile	11,939	17	702	1,296	2,117	22,917	2.01	20	869	1,633	2,817	30,816	2.16	21	866	1,633	2,819	30,850	2.16
4 th Quintile	10,487	11	754	1,346	2,269	21,523	1.98	21	938	1,753	3,109	28,335	2.12	21	939	1,753	3,113	27,216	2.11
5 th Quintile	8,132	11	874	1,471	2,547	20,358	1.90	12	1,120	1,970	3,638	30,966	2.04	12	1,125	1,973	3,626	30,852	2.04
Population Density																			
Large MSA ¹	21,539	11	1,066	1,614	2,597	22,917	1.78	12	1,334	2,096	3,567	32,038	1.92	12	1,332	2,099	3,564	32,049	1.92
Medium to Small MSA ¹	24,114	17	512	1,089	1,843	24,560	2.19	23	640	1,347	2,379	25,798	2.32	20	640	1,346	2,379	26,863	2.32
Non-MSA, ¹ Urban	1,246	77	425	969	1,786	16,212	2.34	86	565	1,217	2,340	30,966	2.51	81	565	1,217	2,343	30,852	2.51
Non-MSA, ¹ Rural	2,773	40	296	854	1,573	11,391	2.44	44	373	1,065	2,096	16,621	2.58	33	373	1,068	2,093	16,620	2.59
Group Quarters																			
Group	604	54	273	612	1,612	20,358	2.90	58	299	678	1,904	26,001	3.36	59	298	692	1,903	26,023	3.36
Non-Group	49,068	11	731	1,315	2,190	24,560	2.01	12	899	1,669	2,938	32,038	2.15	12	899	1,671	2,933	32,049	2.15
Household Size																			
One	6,661	83	1,242	2,603	4,965	24,560	1.77	95	1,596	3,426	6,783	32,038	1.84	85	1,593	3,425	6,780	32,049	1.84
Two	20,354	22	835	1,542	2,558	21,523	1.69	23	1,053	2,032	3,515	21,825	1.79	20	1,049	2,032	3,513	20,187	1.79
Three	12,242	18	606	1,078	1,578	9,527	1.49	22	732	1,345	2,042	11,227	1.60	21	730	1,344	2,049	11,574	1.61
Four or More	10,415	11	551	1,030	1,484	9,624	1.47	12	645	1,245	1,841	16,260	1.57	12	643	1,244	1,845	12,872	1.55

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

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

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

Appendix H: GEM Modeling Summary for the Pair Weights

This page intentionally left blank

Appendix H: GEM Modeling Summary for the Pair Weights

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

For 2014, modeling involved taking two model groups through four adjustment steps: (1) selected pair poststratification, (2) pair nonresponse adjustment, (3) responding pair poststratification, and (4) responding pair extreme value adjustment.

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

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

Final Model Explanatory Variables

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

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

Group Quarter Indicator

1: College Dorm, 2: Other Group Quarter, 3: Non-Group Quarter¹

Household Size

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

Pair Age (15 Levels)

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

Pair Age (6 Levels)

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

Pair Age (3 Levels)

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

Pair Gender

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

Pair Race/Ethnicity (10 Levels)

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

Pair Race/Ethnicity (5 Levels)

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

Pair Race/Ethnicity (4 Levels)

1: Two or More Races Pair or Other and Other, 2: Hispanic or Latino Pair, 3: Black or African American Pair, 4: White Pair¹

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 Quintile¹

Population Density

1: MSA 1,000,000 or More, 2: MSA Less than 1,000,000, 3: Non-MSA Urban, 4: Non-MSA Rural¹

Quarter

1: Quarter 1, 2: Quarter 2, 3: Quarter 3, 4: Quarter 4¹

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)

States³

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 2014 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)

This page intentionally left blank

Table H.1a 2014 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
<i>sel.pr.ps</i>	4.27	18.98	8.31	10.3007	213	(0.27, 1.30)	(0.27, 1.30)
	0.89	2.48	0.26	4.1421	201	(0.25, 2.44)	(0.25, 2.44)
						(0.90, 1.39)	(1.39, 1.39)
<i>res.pr.nr</i>	1.18	4.96	0.43	4.0224	213	(1.00, 1.70)	(1.00, 1.70)
	2.15	8.46	1.84	5.5871	212	(1.00, 5.00)	(1.00, 5.00)
						N/A	N/A
<i>res.pr.ps</i>	2.15	8.33	2.00	5.5871	223	(0.45, 2.00)	(0.46, 2.00)
	1.46	7.54	1.12	5.7419	211	(0.23, 2.81)	(0.24, 2.81)
						(0.90, 5.00)	(0.90, 0.90)
<i>res.pr.ev</i>	1.46	7.54	1.12	5.7419	223	(0.95, 1.40)	(0.95, 1.40)
	0.74	4.69	0.35	5.6595	211	(0.85, 1.47)	(0.85, 1.47)
						N/A	N/A

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

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

² Unequal weighting effect (UWE) defined as $1 + [(n - 1)/n] * CV^2$, 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 2014 Distribution of Weight Adjustment Factors and Weight Products (Model Group 1: Northeast and South)

	SDU Weight	Pair Selection Prob		sel.pr.ps ¹		res.pr.nr ¹		res.pr.ps ¹		res.pr.ev ¹	
	1-11	pairwt12	1-12	pairwt13	1-13	pairwt14	1-14	pairwt15	1-15	pairwt16	1-16
Minimum	14	1.02	21	0.08	34	0.51	36	0.24	21	0.54	18
1%	93	1.26	255	0.27	239	0.97	270	0.33	210	0.85	205
5%	156	1.45	579	0.47	564	1.01	660	0.51	587	0.93	572
10%	248	1.66	995	0.67	919	1.03	1,096	0.64	987	0.95	964
25%	681	2.62	2,035	0.88	1,959	1.11	2,288	0.86	2,186	0.98	2,157
Median	1,038	4.06	4,014	1.06	4,156	1.29	5,081	1.03	5,022	1.00	5,001
75%	1,401	8.47	8,785	1.27	9,094	1.57	11,859	1.18	11,643	1.01	11,620
90%	1,844	16.69	18,121	1.52	20,004	2.04	28,415	1.32	28,516	1.03	28,585
95%	2,136	26.41	30,602	1.70	32,603	2.54	51,147	1.41	51,202	1.05	51,774
99%	3,168	53.88	64,762	2.06	72,318	4.11	136,722	1.65	136,217	1.12	137,732
Maximum	7,583	989.83	2,040,767	2.44	226,138	5.00	470,163	2.81	514,719	1.47	419,968
<i>n</i>	13,969	-	13,969	-	13,969	-	9,436	-	9,436	-	9,436
Mean	1,079	8.07	8,688	1.08	8,823	1.46	13,061	1.01	13,061	0.99	13,061
Max/Mean	7	-	235	-	26	-	36	-	39	-	32

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, 201 of 213 proposed factors were retained in the final model. All main and two-factor effects were retained at proposed levels. None of the 12 three-factor effects were kept in the model because of convergence problems.

Respondent Pair-Level Nonresponse

In the respondent pair-level nonresponse step, 212 of 213 proposed factors were retained in the final model. State/Region by Pair Race/Ethnicity collapsed the Black or African American pair level with the Other pair level for New York because of convergence problems.

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 at the proposed levels. None of the 12 three-factor effects were kept in the model because of convergence problems.

Respondent Pair-Level Extreme Value Adjustment

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

Exhibit H.1.1 Covariates for 2014 NSDUH Pair Weights (sel.pr.ps) Model Group 1: Northeast and South

Variables	Level	Proposed	Final	Comments
One-Factor Effects		76	76	
Intercept	1	1	1	All levels present.
State	26	25	25	All levels present.
Quarter	4	3	3	All levels present.
Population Density	4	3	3	All levels present.
Group Quarter	3	2	2	All levels present.
Household Size	3	2	2	All levels present.
Pair Age	15	14	14	All levels present.
Pair Gender	3	2	2	All levels present.
Pair Race/Ethnicity	10	9	9	All levels present.
Race/Ethnicity of Householder	6	5	5	All levels present.
Rent/Housing	5	4	4	All levels present.
Segment % Black or African American	3	2	2	All levels present.
Segment % Hispanic or Latino	3	2	2	All levels present.
% Owner-Occupied	3	2	2	All levels present.
Two-Factor Effects		125	125	
Pair Race/Ethnicity (5 Levels) × Pair Age (6 Levels)	5 × 6	20	20	All levels present.
Pair Race/Ethnicity (5 Levels) × Pair Gender	5 × 3	8	8	All levels present.
Pair Gender × Pair Age (6 Levels)	3 × 6	10	10	All levels present.
State/Region × Pair Race/Ethnicity (5 Levels)	6 × 5	20	20	All levels present.
State/Region × Pair Age (6 Levels)	6 × 6	25	25	All levels present.
State/Region × Pair Gender	6 × 3	10	10	All levels present.
Rent/Housing × % Black or African American	5 × 3	8	8	All levels present.
Rent/Housing × % Hispanic or Latino	5 × 3	8	8	All levels present.
Rent/Housing × % Owner-Occupied	5 × 3	8	8	All levels present.
% Owner-Occupied × % Black or African American	3 × 3	4	4	All levels present.
% Owner-Occupied × % Hispanic or Latino	3 × 3	4	4	All levels present.
Three-Factor Effects		12	0	
Pair Race/Ethnicity (4 Levels) × Pair Gender × Pair Age (3 Levels)	4 × 3 × 3	12	0	Drop all; conv.
Total		213	201	

Exhibit H.1.2 Covariates for 2014 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	124	
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	19	Coll. (5,3) & (5,5); conv.
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	212	

Exhibit H.1.3 Covariates for 2014 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 2014 NSDUH Pair Weights (res.pr.ev) Model Group 1: Northeast and South

This step used the same variables as the respondent pair-level nonresponse and poststratification steps in [Exhibits H.1.2](#) and [H.1.3](#), respectively.

This page intentionally left blank

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

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

This page intentionally left blank

Table H.2a 2014 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
<i>sel.pr.ps</i>	3.89	19.77	9.37	9.9230	212	(0.35, 1.50)	(0.35, 1.50)
	1.51	4.27	0.36	4.2405	200	(0.26, 2.04)	(0.27, 2.03)
						(0.90, 1.53)	(1.53, 1.53)
<i>res.pr.nr</i>	1.92	6.21	0.78	4.4648	212	(1.00, 1.60)	(1.00, 1.60)
	1.73	4.89	0.91	5.1655	211	(1.00, 4.45)	(1.00, 4.42)
						N/A	N/A
<i>res.pr.ps</i>	1.84	5.22	1.19	5.1655	222	(0.63, 1.30)	(0.64, 1.30)
	1.69	4.01	0.26	5.2648	209	(0.32, 1.46)	(0.33, 1.45)
						N/A	N/A
<i>res.pr.ev</i>	1.69	4.01	0.26	5.2648	222	(0.99, 1.10)	(0.99, 1.10)
	0.10	0.10	0.00	5.2522	209	(0.94, 1.20)	(0.94, 1.20)
						(0.99, 1.50)	(0.99, 0.99)

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

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

² Unequal weighting effect (UWE) defined as $1 + [(n - 1)/n] * CV^2$, 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 2014 Distribution of Weight Adjustment Factors and Weight Products (Model Group 2: Midwest and West)

	SDU Weight	Pair Selection		sel.pr.ps ¹		res.pr.nr ¹		res.pr.ps ¹		res.pr.ev ¹	
	1-11	pairwt12	1-12	pairwt13	1-13	pairwt14	1-14	pairwt15	1-15	pairwt16	1-16
Minimum	38	1.02	94	0.08	32	0.54	35	0.07	22	0.68	21
1%	99	1.19	233	0.31	180	0.93	211	0.38	187	0.90	184
5%	138	1.43	449	0.51	404	1.06	532	0.62	501	0.96	496
10%	183	1.69	740	0.66	700	1.10	886	0.78	830	0.98	821
25%	466	2.61	1,592	0.84	1,587	1.20	1,973	0.93	1,930	0.99	1,916
Median	915	4.17	3,575	1.04	3,607	1.35	4,710	1.04	4,729	1.00	4,739
75%	1,367	8.17	8,044	1.26	8,269	1.59	10,968	1.12	10,840	1.01	10,875
90%	1,812	16.13	16,931	1.47	18,234	1.91	25,513	1.21	25,463	1.02	25,304
95%	2,128	27.27	28,487	1.60	30,858	2.20	47,603	1.27	46,663	1.04	46,572
99%	3,063	50.63	64,088	1.85	68,783	2.92	118,611	1.37	120,397	1.12	121,512
Maximum	8,241	643.22	1,032,769	2.64	243,331	4.42	455,189	1.45	394,739	1.77	358,720
<i>n</i>	12,875	-	12,875	-	12,875	-	8,793	-	8,793	-	8,793
Mean	991	8.07	8,214	1.05	8,019	1.45	11,742	1.01	11,742	1.00	11,742
Max/Mean	8	-	126	-	30	-	39	-	34	-	31

SDU = screener dwelling unit.

