



2019 National Survey on Drug Use and Health (NSDUH) Methodological Resource Book

Section 14: Sample Experience Report

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

July 2020

2019 National Survey on Drug Use and Health (NSDUH) Methodological Resource Book: Section 14: Sample Experience Report

Acknowledgments

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

Public Domain Notice

All material appearing in this report is in the public domain and may be reproduced or copied without permission from SAMHSA. Citation of the source is appreciated. However, this publication may *not* be reproduced or distributed for a fee without the specific, written authorization of the Office of Communications, SAMHSA, HHS.

Electronic Access

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

Recommended Citation

Center for Behavioral Health Statistics and Quality. (2020). *2019 National Survey on Drug Use and Health (NSDUH) methodological resource book: Section 14: Sample experience report*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Originating Office

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

Nondiscrimination Notice

SAMHSA complies with applicable federal civil rights laws and does not discriminate on the basis of race, color, national origin, age, disability, or sex. SAMHSA cumple con las leyes federales de derechos civiles aplicables y no discrimina por motivos de raza, color, nacionalidad, edad, discapacidad o sexo.

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

July 2020

Table of Contents

| Chapter | Page |
|--------------|---|
| <u>1.</u> | <u>Introduction..... 1</u> |
| <u>2.</u> | <u>Overview of the 2019 Sample Design 3</u> |
| <u>2.1</u> | <u>Target Population..... 3</u> |
| <u>2.2</u> | <u>Design Overview 3</u> |
| <u>2.2.1</u> | <u>Multiyear Coordinated Design..... 4</u> |
| <u>2.2.2</u> | <u>Sample Frame 7</u> |
| <u>2.2.3</u> | <u>Sample Selection at Fourth and Fifth Stages 9</u> |
| <u>2.2.4</u> | <u>Creation of Variance Strata and Variance PSUs..... 11</u> |
| <u>3.</u> | <u>Segment (Third-Stage) Sample Experience..... 13</u> |
| <u>4.</u> | <u>Dwelling Unit (Fourth-Stage) Sample Experience 17</u> |
| <u>5.</u> | <u>Person (Fifth-Stage) Sample Experience 23</u> |
| <u>6.</u> | <u>Sampling Error..... 29</u> |
| <u>6.1</u> | <u>Computation of Relative Standard Errors and Design Effects 29</u> |
| <u>6.2</u> | <u>Derivation of the Approximation for the RSE of a Prevalence Estimate's Negative Natural Logarithm 31</u> |
| <u>6.3</u> | <u>Comparison of Observed Precision with Expected Precision..... 31</u> |
| <u>6.3.1</u> | <u>Sample and Precision Requirements..... 32</u> |
| <u>6.3.2</u> | <u>Observed versus Expected Precision 34</u> |
| <u>6.4</u> | <u>Comparison of Median and Mean Design Effects 34</u> |
| <u>7.</u> | <u>Issue Encountered 37</u> |
| | <u>References..... 39</u> |
| | <u>List of Contributors..... 41</u> |

This page intentionally left blank

List of Tables

| Table | | Page |
|--------------|--|-------------|
| <u>3.1</u> | <u>Duplicated Segments in the 2019 NSDUH Sample.....</u> | <u>14</u> |
| <u>3.2</u> | <u>2019 NSDUH Segment Quarter Switches in the Field.....</u> | <u>15</u> |
| <u>4.1</u> | <u>2019 NSDUH Segment and Dwelling Unit Summary</u> | <u>17</u> |
| <u>4.2</u> | <u>2019 NSDUH Quarterly Dwelling Unit Sample Sizes and Percentages Released.....</u> | <u>20</u> |
| <u>5.1</u> | <u>2019 NSDUH Yields, by Age Group and State.....</u> | <u>23</u> |
| <u>5.2</u> | <u>2019 NSDUH Projected and Observed Pair Selection Counts, by Age Group Pairs (Three Age Groups: 12 to 17, 18 to 25, and 26 or Older).....</u> | <u>25</u> |
| <u>5.3</u> | <u>2019 NSDUH Projected and Observed Pair Response Rates, by Age Group Pairs (Three Age Groups: 12 to 17, 18 to 25, and 26 or Older).....</u> | <u>26</u> |
| <u>5.4</u> | <u>2019 NSDUH Weighted Screening and Interview Response Rates, by State</u> | <u>27</u> |
| <u>6.1</u> | <u>Comparisons of Projected and Observed Relative Standard Errors and Sample Sizes for Key Outcome Measures, by Demographic Domain, 2019 NSDUH</u> | <u>33</u> |
| <u>6.2</u> | <u>Comparison of Median and Mean Design Effects of 53 Outcomes: 2019</u> | <u>34</u> |
| <u>6.3</u> | <u>Median and Mean Design Effects of 53 Outcomes: 2017 through 2019.....</u> | <u>35</