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

Model Group 2 Overview

Selected Pair-Level Poststratification

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

Respondent Pair-Level Nonresponse

In the respondent pair-level nonresponse step, 211 of 212 proposed factors were retained in the final model. The fourth quintile of segment-combined median rent and housing value crossed with the 50-100% level of percentage of segments that are Black or African American was removed because of linear combination.

Respondent Pair-Level Poststratification

In the respondent pair-level poststratification step, 209 of 222 proposed factors were retained in the final model. Similar to the respondent pair-level nonresponse step, the fourth quintile of segment-combined median rent and housing value crossed with the 50-100% level of percentage of segments that are Black or African American was a linear combination, but in this step, the third and fourth quintiles were collapsed across the 50-100% level to address the issue.

Respondent Pair-Level Extreme Value Adjustment

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

Exhibit H.2.1 Covariates for 2014 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 2014 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	124	
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	7	Drop (4,1); sing.
Rent/Housing × % Hispanic or Latino	5 × 3	8	8	All levels present.
Rent/Housing × % Owner-Occupied	5 × 3	8	8	All levels present.
% Owner-Occupied × % Black or African American	3 × 3	4	4	All levels present.
% Owner-Occupied × % Hispanic or Latino	3 × 3	4	4	All levels present.
Three-Factor Effects		12	12	
Pair Race/Ethnicity (4 Levels) × Pair Gender × Pair Age (3 Levels)	4 × 3 × 3	12	12	All levels present.
Total		212	211	

Exhibit H.2.3 Covariates for 2014 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	124	
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	7	Coll. (3,1) & (4,1); sing.
Rent/Housing × % Hispanic or Latino	5 × 3	8	8	All levels present.
Rent/Housing × % Owner-Occupied	5 × 3	8	8	All levels present.
% Owner-Occupied × % Black or African American	3 × 3	4	4	All levels present.
% Owner-Occupied × % Hispanic or Latino	3 × 3	4	4	All levels present.
Three-Factor Effects		12	0	
Pair Race/Ethnicity (4 Levels) × Pair Gender × Pair Age (3 Levels)	4 × 3 × 3	12	0	Drop all; conv.
Total		222	209	

Exhibit H.2.4 Covariates for 2014 NSDUH Pair Weights (res.pr.ev) Model Group 2: Midwest and West

This step used the same variables as the respondent pair-level poststratification step in [Exhibit H.2.3](#).

This page intentionally left blank

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

This page intentionally left blank

Table I.1 2014 NSDUH Person Pair-Level Response Rates

Domain	Selected Pairs	Respondent Pairs	% Interview Response Rate¹
Total	26,844	18,229	61.34
<i>Pair Age Group</i>			
12-17, 12-17	3,070	2,415	78.91
12-17, 18-25	2,443	1,832	75.74
12-17, 26-34	1,297	951	72.30
12-17, 35-49	5,530	3,932	72.67
12-17, 50+	1,132	771	65.97
18-25, 18-25	3,743	2,625	69.43
18-25, 26-34	1,378	859	59.54
18-25, 35-49	1,906	1,189	63.91
18-25, 50+	1,263	741	59.15
26-34, 26-34	1,356	855	63.40
26-34, 35-49	737	431	56.49
26-34, 50+	420	220	50.61
35-49, 35-49	1,160	667	56.14
35-49, 50+	498	253	50.44
50+, 50+	911	488	52.44
<i>Pair Race/Ethnicity</i>			
Hispanic or Latino	4,454	3,128	63.79
Black or African American	2,644	2,001	66.20
White	15,161	10,019	60.87
Other	2,075	1,308	44.99
White & Black or African American	266	193	65.52
White & Hispanic or Latino	970	664	65.98
White & Other	846	607	66.06
Black or African American & Hispanic or Latino	135	94	77.57
Black or African American & Other	125	94	74.16
Hispanic or Latino & Other	168	121	60.64
<i>Pair Gender</i>			
Male, Male	5,676	3,735	58.88
Female, Female	5,853	4,227	67.92
Male, Female	15,315	10,267	60.18
<i>Household Size</i>			
Two	6,790	4,431	58.76
Three	8,591	5,841	61.04
Four or More	11,463	7,957	62.76

Table I.1 2014 NSDUH Person Pair-Level Response Rates (continued)

Domain	Selected Pairs	Respondent Pairs	% Interview Response Rate¹
<i>Census Region</i>			
Northeast	5,225	3,335	53.92
South	8,744	6,101	62.50
Midwest	6,247	4,207	62.87
West	6,628	4,586	63.71
<i>Quarter</i>			
Quarter 1	6,141	4,254	62.78
Quarter 2	7,520	5,113	61.12
Quarter 3	6,588	4,455	61.12
Quarter 4	6,595	4,407	60.34
<i>% Hispanic or Latino in Segment</i>			
50-100%	2,535	1,799	61.72
10-<50%	6,914	4,664	59.00
<10%	17,395	11,766	62.45
<i>% Black or African American in Segment</i>			
50-100%	1,857	1,390	65.43
10-<50%	5,203	3,610	62.69
<10%	19,784	13,229	60.52
<i>% Owner-Occupied DUs in Segment</i>			
50-100%	20,074	13,482	61.07
10-<50%	5,353	3,722	62.65
<10%	1,417	1,025	59.45
<i>Combined Median Rent/Housing Value</i>			
1 st Quintile	4,261	3,051	63.39
2 nd Quintile	5,619	3,982	64.30
3 rd Quintile	6,565	4,517	61.42
4 th Quintile	5,836	3,855	59.18
5 th Quintile	4,563	2,824	59.49
<i>Population Density</i>			
Large MSA	12,179	8,000	58.69
Medium to Small MSA	12,659	8,833	64.54
Non-MSA, Urban	628	431	66.06
Non-MSA, Rural	1,378	965	68.29
<i>Group Quarters</i>			
Group	272	205	69.75
Non-Group	26,572	18,024	61.30

DU = dwelling unit, MSA = metropolitan statistical area.

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

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

This page intentionally left blank

Table J.1 2014 NSDUH Selected Pair-Level Proportions of Extreme Values and Outwinors

Domain	n	SDU-Level Weights ¹ (SDUWT: WT1*...*WT11)			Before sel.pr.ps ¹ (SDUWT*PRWT12)			After sel.pr.ps ¹ (SDUWT*PRWT12*PRWT13)		
		% Unweighted	% Weighted ²	% Outwinor ³	% Unweighted	% Weighted ²	% Outwinor ³	% Unweighted	% Weighted ²	% Outwinor ³
Total	26,844	1.99	5.29	1.41	4.15	20.50	9.96	1.29	5.59	1.18
<i>Pair Age Group</i>										
12-17, 12-17	3,070	1.14	2.64	0.52	2.96	13.11	3.66	0.23	0.84	0.06
12-17, 18-25	2,443	2.17	6.13	1.42	7.37	30.91	13.87	0.57	2.20	0.22
12-17, 26-34	1,297	2.47	6.13	1.95	1.77	8.43	2.03	0.46	2.46	0.44
12-17, 35-49	5,530	1.28	4.11	1.31	1.54	8.80	2.29	0.56	2.65	0.26
12-17, 50+	1,132	1.59	4.56	0.96	0.97	5.51	0.55	0.27	0.47	0.05
18-25, 18-25	3,743	2.48	5.38	1.17	7.88	27.95	11.43	1.79	4.90	0.46
18-25, 26-34	1,378	5.22	12.57	3.54	5.15	16.31	5.65	2.90	5.67	0.43
18-25, 35-49	1,906	2.78	7.32	2.19	7.45	25.27	10.58	2.73	4.38	0.39
18-25, 50+	1,263	1.35	3.88	1.12	2.93	15.64	7.81	1.43	2.76	0.30
26-34, 26-34	1,356	2.51	6.65	1.78	1.70	7.61	2.27	1.18	3.23	0.40
26-34, 35-49	737	2.58	8.00	2.58	2.71	17.94	7.51	3.39	6.19	0.72
26-34, 50+	420	0.24	0.48	0.04	1.90	16.23	10.30	0.00	0.00	0.00
35-49, 35-49	1,160	1.55	3.92	1.03	4.22	28.68	16.72	1.38	1.48	0.18
35-49, 50+	498	1.61	5.35	1.26	6.22	39.82	22.53	3.21	16.12	4.09
50+, 50+	911	0.99	2.05	0.44	5.38	25.69	14.40	3.84	13.17	3.75
<i>Pair Race/Ethnicity</i>										
Hispanic or Latino	4,454	3.12	8.79	2.99	4.58	25.07	12.12	1.82	6.49	1.10
Black or African American	2,644	3.63	8.08	1.89	6.32	23.14	9.68	2.08	6.73	1.18
White	15,161	0.66	1.61	0.27	3.28	17.64	8.61	0.64	4.63	1.13
Other	2,075	4.39	10.47	1.85	4.82	18.54	7.54	2.94	8.62	1.99
White & Black or African American	266	3.38	6.02	1.83	7.52	28.12	14.79	0.75	1.58	0.01
White & Hispanic or Latino	970	3.51	9.37	3.37	5.36	17.86	6.29	1.75	3.32	0.18
White & Other	846	3.07	6.82	1.39	4.49	13.13	3.83	2.13	6.57	0.77
Black or African American & Hispanic or Latino	135	15.56	39.31	15.03	11.85	74.57	64.15	2.96	19.83	6.37
Black or African American & Other	125	4.00	8.27	0.80	1.60	11.75	1.21	2.40	6.00	0.71
Hispanic or Latino & Other	168	7.14	18.40	5.51	10.71	23.96	4.83	4.76	8.84	1.57
<i>Pair Gender</i>										
Male, Male	5,676	2.34	6.23	1.64	5.78	19.07	7.75	1.81	6.63	1.11
Female, Female	5,853	1.78	4.70	1.20	4.37	22.33	11.09	1.14	3.72	0.64
Male, Female	15,315	1.93	5.15	1.40	3.47	20.34	10.23	1.15	5.82	1.35
<i>Household Size</i>										
Two	6,790	1.63	4.17	1.09	0.91	2.41	0.57	0.46	0.81	0.09
Three	8,591	1.78	4.75	1.25	1.66	20.10	11.35	1.08	5.10	1.04
Four or More	11,463	2.35	6.29	1.70	7.94	29.50	13.81	1.94	8.20	1.78

Table J.1 2014 NSDUH Selected Pair-Level Proportions of Extreme Values and Outwinors (continued)

Domain	n	SDU-Level Weights ¹ (SDUWT: WT1*...*WT11)			Before sel.pr.ps ¹ (SDUWT*PRWT12)			After sel.pr.ps ¹ (SDUWT*PRWT12*PRWT13)		
		% Unweighted	% Weighted ²	% Outwinors ³	% Unweighted	% Weighted ²	% Outwinors ³	% Unweighted	% Weighted ²	% Outwinors ³
<i>Census Region</i>										
Northeast	5,225	3.23	9.21	2.52	4.75	25.60	14.80	1.19	5.51	1.29
South	8,744	1.69	4.49	1.16	4.08	16.98	6.83	0.94	5.88	1.39
Midwest	6,247	1.31	3.05	0.86	4.24	20.05	9.10	1.50	6.08	0.93
West	6,628	2.02	5.70	1.49	3.70	21.85	11.31	1.63	4.85	0.98
<i>Quarter</i>										
Quarter 1	6,141	2.62	6.56	1.88	5.23	24.82	12.90	1.47	6.01	1.36
Quarter 2	7,520	1.52	4.06	0.93	3.64	15.97	5.61	1.18	5.61	1.00
Quarter 3	6,588	2.00	5.63	1.47	3.87	21.33	10.67	1.31	5.38	1.12
Quarter 4	6,595	1.91	4.93	1.37	4.02	19.77	10.60	1.23	5.36	1.23
<i>% Hispanic or Latino in Segment</i>										
50-100%	2,535	2.17	5.99	1.77	3.51	15.36	5.42	1.50	5.27	0.73
10-<50%	6,914	2.50	6.52	2.03	4.74	22.88	10.72	1.75	5.80	1.23
<10%	17,395	1.75	4.49	1.01	4.01	20.27	10.43	1.08	5.55	1.24
<i>% Black or African American in Segment</i>										
50-100%	1,857	4.36	11.51	3.58	6.84	21.95	7.99	1.94	6.58	1.06
10-<50%	5,203	2.42	6.19	1.72	4.90	24.85	13.49	1.48	6.08	1.24
<10%	19,784	1.65	4.33	1.08	3.71	19.09	9.12	1.18	5.34	1.17
<i>% Owner-Occupied DUs¹ in Segment</i>										
50-100%	20,074	1.59	4.43	1.15	3.84	19.37	9.42	1.28	5.71	1.24
10-<50%	5,353	2.73	6.36	1.71	4.60	22.31	10.90	1.63	5.61	1.01
<10%	1,417	4.73	11.83	3.50	6.92	32.03	15.19	0.14	0.05	0.01
<i>Combined Median Rent/Housing Value</i>										
1 st Quintile	4,261	1.60	4.44	1.26	3.61	15.59	6.45	1.15	4.51	0.65
2 nd Quintile	5,619	2.03	5.67	1.60	4.31	17.37	6.53	1.32	4.77	0.58
3 rd Quintile	6,565	1.75	4.50	1.21	3.76	17.48	7.55	1.14	4.95	1.15
4 th Quintile	5,836	2.43	6.09	1.65	4.73	23.57	11.77	1.35	6.13	1.32
5 th Quintile	4,563	2.06	5.54	1.28	4.30	26.23	15.86	1.51	7.26	1.98
<i>Population Density</i>										
Large MSA ¹	12,179	2.62	6.41	1.74	4.81	24.11	12.19	1.52	6.83	1.62
Medium to Small MSA ¹	12,659	1.53	4.00	1.01	3.60	16.12	7.26	1.13	4.25	0.65
Non-MSA, ¹ Urban	628	0.64	2.26	0.50	3.98	10.55	2.80	0.32	0.33	0.03
Non-MSA, ¹ Rural	1,378	1.16	2.85	0.85	3.48	10.02	3.86	1.16	2.15	0.11
<i>Group Quarters</i>										
Group	272	4.04	6.59	1.63	14.34	39.48	19.22	2.94	3.70	0.96
Non-Group	26,572	1.96	5.28	1.41	4.05	20.43	9.93	1.27	5.60	1.18

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

Table J.2 2014 NSDUH Respondent Pair-Level Proportions of Extreme Values and Outwinsors

Domain	n	Before res.pr.nr ¹ (SDUWT*PRWT12*PRWT13)			After res.pr.nr ¹ (SDUWT*PRWT12*...*PRWT14)		
		% Unweighted	% Weighted ²	% Outwinsor ³	% Unweighted	% Weighted ²	% Outwinsor ³
Total	18,229	1.63	7.80	1.57	2.08	10.37	1.97
<i>Pair Age Group</i>							
12-17, 12-17	2,415	0.29	1.43	0.08	0.25	1.57	0.28
12-17, 18-25	1,832	0.93	3.42	0.34	0.27	0.96	0.16
12-17, 26-34	951	0.42	2.22	0.12	0.42	1.89	0.17
12-17, 35-49	3,932	0.79	3.40	0.45	0.36	1.60	0.25
12-17, 50+	771	0.13	0.23	0.00	0.13	0.23	0.03
18-25, 18-25	2,625	2.25	5.40	0.51	3.54	9.22	1.50
18-25, 26-34	859	3.49	7.03	0.86	7.33	18.41	4.63
18-25, 35-49	1,189	3.62	6.10	0.55	4.79	9.82	1.87
18-25, 50+	741	1.75	5.27	0.30	2.56	6.89	1.54
26-34, 26-34	855	1.99	4.70	0.63	1.87	6.66	1.95
26-34, 35-49	431	4.41	10.02	1.63	4.64	7.64	1.32
26-34, 50+	220	0.45	2.48	0.37	2.27	9.15	1.73
35-49, 35-49	667	3.30	14.79	3.12	4.95	13.98	3.09
35-49, 50+	253	5.93	27.37	6.25	7.51	28.90	4.24
50+, 50+	488	3.89	13.63	4.20	5.12	18.06	3.44
<i>Pair Race/Ethnicity</i>							
Hispanic or Latino	3,128	2.30	11.38	2.08	1.92	7.80	1.35
Black or African American	2,001	2.20	7.25	1.15	2.15	13.74	2.85
White	10,019	0.88	6.24	1.51	1.41	8.24	1.24
Other	1,308	3.21	8.40	1.72	7.11	32.60	8.62
White & Black or African American	193	1.55	4.78	0.16	2.07	2.64	1.11
White & Hispanic or Latino	664	1.96	6.62	0.44	2.86	7.38	1.41
White & Other	607	2.47	6.85	0.78	2.31	5.85	0.34
Black or African American & Hispanic or Latino	94	10.64	48.15	9.67	0.00	0.00	0.00
Black or African American & Other	94	2.13	1.27	0.19	1.06	0.31	0.01
Hispanic or Latino & Other	121	7.44	13.76	1.78	4.13	10.75	3.27
<i>Pair Gender</i>							
Male, Male	3,735	2.12	7.21	0.80	3.45	11.46	2.26
Female, Female	4,227	1.68	6.45	1.01	1.35	6.31	1.06
Male, Female	10,267	1.44	8.39	1.95	1.89	11.21	2.15
<i>Household Size</i>							
Two	4,431	0.50	1.07	0.15	1.08	3.75	0.86
Three	5,841	1.20	9.23	2.08	1.68	12.38	1.98
Four or More	7,957	2.59	10.20	1.96	2.94	12.61	2.52

Table J.2 2014 NSDUH Respondent Pair-Level Proportions of Extreme Values and Outwinsors (continued)

Domain	n	Before res.pr.nr ¹ (SDUWT*PRWT12*PRWT13)			After res.pr.nr ¹ (SDUWT*PRWT12*...*PRWT14)		
		% Unweighted	% Weighted ²	% Outwinsor ³	% Unweighted	% Weighted ²	% Outwinsor ³
<i>Census Region</i>							
Northeast	3,335	1.74	6.88	1.46	2.85	16.79	3.77
South	6,101	1.03	7.47	1.59	2.03	10.83	2.04
Midwest	4,207	1.78	7.26	1.38	1.90	6.36	0.84
West	4,586	2.22	9.23	1.74	1.77	8.34	1.48
<i>Quarter</i>							
Quarter 1	4,254	2.33	9.01	1.78	2.68	11.74	2.19
Quarter 2	5,113	1.21	5.24	0.79	1.56	7.11	1.57
Quarter 3	4,455	1.50	8.25	1.99	2.22	11.89	2.29
Quarter 4	4,407	1.59	8.75	1.73	1.97	10.83	1.85
<i>% Hispanic or Latino in Segment</i>							
50-100%	1,799	1.72	8.34	0.99	2.28	7.58	1.51
10-<50%	4,664	2.34	8.98	1.83	3.17	14.66	2.92
<10%	11,766	1.34	7.13	1.55	1.62	8.76	1.58
<i>% Black or African American in Segment</i>							
50-100%	1,390	2.16	6.70	0.81	1.73	13.95	3.59
10-<50%	3,610	1.69	8.91	1.79	2.47	12.12	2.25
<10%	13,229	1.56	7.59	1.58	2.02	9.49	1.72
<i>% Owner-Occupied DUs¹ in Segment</i>							
50-100%	13,482	1.60	7.58	1.60	1.84	10.43	1.98
10-<50%	3,722	2.20	9.49	1.59	3.25	10.55	1.98
<10%	1,025	0.00	0.00	0.00	1.07	6.29	1.66
<i>Combined Median Rent/Housing Value</i>							
1 st Quintile	3,051	1.38	7.77	0.86	1.28	8.05	1.70
2 nd Quintile	3,982	1.61	6.04	0.80	1.53	8.59	1.90
3 rd Quintile	4,517	1.37	5.41	0.92	1.90	6.56	1.38
4 th Quintile	3,855	1.82	8.73	1.84	2.80	13.49	2.50
5 th Quintile	2,824	2.12	11.59	3.36	3.05	14.62	2.31
<i>Population Density</i>							
Large MSA ¹	8,000	1.96	9.41	2.21	2.80	13.29	2.63
Medium to Small MSA ¹	8,833	1.38	5.93	0.85	1.62	7.06	1.18
Non-MSA, ¹ Urban	431	0.93	2.25	0.21	0.93	0.96	0.12
Non-MSA, ¹ Rural	965	1.55	6.61	0.43	0.93	2.50	0.61
<i>Group Quarters</i>							
Group	205	6.34	10.22	1.45	8.29	13.54	2.45
Non-Group	18,024	1.58	7.79	1.57	2.01	10.36	1.97

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: $100 * \frac{\sum_k w_{ek}}{\sum_k w_k}$, 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: $100 * \frac{\sum_k (w_{ek} - b_k)}{\sum_k w_k}$, where b_k denotes the winsorized weight.

Table J.3 2014 NSDUH Respondent Pair-Level Proportions of Extreme Values and Outwinors

Domain	n	Before res.pr.ps ¹ (SDUWT*PRWT12*...*PRWT14)			After res.pr.ps ¹ (SDUWT*PRWT12*...*PRWT15)			Final Weight: After res.pr.ev ¹ (SDUWT*PRWT12*...*PRWT16)		
		% Unweighted	% Weighted ²	% Outwinors ³	% Unweighted	% Weighted ²	% Outwinors ³	% Unweighted	% Weighted ²	% Outwinors ³
Total	18,229	2.01	6.95	1.63	1.59	5.97	0.74	0.44	2.64	0.20
<i>Pair Age Group</i>										
12-17, 12-17	2,407	0.25	1.57	0.28	0.21	1.61	0.35	0.08	0.74	0.10
12-17, 18-25	1,832	0.44	1.60	0.29	0.44	2.43	0.66	0.27	1.40	0.14
12-17, 26-34	941	0.53	4.29	2.49	0.53	1.19	0.07	0.00	0.00	0.00
12-17, 35-49	3,940	0.43	1.79	0.28	0.41	1.27	0.30	0.25	0.73	0.12
12-17, 50+	776	0.26	0.74	0.05	0.26	1.06	0.30	0.39	4.36	0.35
18-25, 18-25	2,585	3.64	9.43	1.54	2.90	7.10	0.93	0.46	2.07	0.24
18-25, 26-34	870	6.67	17.18	5.23	5.17	16.28	2.35	1.72	6.59	0.69
18-25, 35-49	1,180	5.51	12.14	2.57	3.64	8.73	0.95	0.59	1.31	0.10
18-25, 50+	750	2.93	7.54	1.79	2.53	5.13	0.53	0.67	1.08	0.09
26-34, 26-34	865	1.97	6.84	2.06	1.73	4.23	0.83	0.35	1.49	0.16
26-34, 35-49	442	4.98	8.80	1.66	2.04	3.41	0.17	0.45	0.62	0.00
26-34, 50+	235	1.28	3.87	0.37	1.28	4.09	0.24	0.43	2.58	0.28
35-49, 35-49	658	4.56	10.07	2.72	3.65	10.05	1.30	0.91	3.74	0.23
35-49, 50+	262	2.67	7.24	1.28	3.82	11.65	1.17	1.15	3.61	0.09
50+, 50+	486	2.26	9.07	2.36	2.06	7.44	0.85	1.44	5.11	0.32
<i>Pair Race/Ethnicity</i>										
Hispanic or Latino	3,170	1.86	5.92	1.09	0.98	3.22	0.20	0.13	0.41	0.06
Black or African American	1,928	2.02	8.53	2.45	1.30	4.05	0.52	0.26	1.71	0.19
White	9,683	1.34	3.98	0.69	1.09	5.07	0.58	0.15	2.12	0.08
Other	1,273	7.31	31.37	9.25	8.01	23.88	4.01	4.08	14.36	1.51
White & Black or African American	186	2.15	2.75	1.18	3.23	21.31	1.52	1.08	13.51	0.93
White & Hispanic or Latino	705	2.70	6.29	1.17	1.99	4.03	0.32	0.14	0.21	0.01
White & Other	793	1.89	3.68	0.49	0.38	1.06	0.09	0.13	0.18	0.00
Black or African American & Hispanic or Latino	127	0.79	1.56	0.59	0.79	1.41	0.28	0.00	0.00	0.00
Black or African American & Other	191	0.52	1.32	0.16	0.00	0.00	0.00	0.00	0.00	0.00
Hispanic or Latino & Other	173	3.47	10.20	3.77	0.58	0.10	0.01	0.58	0.09	0.00
<i>Pair Gender</i>										
Male, Male	3,722	3.44	10.31	2.35	2.47	8.22	1.13	0.54	2.53	0.30
Female, Female	4,219	1.30	5.05	1.00	1.35	4.57	0.64	0.45	2.95	0.25
Male, Female	10,288	1.79	6.55	1.61	1.36	5.74	0.66	0.41	2.58	0.16
<i>Household Size</i>										
Two	4,431	1.13	3.86	0.91	0.88	2.19	0.23	0.23	0.68	0.07
Three	5,841	1.49	7.96	1.69	1.16	6.51	0.82	0.43	3.93	0.26
Four or More	7,957	2.89	7.97	1.96	2.29	7.57	0.95	0.58	2.95	0.23

Table J.3 2014 NSDUH Respondent Pair-Level Proportions of Extreme Values and Outwinors (continued)

Domain	n	Before res.pr.ps ¹ (SDUWT*PRWT12*...*PRWT14)			After res.pr.ps ¹ (SDUWT*PRWT12*...*PRWT15)			Final Weight: After res.pr.ev ¹ (SDUWT*PRWT12*...*PRWT16)		
		% Unweighted	% Weighted ²	% Outwinor ³	% Unweighted	% Weighted ²	% Outwinor ³	% Unweighted	% Weighted ²	% Outwinor ³
<i>Census Region</i>										
Northeast	3,335	2.79	12.93	3.15	2.25	13.04	2.18	1.41	9.87	0.79
South	6,101	1.84	6.16	1.44	1.07	4.93	0.62	0.41	2.23	0.16
Midwest	4,207	1.95	4.39	1.06	1.66	2.94	0.23	0.17	0.21	0.00
West	4,586	1.74	5.87	1.28	1.72	4.84	0.29	0.04	0.02	0.00
<i>Quarter</i>										
Quarter 1	4,254	2.54	8.05	2.08	1.88	5.21	0.59	0.47	2.66	0.23
Quarter 2	5,113	1.56	4.57	1.18	1.27	3.92	0.59	0.33	1.49	0.11
Quarter 3	4,455	2.11	8.59	1.93	1.64	7.42	0.82	0.58	3.54	0.26
Quarter 4	4,407	1.93	6.66	1.35	1.61	7.37	0.96	0.41	2.89	0.20
<i>% Hispanic or Latino in Segment</i>										
50-100%	1,799	2.11	6.46	1.35	1.06	2.78	0.26	0.11	0.43	0.05
10-<50%	4,664	3.02	10.28	2.58	2.19	7.72	0.96	0.73	3.54	0.32
<10%	11,766	1.60	5.36	1.21	1.43	5.74	0.72	0.38	2.63	0.17
<i>% Black or African American in Segment</i>										
50-100%	1,390	1.65	9.48	3.18	1.22	6.67	0.97	0.50	3.38	0.41
10-<50%	3,610	2.27	7.87	1.83	1.63	6.66	0.76	0.61	2.93	0.18
<10%	13,229	1.98	6.43	1.41	1.61	5.70	0.71	0.39	2.48	0.18
<i>% Owner-Occupied DUs¹ in Segment</i>										
50-100%	13,482	1.77	6.84	1.65	1.51	6.26	0.81	0.46	2.94	0.21
10-<50%	3,722	3.14	7.53	1.51	1.93	4.69	0.39	0.35	1.23	0.13
<10%	1,025	1.07	6.29	1.84	1.27	6.68	1.16	0.59	3.78	0.47
<i>Combined Median Rent/Housing Value</i>										
1 st Quintile	3,051	1.34	5.15	1.59	1.74	5.63	0.88	0.52	2.59	0.11
2 nd Quintile	3,982	1.38	5.72	1.30	1.08	5.16	0.51	0.20	1.19	0.13
3 rd Quintile	4,517	1.95	5.30	1.20	1.37	4.53	0.50	0.29	1.43	0.13
4 th Quintile	3,855	2.67	9.80	2.41	1.82	6.40	0.90	0.60	3.44	0.32
5 th Quintile	2,824	2.83	8.00	1.58	2.16	8.25	0.97	0.74	4.61	0.27
<i>Population Density</i>										
Large MSA ¹	8,000	2.66	8.97	2.12	1.80	7.02	0.90	0.73	3.54	0.27
Medium to Small MSA ¹	8,833	1.60	4.54	1.06	1.37	4.60	0.53	0.20	1.46	0.12
Non-MSA, ¹ Urban	431	0.93	1.26	0.08	1.16	4.51	0.21	0.00	0.00	0.00
Non-MSA, ¹ Rural	965	0.93	2.50	0.46	1.97	4.31	0.60	0.52	1.77	0.02
<i>Group Quarters</i>										
Group	205	8.29	13.54	2.45	7.32	15.83	1.95	1.46	4.62	0.18
Non-Group	18,024	1.94	6.93	1.63	1.52	5.94	0.73	0.43	2.63	0.20

Table J.3 2014 NSDUH Respondent Pair-Level Proportions of Extreme Values and Outwinsors (continued)

Domain	n	Before res.pr.ps ¹ (SDUWT*PRWT12*...*PRWT14)			After res.pr.ps ¹ (SDUWT*PRWT12*...*PRWT15)			Final Weight: After res.pr.ev ¹ (SDUWT*PRWT12*...*PRWT16)		
		% Unweighted	% Weighted ²	% Outwinsor ³	% Unweighted	% Weighted ²	% Outwinsor ³	% Unweighted	% Weighted ²	% Outwinsor ³
<i>Pair Relationship Domain⁴</i>										
Parent-Child (12-14)	2,813	0.43	2.52	0.91	0.28	1.03	0.23	0.18	1.66	0.20
Parent-Child (12-17)	5,178	0.41	1.81	0.53	0.42	1.05	0.22	0.23	1.55	0.16
Parent-Child (12-20)	6,040	0.88	3.44	0.82	0.83	2.99	0.35	0.25	1.34	0.12
Sibling (12-14)-Sibling (15-17)	1,418	0.21	1.31	0.28	0.28	2.25	0.56	0.14	1.21	0.17
Sibling (12-17)-Sibling (18-25)	1,671	0.48	1.75	0.31	0.48	2.65	0.72	0.30	1.53	0.16
Spouse-Spouse/Partner-Partner	3,566	1.79	6.26	1.57	1.40	4.92	0.52	0.34	1.64	0.05
Spouse-Spouse/Partner-Partner with Children (Younger Than 18)	1,796	2.06	6.02	1.55	2.51	11.02	1.25	0.67	4.00	0.13

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: $100 * \frac{\sum_k w_{ek}}{\sum_k w_k}$, 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: $100 * \frac{\sum_k (w_{ek} - b_k)}{\sum_k w_k}$, where b_k denotes the winsorized weight.

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

This page intentionally left blank

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

This page intentionally left blank

Table K.1 2014 NSDUH Respondent Pair-Level Slippage Rates

Domain	<i>n</i>	Initial Total (<i>I</i>)¹	Final Total (<i>F</i>)²	Control Total from SDU (<i>C</i>)	(<i>I</i> - <i>C</i>)/<i>C</i>%	(<i>F</i> - <i>C</i>)/<i>C</i>%
Total	18,229	226,491,634	226,491,634	226,491,634	0.00	0.00
<i>Pair Age Group</i>						
12-17, 12-17	2,407	7,248,958	7,250,253	7,250,253	-0.02	-0.00
12-17, 18-25	1,832	8,585,195	8,544,102	8,544,102	0.48	-0.00
12-17, 26-34	941	5,318,050	5,213,991	5,213,991	2.00	0.00
12-17, 35-49	3,940	29,326,152	29,000,052	29,000,052	1.12	0.00
12-17, 50+	776	12,224,436	12,387,714	12,387,714	-1.32	0.00
18-25, 18-25	2,585	12,536,548	12,661,498	12,661,498	-0.99	-0.00
18-25, 26-34	870	7,437,562	7,375,281	7,375,281	0.84	-0.00
18-25, 35-49	1,180	17,272,856	17,484,908	17,484,908	-1.21	-0.00
18-25, 50+	750	19,494,448	19,802,773	19,802,773	-1.56	-0.00
26-34, 26-34	865	10,792,194	10,932,453	10,932,453	-1.28	0.00
26-34, 35-49	442	8,610,431	8,571,921	8,571,921	0.45	0.00
26-34, 50+	235	13,356,220	12,537,407	12,537,407	6.53	0.00
35-49, 35-49	658	18,132,311	18,091,428	18,091,428	0.23	-0.00
35-49, 50+	262	16,518,566	16,738,461	16,738,461	-1.31	0.00
50+, 50+	486	39,637,707	39,899,392	39,899,392	-0.66	0.00
<i>Pair Race/Ethnicity</i>						
Hispanic or Latino	3,170	40,100,006	39,738,046	39,738,046	0.91	0.00
Black or African American	1,928	25,172,788	25,350,552	25,350,552	-0.70	0.00
White	9,683	119,827,457	123,892,916	123,892,916	-3.28	0.00
Other	1,273	16,816,802	16,919,733	16,919,733	-0.61	0.00
White & Black or African American	186	2,105,062	2,190,967	2,190,967	-3.92	0.00
White & Hispanic or Latino	705	9,159,337	8,577,043	8,577,043	6.79	0.00
White & Other	793	8,158,997	6,121,889	6,121,889	33.28	0.00
Black or African American & Hispanic or Latino	127	1,531,253	1,488,049	1,488,049	2.90	0.00
Black or African American & Other	191	1,591,415	925,895	925,895	71.88	0.00
Hispanic or Latino & Other	173	2,028,517	1,286,543	1,286,543	57.67	0.00
<i>Pair Gender</i>						
Male, Male	3,722	40,395,852	40,627,453	40,627,453	-0.57	0.00
Female, Female	4,219	40,479,827	40,648,289	40,648,289	-0.41	0.00
Male, Female	10,288	145,615,955	145,215,892	145,215,892	0.28	0.00
<i>Pair Relationship Domain</i>^{3,4,5}						
Parent-Child (12-14)*	2,813	11,405,656	12,534,068	12,534,068	-9.00	0.00
Parent-Child (12-17)*	5,178	23,764,830	25,182,498	25,182,498	-5.63	0.00
Parent-Child (15-17)*	2,365	12,359,174	12,648,430	12,648,430	-2.29	0.00
Parent-Child (12-20)*	6,040	32,702,707	34,607,631	34,607,631	-5.50	0.00
Parent*-Child (12-14)	2,813	17,734,367	19,340,763	19,340,763	-8.31	0.00
Parent*-Child (12-17)	5,178	30,294,926	32,424,768	32,424,768	-6.57	0.00
Parent*-Child (15-17)	2,365	18,853,477	19,192,491	19,320,073	-2.42	-0.66
Parent*-Child (12-20)	6,040	38,585,284	40,493,782	40,493,782	-4.71	0.00
Sibling (12-14)-Sibling (15-17)*	1,418	3,870,135	4,044,850	4,044,850	-4.32	-0.00
Sibling (12-17)-Sibling (18-25)*	1,671	6,393,975	6,442,409	6,442,409	-0.75	-0.00
Spouse-Spouse/Partner-Partner	3,566	73,487,730	72,518,420	72,518,420	1.34	-0.00
Spouse-Spouse/Partner-Partner with Children (Younger Than 18)	1,796	25,905,213	29,710,120	29,710,120	-12.81	-0.00

Table K.1 2014 NSDUH Respondent Pair-Level Slippage Rates (continued)

Domain	<i>n</i>	Initial Total (<i>I</i>)¹	Final Total (<i>F</i>)²	Control Total from SDU (<i>C</i>)	(<i>I</i> - <i>C</i>)/<i>C</i>%	(<i>F</i> - <i>C</i>)/<i>C</i>%
<i>Household Size</i>						
Two	4,431	55,696,048	55,696,048	55,696,048	-0.00	0.00
Three	5,841	57,834,499	57,834,499	57,834,499	-0.00	0.00
Four or More	7,957	112,961,087	112,961,087	112,961,087	0.00	0.00
<i>Census Region</i>						
Northeast	3,335	40,920,183	40,920,183	40,920,183	-0.00	0.00
South	6,101	82,323,153	82,323,153	82,323,153	0.00	0.00
Midwest	4,207	45,244,345	45,244,345	45,244,345	-0.00	-0.00
West	4,586	58,003,952	58,003,952	58,003,952	0.00	-0.00
<i>Quarter</i>						
Quarter 1	4,254	55,741,969	55,741,969	55,741,969	-0.00	0.00
Quarter 2	5,113	57,662,172	57,662,172	57,662,172	0.00	0.00
Quarter 3	4,455	56,844,052	56,844,052	56,844,052	-0.00	0.00
Quarter 4	4,407	56,243,442	56,243,442	56,243,442	-0.00	0.00
<i>% Hispanic or Latino in Segment</i>						
50-100%	1,799	26,937,885	26,937,885	26,937,885	-0.00	-0.00
10-<50%	4,664	67,423,975	67,423,975	67,423,975	0.00	0.00
<10%	11,766	132,129,773	132,129,773	132,129,773	-0.00	0.00
<i>% Black or African American in Segment</i>						
50-100%	1,390	16,705,860	16,705,860	16,705,860	0.00	-0.00
10-<50%	3,610	47,696,643	47,696,643	47,696,643	-0.00	0.00
<10%	13,229	162,089,131	162,089,131	162,089,131	-0.00	0.00
<i>% Owner-Occupied DUs in Segment</i>						
50-100%	13,482	180,344,901	180,344,901	180,344,901	-0.00	0.00
10-<50%	3,722	41,920,064	41,920,064	41,920,064	0.00	0.00
<10%	1,025	4,226,670	4,226,670	4,226,670	-0.00	-0.00
<i>Combined Median Rent/Housing Value</i>						
1 st Quintile	3,051	29,901,019	29,901,019	29,901,019	-0.00	0.00
2 nd Quintile	3,982	44,272,086	44,272,086	44,272,086	0.00	0.00
3 rd Quintile	4,517	54,241,866	54,241,866	54,241,866	0.00	0.00
4 th Quintile	3,855	53,296,182	53,296,182	53,296,182	0.00	0.00
5 th Quintile	2,824	44,780,481	44,780,481	44,780,481	-0.00	0.00
<i>Population Density</i>						
Large MSA	8,000	130,147,419	130,147,419	130,147,419	0.00	0.00
Medium to Small MSA	8,833	84,429,828	84,429,828	84,429,828	0.00	0.00
Non-MSA, Urban	431	3,944,947	3,944,947	3,944,947	0.00	0.00
Non-MSA, Rural	965	7,969,441	7,969,441	7,969,441	0.00	0.00
<i>Group Quarters</i>						
Group	205	835,845	835,845	835,845	-0.00	0.00
Non-Group	18,024	225,655,789	225,655,789	225,655,789	0.00	0.00

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

¹ WT1*...*WT11*PRWT12*...*PRWT14 (before person pair poststratification).

² WT1*...*WT11*PRWT12*...*PRWT15 (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**

This page intentionally left blank

Table L.1 2014 NSDUH Selected Pair-Level Weight Summary Statistics

Domain	n	SDU-Level Weights ¹ (SDUWT: WT1*...*WT11)						Before sel.pr.ps ¹ (SDUWT*PRWT12)					After sel.pr.ps ¹ (SDUWT*PRWT12*PRWT13)						
		Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³
Total	26,844	14	565	989	1,386	8,241	1.42	21	1,820	3,807	8,444	2,040,767	10.14	32	1,771	3,887	8,691	243,331	4.19
<i>Pair Age Group</i>																			
12-17, 12-17	3,070	14	433	896	1,249	5,321	1.41	21	887	1,665	2,836	40,664	2.18	34	862	1,693	3,066	19,610	1.91
12-17, 18-25	2,443	16	627	1,036	1,439	6,287	1.42	46	1,266	2,134	3,959	176,016	3.70	59	1,358	2,470	4,747	23,287	1.83
12-17, 26-34	1,297	43	558	958	1,350	5,572	1.39	144	1,808	2,984	4,529	53,859	2.29	49	1,414	2,731	4,669	52,762	2.38
12-17, 35-49	5,530	17	491	917	1,279	7,182	1.41	82	2,021	3,553	6,019	84,106	2.36	87	1,712	3,387	6,299	76,465	2.30
12-17, 50+	1,132	23	588	1,048	1,465	5,683	1.42	191	4,881	8,643	13,103	88,126	1.89	269	4,139	8,008	13,292	76,006	1.94
18-25, 18-25	3,743	34	584	1,053	1,479	7,583	1.43	96	1,320	2,268	3,719	64,762	2.40	32	1,220	2,375	4,479	21,729	1.87
18-25, 26-34	1,378	50	663	1,081	1,529	6,749	1.45	237	2,350	3,979	6,171	79,723	2.31	154	1,943	3,629	6,388	56,228	2.29
18-25, 35-49	1,906	33	680	1,038	1,423	6,812	1.38	220	3,267	5,623	9,771	241,294	3.17	208	3,205	6,023	11,200	58,565	2.05
18-25, 50+	1,263	36	742	1,151	1,554	6,379	1.37	509	7,033	11,106	16,783	454,863	3.00	474	6,711	11,939	19,701	91,564	1.76
26-34, 26-34	1,356	51	537	961	1,350	8,241	1.45	298	3,454	6,141	9,654	255,073	2.69	402	3,209	6,049	9,593	108,940	2.23
26-34, 35-49	737	60	606	960	1,302	6,237	1.47	597	4,733	7,964	11,503	303,196	4.38	456	4,251	8,180	13,274	209,865	2.86
26-34, 50+	420	65	704	1,127	1,549	3,303	1.29	1,894	12,956	21,966	33,124	805,915	4.64	1,493	12,294	22,952	39,627	171,294	1.67
35-49, 35-49	1,160	17	506	924	1,314	4,695	1.38	453	4,536	9,296	13,528	644,951	8.15	458	4,333	8,963	15,385	209,566	3.34
35-49, 50+	498	47	667	1,012	1,420	4,612	1.40	1,423	12,946	21,157	30,451	1,203,720	6.73	1,469	13,238	24,032	37,810	222,239	2.20
50+, 50+	911	31	593	1,027	1,486	5,906	1.35	1,282	18,280	35,886	48,413	2,040,767	4.46	2,012	20,447	37,188	54,143	243,331	1.68
<i>Pair Race/Ethnicity</i>																			
Hispanic or Latino	4,454	14	706	1,143	1,557	7,583	1.39	21	2,302	4,373	8,932	840,885	9.19	32	2,048	4,373	9,650	223,650	3.91
Black or African American	2,644	43	829	1,137	1,492	7,083	1.33	98	2,215	4,269	9,025	608,925	5.92	34	2,145	4,647	9,719	222,239	3.82
White	15,161	46	522	959	1,266	5,003	1.36	67	1,728	3,621	8,256	1,203,720	8.89	47	1,730	3,668	8,255	243,331	4.39
Other	2,075	37	240	773	1,597	4,836	1.69	70	1,171	3,194	7,871	439,880	5.67	49	1,178	3,256	8,457	224,084	4.57
White & Black or African American	266	56	734	1,065	1,473	4,243	1.36	197	2,406	4,248	9,817	303,196	7.09	86	2,013	3,986	9,732	108,857	3.18
White & Hispanic or Latino	970	44	625	1,010	1,518	8,241	1.47	113	2,125	4,194	9,552	223,323	3.86	57	1,830	4,262	9,665	152,449	3.56
White & Other	846	45	305	786	1,303	5,683	1.62	123	1,504	3,387	7,913	167,607	3.54	80	1,425	3,554	7,775	119,383	3.46
Black or African American & Hispanic or Latino	135	48	663	1,254	1,957	6,618	1.71	109	2,201	4,511	10,666	2,040,767	46.33	154	2,133	4,309	10,502	226,138	5.90
Black or African American & Other	125	41	433	766	1,155	3,396	1.55	241	1,646	3,000	5,134	44,841	2.96	127	2,098	3,884	6,709	67,694	3.17
Hispanic or Latino & Other	168	31	282	803	1,353	5,813	1.90	207	1,313	3,647	7,432	75,986	3.13	143	1,441	3,856	9,573	68,015	3.01

Table L.1 2014 NSDUH Selected Pair-Level Weight Summary Statistics (continued)

Domain	n	SDU-Level Weights ¹ (SDUWT: WT1*...*WT11)						Before sel.pr.ps ¹ (SDUWT*PRWT12)						After sel.pr.ps ¹ (SDUWT*PRWT12*PRWT13)					
		Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³
<i>Pair Gender</i>																			
Male, Male	5,676	16	570	1,010	1,416	7,182	1.45	46	1,732	3,546	7,816	454,863	4.76	34	1,640	3,633	8,005	220,512	3.65
Female, Female	5,853	33	564	979	1,382	6,276	1.40	70	1,723	3,505	7,205	1,032,769	11.87	32	1,659	3,530	7,415	222,239	3.85
Male, Female	15,315	14	564	985	1,379	8,241	1.41	21	1,898	4,060	9,147	2,040,767	10.56	38	1,904	4,162	9,530	243,331	4.26
<i>Household Size</i>																			
Two	6,790	31	545	957	1,339	6,409	1.40	96	1,902	4,065	9,075	157,949	2.93	32	1,676	3,643	8,707	152,449	3.27
Three	8,591	17	578	979	1,353	7,182	1.40	67	1,593	3,252	5,957	1,203,720	14.24	34	1,522	3,267	6,335	221,116	5.03
Four or More	11,463	14	567	1,017	1,438	8,241	1.44	21	1,918	4,391	9,953	2,040,767	11.13	38	2,069	4,799	10,419	243,331	4.10
<i>Census Region</i>																			
Northeast	5,225	33	342	921	1,164	6,078	1.46	82	1,586	3,323	7,928	2,040,767	20.81	49	1,556	3,410	8,197	226,138	4.33
South	8,744	14	793	1,143	1,524	7,583	1.32	21	2,345	4,362	9,298	608,925	5.07	34	2,245	4,537	9,611	223,650	4.02
Midwest	6,247	38	647	925	1,181	6,276	1.28	103	1,758	3,415	7,297	805,915	7.66	32	1,812	3,524	7,398	243,331	3.98
West	6,628	45	355	878	1,584	8,241	1.60	94	1,372	3,764	8,879	1,032,769	11.29	51	1,328	3,703	9,360	232,317	4.35
<i>Quarter</i>																			
Quarter1	6,141	26	626	1,073	1,503	8,241	1.42	56	1,971	4,114	8,930	840,885	9.47	32	1,978	4,235	9,229	232,317	4.10
Quarter2	7,520	36	513	897	1,251	6,749	1.41	70	1,675	3,500	7,634	351,409	4.76	38	1,617	3,518	7,785	224,084	4.27
Quarter3	6,588	14	537	1,001	1,403	6,379	1.44	21	1,763	3,812	8,799	2,040,767	15.20	34	1,727	3,954	9,031	243,331	4.13
Quarter4	6,595	16	611	1,014	1,415	6,812	1.39	46	1,898	3,913	8,555	1,203,720	10.10	60	1,853	3,962	8,867	222,239	4.23
<i>% Hispanic or Latino in Segment</i>																			
50-100%	2,535	16	857	1,334	1,625	7,583	1.28	49	2,769	5,045	10,520	395,578	4.03	60	2,417	5,306	11,467	223,650	3.45
10-<50%	6,914	14	732	1,137	1,618	8,241	1.36	21	2,309	4,547	9,961	1,032,769	8.60	38	2,212	4,763	10,539	232,317	3.82
<10%	17,395	16	406	915	1,236	6,409	1.44	46	1,574	3,328	7,561	2,040,767	12.31	32	1,587	3,414	7,681	243,331	4.50
<i>% Black or African American in Segment</i>																			
50-100%	1,857	14	780	1,084	1,427	6,276	1.39	21	2,168	4,261	9,025	277,744	3.77	34	2,011	4,467	9,575	220,345	3.58
10-<50%	5,203	16	779	1,082	1,472	8,241	1.32	49	2,260	4,221	8,903	2,040,767	16.45	60	2,172	4,434	9,530	226,138	3.89
<10%	19,784	35	452	952	1,354	7,583	1.45	70	1,660	3,645	8,290	1,203,720	8.67	32	1,656	3,682	8,439	243,331	4.35
<i>% Owner-Occupied DUs¹ in Segment</i>																			
50-100%	20,074	14	531	973	1,356	8,241	1.42	21	1,820	3,864	8,738	2,040,767	9.85	38	1,910	4,153	9,276	243,331	4.07
10-<50%	5,353	17	614	1,022	1,427	6,749	1.41	67	1,806	3,649	7,803	1,032,769	10.52	47	1,823	3,861	8,351	220,345	4.12
<10%	1,417	43	733	1,125	1,585	7,083	1.40	96	1,861	3,492	6,974	840,885	13.56	32	734	1,495	3,140	95,300	5.01
<i>Combined Median Rent/Housing Value</i>																			
1 st Quintile	4,261	36	366	850	1,221	6,640	1.49	94	1,411	3,022	6,792	395,578	6.32	51	1,353	2,948	6,711	198,823	4.84
2 nd Quintile	5,619	33	511	944	1,307	7,182	1.44	107	1,726	3,501	7,720	746,626	5.75	34	1,666	3,635	7,998	220,345	4.24
3 rd Quintile	6,565	26	549	993	1,401	7,583	1.41	56	1,827	3,779	8,330	840,885	6.95	32	1,729	3,812	8,546	224,084	4.06
4 th Quintile	5,836	14	637	1,048	1,474	8,241	1.40	21	1,985	4,151	9,477	1,203,720	10.34	38	1,992	4,314	9,806	232,317	3.99
5 th Quintile	4,563	16	747	1,094	1,490	7,083	1.35	49	2,167	4,680	9,768	2,040,767	16.40	51	2,283	4,842	10,212	243,331	4.00

Table L.1 2014 NSDUH Selected Pair-Level Weight Summary Statistics (continued)

Domain	n	SDU-Level Weights ¹ (SDUWT: WT1*... ² WT11)						Before sel.pr.ps ¹ (SDUWT*PRWT12)						After sel.pr.ps ¹ (SDUWT*PRWT12*PRWT13)					
		Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³
<i>Population Density</i>																			
Large MSA ¹	12,179	14	898	1,193	1,597	8,241	1.27	21	2,695	5,168	10,773	2,040,767	9.94	34	2,645	5,343	11,234	243,331	3.70
Medium to Small MSA ¹	12,659	37	336	805	1,176	6,640	1.51	70	1,368	2,978	6,494	1,203,720	8.82	32	1,345	2,987	6,736	213,433	4.52
Non-MSA, ¹ Urban	628	86	320	714	1,085	3,958	1.53	126	1,177	2,483	5,836	80,074	3.71	109	1,145	2,636	5,937	121,421	4.37
Non-MSA, ¹ Rural	1,378	37	184	503	990	4,721	1.73	82	859	1,991	4,975	121,773	4.69	67	942	2,279	5,559	131,150	4.80
<i>Group Quarters</i>																			
Group	272	46	282	651	1,423	3,027	1.74	103	1,080	2,073	3,003	34,501	2.71	32	1,067	2,133	4,636	24,743	1.85
Non-Group	26,572	14	571	991	1,386	8,241	1.42	21	1,828	3,841	8,486	2,040,767	10.11	34	1,784	3,906	8,766	243,331	4.18

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 2014 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.nr)

Domain	n	Before res.pr.nr ¹ (SDUWT*PRWT12*PRWT13)						After res.pr.nr ¹ (SDUWT*PRWT12*...*PRWT14)					
		Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³
Total	18,229	32	1,681	3,601	7,894	243,331	4.23	35	2,149	4,919	11,443	470,163	5.42
<i>Pair Age Group</i>													
12-17, 12-17	2,415	34	859	1,699	3,118	19,610	1.90	36	1,042	2,081	3,944	25,533	1.97
12-17, 18-25	1,832	59	1,309	2,488	4,835	23,287	1.84	60	1,667	3,281	6,319	24,760	1.83
12-17, 26-34	951	49	1,400	2,682	4,617	38,046	2.35	49	1,773	3,416	6,432	54,721	2.49
12-17, 35-49	3,932	102	1,712	3,389	6,295	76,465	2.38	133	2,183	4,475	8,672	95,474	2.49
12-17, 50+	771	269	3,825	7,763	13,020	69,891	1.91	286	5,388	11,150	18,714	137,805	2.15
18-25, 18-25	2,625	32	1,227	2,340	4,479	20,555	1.84	35	1,631	3,131	6,387	40,273	1.98
18-25, 26-34	859	260	1,927	3,438	6,107	51,614	2.39	263	2,618	5,169	9,431	157,568	3.05
18-25, 35-49	1,189	208	3,222	6,045	11,215	58,565	2.08	355	4,497	8,769	18,524	98,439	2.19
18-25, 50+	741	474	6,815	11,780	19,726	91,564	1.82	506	10,281	18,973	33,574	209,970	1.92
26-34, 26-34	855	437	3,287	6,059	9,588	108,112	2.25	470	4,657	7,956	13,981	206,031	3.07
26-34, 35-49	431	722	4,020	7,811	12,289	151,007	2.86	773	6,269	12,224	22,693	229,885	3.05
26-34, 50+	220	1,493	11,970	22,852	38,424	157,074	1.64	2,254	18,727	40,100	72,225	437,673	1.95
35-49, 35-49	667	458	4,156	8,722	15,431	209,566	3.29	593	6,051	12,854	27,975	380,775	3.29
35-49, 50+	253	1,639	13,011	22,571	34,650	222,239	2.33	2,574	20,425	41,041	73,786	407,167	2.22
50+, 50+	488	2,012	17,992	37,155	53,665	243,331	1.70	2,181	32,646	65,267	103,085	470,163	1.78
<i>Pair Race/Ethnicity</i>													
Hispanic or Latino	3,128	32	1,908	4,136	8,845	210,727	3.89	35	2,337	5,473	12,547	314,883	4.60
Black or African American	2,001	34	2,055	4,321	8,943	222,239	3.45	36	2,413	5,302	11,587	437,673	5.48
White	10,019	47	1,656	3,414	7,586	243,331	4.53	48	2,154	4,810	11,100	470,163	5.58
Other	1,308	49	985	2,598	6,339	206,306	4.52	49	1,228	3,716	10,542	455,189	7.21
White & Black or African American	193	86	1,921	3,748	8,648	108,857	3.29	96	2,469	4,682	12,148	202,675	4.23
White & Hispanic or Latino	664	57	1,828	4,035	8,635	128,397	3.62	60	2,457	5,496	12,886	189,201	4.02
White & Other	607	80	1,367	3,348	7,249	99,461	3.30	82	1,706	4,379	10,429	158,043	3.95
Black or African American & Hispanic or Latino	94	196	2,186	4,820	10,860	226,138	6.28	198	2,787	5,256	13,755	255,707	6.74
Black or African American & Other	94	127	1,784	4,028	6,854	67,694	2.95	235	2,263	4,487	8,896	121,400	4.05
Hispanic or Latino & Other	121	195	1,387	3,087	6,936	58,254	3.10	301	1,706	4,669	10,484	114,324	3.96
<i>Pair Gender</i>													
Male, Male	3,735	34	1,540	3,314	7,109	151,007	3.58	36	2,063	4,687	10,768	431,487	5.01
Female, Female	4,227	32	1,603	3,384	7,077	222,239	3.91	35	1,987	4,413	9,352	314,883	4.90
Male, Female	10,267	38	1,799	3,807	8,655	243,331	4.33	39	2,275	5,240	12,590	470,163	5.42
<i>Household Size</i>													
Two	4,431	32	1,574	3,205	7,793	110,137	3.39	35	1,941	4,350	11,268	357,571	4.51
Three	5,841	34	1,406	3,060	5,646	221,116	5.67	36	1,834	4,047	8,151	411,221	7.33
Four or More	7,957	53	1,963	4,540	9,705	243,331	3.88	56	2,592	6,216	14,140	470,163	4.93

Table L.2 2014 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.nr) (continued)

Domain	n	Before res.pr.nr ¹ (SDUWT*PRWT12*PRWT13)						After res.pr.nr ¹ (SDUWT*PRWT12*...*PRWT14)					
		Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³
Census Region													
Northeast	3,335	49	1,425	2,969	7,042	226,138	4.33	49	1,818	4,204	11,325	470,163	6.58
South	6,101	34	2,140	4,271	8,836	222,239	3.85	36	2,609	5,552	12,216	407,167	5.12
Midwest	4,207	32	1,705	3,254	6,805	243,331	4.19	35	2,326	4,732	10,127	305,400	4.51
West	4,586	51	1,280	3,486	8,572	232,317	4.58	53	1,616	4,676	12,041	455,189	5.54
Quarter													
Quarter1	4,254	32	1,839	3,892	8,453	232,317	4.11	35	2,306	5,267	11,906	431,487	5.23
Quarter2	5,113	38	1,520	3,228	7,182	178,076	4.08	39	1,937	4,413	10,176	455,189	5.59
Quarter3	4,455	34	1,630	3,603	8,066	243,331	4.42	36	2,095	4,970	11,932	470,163	5.45
Quarter4	4,407	60	1,772	3,713	8,171	222,239	4.23	74	2,358	5,145	11,942	437,673	5.36
% Hispanic or Latino in Segment													
50-100%	1,799	60	2,274	4,776	10,121	156,918	3.37	74	2,813	6,303	14,987	311,494	4.33
10-<50%	4,664	47	2,006	4,326	9,350	232,317	3.83	48	2,589	5,962	14,203	455,189	5.07
<10%	11,766	32	1,514	3,198	7,082	243,331	4.59	35	1,973	4,365	9,979	470,163	5.77
% Black or African American in Segment													
50-100%	1,390	34	1,879	4,038	8,486	99,821	3.18	36	2,181	4,925	10,831	437,673	5.89
10-<50%	3,610	60	2,056	4,172	8,565	226,138	3.95	74	2,570	5,670	12,341	380,775	4.99
<10%	13,229	32	1,563	3,407	7,663	243,331	4.43	35	2,035	4,715	11,268	470,163	5.50
% Owner-Occupied DUs¹ in Segment													
50-100%	13,482	51	1,827	3,863	8,489	243,331	4.10	53	2,340	5,286	12,160	470,163	5.29
10-<50%	3,722	47	1,733	3,491	7,559	206,306	4.13	48	2,220	4,754	10,765	437,673	5.12
<10%	1,025	32	675	1,334	2,823	62,644	3.69	35	866	1,795	4,384	130,806	5.03
Combined Median Rent/Housing Value													
1 st Quintile	3,051	51	1,326	2,869	6,283	178,076	4.55	53	1,680	3,852	8,825	384,587	6.02
2 nd Quintile	3,982	34	1,624	3,413	7,449	172,728	3.97	36	2,001	4,452	10,097	455,189	5.86
3 rd Quintile	4,517	32	1,647	3,506	7,840	213,433	3.80	35	2,146	4,888	11,345	470,163	4.99
4 th Quintile	3,855	47	1,851	3,996	8,716	232,317	4.06	48	2,366	5,433	13,199	431,487	5.21
5 th Quintile	2,824	51	2,124	4,444	9,518	243,331	4.58	63	2,902	6,400	14,483	314,883	4.96
Population Density													
Large MSA ¹	8,000	34	2,431	4,817	10,089	243,331	3.78	36	3,235	6,822	15,469	470,163	4.70
Medium to Small MSA ¹	8,833	32	1,307	2,841	6,293	213,433	4.47	35	1,696	3,784	8,768	455,189	5.86
Non-MSA, ¹ Urban	431	109	1,235	2,656	5,640	107,677	4.33	114	1,556	3,356	8,005	154,240	5.15
Non-MSA, ¹ Rural	965	67	944	2,232	5,335	131,150	4.93	69	1,210	2,962	6,817	218,260	6.03
Group Quarters													
Group	205	32	1,037	1,914	4,304	13,545	1.73	35	1,422	2,600	5,526	21,089	1.98
Non-Group	18,024	34	1,690	3,618	7,963	243,331	4.21	36	2,158	4,948	11,547	470,163	5.40

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 2014 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.ps and res.pr.ev)

Domain	n	Before res.pr.ps ¹ (SDUWT*PRWT12*...*PRWT14)						After res.pr.ps ¹ (SDUWT*PRWT12*...*PRWT15)						Final Weight: After res.pr.ev ¹ (SDUWT*PRWT12*...*PRWT16)						
		Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³	
Total	18,229	35	2,149	4,919	11,443	470,163	5.42	21	2,063	4,897	11,230	514,719	5.55	18	2,046	4,898	11,272	419,968	5.50	
<i>Pair Age Group</i>																				
12-17, 12-17	2,407	36	1,042	2,084	3,956	25,533	1.97	22	939	2,037	3,906	42,183	2.11	21	918	2,018	3,906	26,983	2.09	
12-17, 18-25	1,832	60	1,653	3,281	6,333	24,760	1.84	21	1,505	3,265	6,339	50,379	1.96	23	1,503	3,294	6,390	27,795	1.91	
12-17, 26-34	941	49	1,773	3,432	6,547	129,323	2.96	21	1,712	3,554	6,821	53,413	2.43	18	1,665	3,538	6,917	51,253	2.43	
12-17, 35-49	3,940	133	2,189	4,483	8,715	95,474	2.51	91	2,094	4,562	8,726	110,094	2.51	87	2,074	4,527	8,746	92,198	2.51	
12-17, 50+	776	286	5,263	11,007	18,499	137,805	2.17	178	4,544	10,218	18,851	181,804	2.47	178	4,433	10,157	18,505	233,098	2.65	
18-25, 18-25	2,585	35	1,618	3,140	6,399	45,819	2.01	22	1,519	3,205	6,476	45,912	2.03	21	1,512	3,219	6,574	45,978	2.00	
18-25, 26-34	870	263	2,526	5,037	9,329	157,568	3.11	180	2,386	4,946	9,161	119,865	3.00	180	2,340	4,969	9,317	110,333	2.87	
18-25, 35-49	1,180	355	4,466	8,730	18,481	98,439	2.21	191	4,354	8,637	18,212	105,128	2.24	187	4,334	8,694	18,188	92,999	2.22	
18-25, 50+	750	506	9,659	18,644	32,600	209,970	1.94	514	9,944	18,589	34,126	179,964	1.94	504	10,005	18,686	34,207	167,459	1.92	
26-34, 26-34	865	394	4,396	7,826	13,494	206,031	3.12	326	3,933	7,414	13,086	225,910	3.45	327	3,880	7,369	13,096	231,394	3.50	
26-34, 35-49	442	773	6,068	12,224	21,141	229,885	3.04	793	6,242	12,883	21,622	226,081	2.83	801	6,107	12,697	21,819	230,286	2.84	
26-34, 50+	235	2,254	16,843	40,105	73,919	437,673	1.96	1,883	15,293	35,484	71,879	319,611	1.97	1,908	15,062	35,780	73,041	323,367	1.99	
35-49, 35-49	658	593	6,131	12,852	27,460	380,775	3.45	320	5,430	12,505	28,095	454,069	3.67	310	5,341	12,383	27,750	414,795	3.57	
35-49, 50+	262	2,419	19,898	40,405	70,983	407,167	2.23	2,055	19,131	38,314	72,798	470,535	2.44	2,048	19,340	38,609	73,180	403,119	2.41	
50+, 50+	486	2,181	32,459	65,267	102,485	470,163	1.78	1,581	32,275	66,271	105,044	514,719	1.77	1,539	31,998	66,280	105,179	419,968	1.74	
<i>Pair Race/Ethnicity</i>																				
Hispanic or Latino	3,170	35	2,331	5,466	12,536	314,883	4.60	22	2,270	5,429	12,430	333,708	4.68	21	2,260	5,402	12,389	335,744	4.69	
Black or African American	1,928	36	2,414	5,402	11,797	437,673	5.54	37	2,675	5,796	12,169	351,540	5.15	37	2,630	5,775	12,218	362,114	5.20	
White	9,683	48	2,148	4,791	11,093	470,163	5.58	62	2,211	4,976	11,215	514,719	5.74	64	2,194	4,971	11,190	419,968	5.61	
Other	1,273	49	1,293	3,964	11,106	455,189	7.08	21	1,079	3,234	10,581	394,739	6.79	18	1,065	3,251	11,201	358,720	6.93	
White & Black or African American	186	96	2,469	4,872	12,148	202,675	4.29	114	2,173	5,369	14,009	279,796	5.46	115	2,120	5,358	13,593	278,239	5.43	
White & Hispanic or Latino	705	60	2,412	5,375	12,470	219,007	4.48	58	1,954	5,091	11,971	164,045	4.32	56	1,896	5,171	11,582	163,979	4.32	
White & Other	793	110	1,823	4,359	9,920	248,496	4.62	46	1,194	3,066	7,779	173,726	4.52	46	1,165	3,031	7,890	176,464	4.57	
Black or African American & Hispanic or Latino	127	198	2,341	4,912	11,580	242,429	5.48	160	2,002	4,466	11,758	208,380	5.12	151	1,918	4,185	11,548	211,006	5.27	
Black or African American & Other	191	74	2,048	4,419	8,513	121,400	3.62	22	892	2,019	5,632	82,013	3.92	21	883	2,025	5,717	84,290	4.00	
Hispanic or Latino & Other	173	82	1,683	4,591	13,199	114,324	3.64	30	1,038	2,563	7,640	88,381	3.83	28	1,029	2,505	7,740	87,611	3.86	

Table L.3 2014 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.ps and res.pr.ev) (continued)

Domain	n	Before res.pr.ps ¹ (SDUWT*PRWT12*...*PRWT14)						After res.pr.ps ¹ (SDUWT*PRWT12*...*PRWT15)						Final Weight: After res.pr.ev ¹ (SDUWT*PRWT12*...*PRWT16)					
		Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³
<i>Pair Gender</i>																			
Male, Male	3,722	36	2,063	4,687	10,765	431,487	5.01	22	1,945	4,755	10,653	333,708	4.89	21	1,911	4,779	10,714	335,744	4.85
Female, Female	4,219	35	1,987	4,417	9,352	314,883	4.89	21	1,959	4,482	9,497	351,540	5.21	18	1,947	4,441	9,471	362,114	5.27
Male, Female	10,288	39	2,276	5,236	12,590	470,163	5.41	30	2,180	5,198	12,548	514,719	5.57	27	2,153	5,194	12,596	419,968	5.49
<i>Household Size</i>																			
Two	4,431	35	1,941	4,350	11,268	357,571	4.51	21	1,867	4,305	11,098	227,727	4.50	18	1,832	4,287	11,118	239,176	4.56
Three	5,841	36	1,834	4,047	8,151	411,221	7.33	21	1,740	4,030	8,292	514,719	7.52	21	1,719	4,016	8,298	419,968	7.41
Four or More	7,957	56	2,592	6,216	14,140	470,163	4.93	31	2,502	6,084	13,878	470,535	5.09	27	2,501	6,097	13,905	414,795	5.02
<i>Census Region</i>																			
Northeast	3,335	49	1,818	4,204	11,325	470,163	6.58	21	1,639	4,021	10,725	514,719	6.59	18	1,613	3,968	10,873	419,968	6.39
South	6,101	36	2,609	5,552	12,216	407,167	5.12	21	2,556	5,547	12,071	470,535	5.34	21	2,523	5,536	12,058	403,119	5.31
Midwest	4,207	35	2,326	4,732	10,127	305,400	4.51	22	2,256	4,757	10,119	366,668	4.72	21	2,234	4,760	10,160	351,501	4.73
West	4,586	53	1,616	4,676	12,041	455,189	5.54	26	1,568	4,696	12,099	394,739	5.57	26	1,566	4,720	11,993	358,720	5.54
<i>Quarter</i>																			
Quarter1	4,254	35	2,306	5,267	11,906	431,487	5.23	22	2,275	5,443	11,962	443,856	5.16	21	2,259	5,411	11,842	415,334	5.15
Quarter2	5,113	39	1,937	4,413	10,176	455,189	5.59	21	1,829	4,351	9,841	514,719	5.89	23	1,813	4,327	9,932	419,968	5.88
Quarter3	4,455	36	2,095	4,970	11,932	470,163	5.45	21	2,007	4,990	12,090	449,238	5.47	18	2,008	4,996	12,166	414,795	5.37
Quarter4	4,407	74	2,358	5,145	11,942	437,673	5.36	30	2,248	5,022	11,711	470,535	5.65	28	2,238	4,996	11,795	403,119	5.57
<i>% Hispanic or Latino in Segment</i>																			
50-100%	1,799	74	2,813	6,303	14,987	311,494	4.33	74	2,721	6,356	14,839	333,708	4.47	70	2,714	6,317	14,867	335,744	4.49
10-<50%	4,664	48	2,589	5,962	14,203	455,189	5.07	30	2,502	5,925	14,232	449,238	5.04	28	2,497	5,882	14,184	414,795	5.03
<10%	11,766	35	1,973	4,365	9,979	470,163	5.77	21	1,866	4,311	9,857	514,719	6.00	18	1,849	4,297	9,919	419,968	5.91
<i>% Black or African American in Segment</i>																			
50-100%	1,390	36	2,181	4,925	10,831	437,673	5.89	22	2,324	5,224	11,015	321,057	5.03	21	2,259	5,208	11,026	313,048	5.02
10-<50%	3,610	74	2,570	5,670	12,341	380,775	4.99	53	2,401	5,597	12,410	454,069	5.38	60	2,401	5,584	12,337	415,334	5.30
<10%	13,229	35	2,035	4,715	11,268	470,163	5.50	21	1,945	4,664	11,008	514,719	5.65	18	1,924	4,647	11,023	419,968	5.60
<i>% Owner-Occupied DUs¹ in Segment</i>																			
50-100%	13,482	53	2,340	5,286	12,160	470,163	5.29	21	2,250	5,276	12,078	514,719	5.39	23	2,232	5,260	12,115	419,968	5.32
10-<50%	3,722	48	2,220	4,754	10,765	437,673	5.12	22	2,139	4,778	10,758	394,739	5.38	21	2,113	4,757	10,773	358,720	5.43
<10%	1,025	35	866	1,795	4,384	130,806	5.03	21	761	1,792	4,258	138,736	5.31	18	741	1,766	4,234	138,262	5.29
<i>Combined Median Rent/Housing Value</i>																			
1 st Quintile	3,051	53	1,680	3,852	8,825	384,587	6.02	58	1,590	3,827	8,962	514,719	6.22	56	1,569	3,784	9,024	419,968	6.01
2 nd Quintile	3,982	36	2,001	4,452	10,097	455,189	5.86	21	1,952	4,466	10,182	454,069	5.82	18	1,925	4,439	10,219	377,599	5.80
3 rd Quintile	4,517	35	2,146	4,888	11,345	470,163	4.99	21	2,017	4,861	11,176	385,048	4.98	21	2,008	4,821	11,257	353,409	4.96
4 th Quintile	3,855	48	2,366	5,433	13,199	431,487	5.21	22	2,338	5,488	12,891	443,856	5.24	21	2,321	5,458	12,924	415,334	5.27
5 th Quintile	2,824	63	2,902	6,400	14,483	314,883	4.96	62	2,685	6,125	14,386	470,535	5.40	64	2,688	6,146	14,495	414,795	5.28

Table L.3 2014 NSDUH Respondent Pair-Level Weight Summary Statistics (res.pr.ps and res.pr.ev) (continued)

Domain	n	Before res.pr.ps ¹ (SDUWT*PRWT12*...*PRWT14)						After res.pr.ps ¹ (SDUWT*PRWT12*...*PRWT15)						Final Weight: After res.pr.ev ¹ (SDUWT*PRWT12*...*PRWT16)					
		Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³	Min	Q1 ²	Med	Q3 ²	Max	UWE ³
<i>Population Density</i>																			
Large MSA ¹	8,000	36	3,235	6,822	15,469	470,163	4.70	22	3,087	6,794	15,499	470,535	4.75	21	3,058	6,806	15,528	415,334	4.70
Medium to Small MSA ¹	8,833	35	1,696	3,784	8,768	455,189	5.86	21	1,601	3,768	8,729	514,719	6.13	18	1,580	3,738	8,733	419,968	6.10
Non-MSA, ¹ Urban	431	114	1,556	3,356	8,005	154,240	5.15	88	1,457	3,300	7,997	193,257	5.51	87	1,436	3,308	7,949	207,816	5.63
Non-MSA, ¹ Rural	965	69	1,210	2,962	6,817	218,260	6.03	65	1,193	2,986	6,693	204,772	6.26	62	1,180	2,985	6,711	211,746	6.26
<i>Group Quarters</i>																			
Group	205	35	1,422	2,600	5,526	21,089	1.98	22	1,303	2,789	4,899	26,174	2.08	21	1,267	2,862	5,034	23,438	2.02
Non-Group	18,024	36	2,158	4,948	11,547	470,163	5.40	21	2,079	4,947	11,334	514,719	5.53	18	2,055	4,937	11,359	419,968	5.48
<i>Pair Relationship Domain⁴</i>																			
Parent-Child (12-14)	2,813	49	2,025	4,130	7,954	129,323	2.82	21	2,096	4,560	8,903	181,804	2.87	18	2,076	4,522	8,891	233,098	2.99
Parent-Child (12-17)	5,178	49	2,201	4,548	9,278	129,323	2.70	21	2,264	4,826	9,729	181,804	2.82	18	2,229	4,799	9,674	233,098	2.93
Parent-Child (12-20)	6,040	49	2,439	5,096	11,038	169,559	2.78	21	2,489	5,397	11,289	181,804	2.88	18	2,457	5,363	11,237	233,098	2.93
Sibling (12-14)-Sibling (15-17)	1,418	36	1,080	2,159	3,999	23,954	1.91	22	1,030	2,194	4,182	42,183	2.08	21	1,031	2,171	4,179	26,983	2.02
Sibling (12-17)-Sibling (18-25)	1,671	60	1,667	3,224	6,315	24,760	1.85	21	1,519	3,248	6,364	50,379	1.96	23	1,526	3,275	6,405	27,795	1.91
Spouse-Spouse/ Partner-Partner	3,566	96	2,623	6,523	17,315	470,163	4.80	64	2,390	6,186	17,549	470,535	4.88	60	2,358	6,132	17,350	414,795	4.84
Spouse-Spouse/ Partner-Partner with Children (Younger Than 18)	1,796	199	2,682	6,163	13,730	380,775	5.15	182	3,051	6,909	15,862	470,535	5.60	170	3,029	6,928	15,877	414,795	5.35

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.

Appendix M: Pair Analysis Manual Excerpt

This page intentionally left blank

Appendix M: Pair Analysis Manual Excerpt

This appendix provides background information on pair data analysis for the National Survey on Drug Use and Health (NSDUH). This draft excerpt is from Section 3.2 in *How To Prepare and Analyze Pair Data in the National Survey on Drug Use and Health* by Center for Behavioral Health Statistics and Quality (in press).

Inferential Population and Multiplicities

There are different perspectives through which pair data can be analyzed: (1) with pairs as the focus (e.g., where the focus of the analysis, or dependent variable, is on how the parent and child jointly report behaviors of interest), or (2) with one member of the pair as the focus (e.g., the child's behavior as the dependent variables and his or her parent's behavior as the independent variable). When the focus is on the pair, the *PRANALWT* variable can be used to weight the data directly with no adjustments. However, when the focus is on one member of the pair, an adjustment often needs to be made to the weight to account for the fact that this member of the pair may be a member of more than one pair of this type (Chromy & Singh, 2001).

Section 2.4.2 describes the weights for the three examples introduced in Section 1.6. The first two of these three examples – example (a) and (b) - have the focus on one member of the pair (i.e., behavior of one member of the pair is the dependent variable and the other is the independent variable). The third of these examples has the focus on the pair itself (i.e., the behaviors of both members of the pairs are jointly the dependent variable). When the focus is on the pair itself, as is the case in example (c), no multiplicity is needed and the *PRANALWT* variable can be used directly.

Since the focuses of examples (a) and (b) introduced in Section 1.6 are on one member of the pair, there is potential for the need of an adjustment to be made to the *PRANALWT* variable due to a multiple counting problem. The *PRANALWT* variable weights each pair member by the number of pairs of that type that the respondent represents. In example (a), we are treating the father's behavior as the focus, or dependent variable. The *PRANALWT* variable weights each father-child pair by the number of father-child pairs there are in the household. When we are focused on the father's behavior, however, we do not want fathers who have more children to be more heavily weighted than fathers with only one child. This introduces a multiple counting problem.

Adjustments for this multiple counting problem are done using "multiplicities" (Chromy & Singh, 2001). Many multiplicities have been computed and are on the data file; in analyses these can be accounted for by simply adjusting the weight variable. Analysts should exercise care in computing multiplicities for any pair types not listed. The process by which the multiplicities are created is described in detail in Chapter 11 of Laufenberg et al. (2015). [Table 5](#) lists the pair domains for which multiplicities can be used to perform person-level analyses. These multiplicity variables are used in conjunction with the *PRANALWT* variable when adjusting the weight. The new weight is assigned the value of *PRANALWT* divided by the multiplicity variable.

Table 5. Pair Domains and Multiplicities

Pair Domain		Multiplicity Variable		
<i>IRPRRE</i> <i>L</i> Levels	Description	Focus ¹	Name	Description
1	Parent-child, child aged 12-14	Parent	<i>IRMPCP1</i> 4	Number of children aged 12-14 living with responding parent
		Child	<i>IRMPCC1</i> 4	Number of parents living with responding child aged 12-14
1,2	Parent-child, child aged 12-17	Parent	<i>IRMPCP1</i> 7	Number of children aged 12-17 living with responding parent
		Child	<i>IRMPCC1</i> 7	Number of parents living with responding child aged 12-17
1,2,3	Parent-child, child aged 12-20	Parent	<i>IRMPCP2</i> 0	Number of children aged 12-20 living with responding parent
		Child	<i>IRMPCC2</i> 0	Number of parents living with responding child aged 12-20
2	Parent-child, child aged 15-17	Parent	<i>IRMPCP5</i> 7	Number of children aged 15-17 living with responding parent
		Child	<i>IRMPCC5</i> 7	Number of parents living with responding child aged 15-17
5	Sibling-sibling, older sibling aged 15-17, younger sibling aged 12-14	Older sibling	<i>IRMS1417</i>	Number of siblings aged 12-14 living with responding sibling aged 15-17
		Younger sibling	<i>IRMS1714</i>	Number of siblings aged 15-17 living with responding sibling aged 12-14
6	Sibling-sibling, older sibling aged 18-25, younger sibling aged 12-17	Older sibling	<i>IRMS1725</i>	Number of siblings aged 12-17 living with responding sibling aged 18-25
		Younger sibling	<i>IRMS2517</i>	Number of siblings aged 18-25 living with responding sibling aged 12-17
8,9	Spouse-spouse and partner-partner	No multiplicity necessary: assume only one spouse per person		
8	Spouse-spouse and partner-partner, with children aged 0-17	No multiplicity necessary: assume only one spouse per person		

¹ No weight adjustment is needed when the inferential focus is on the pair.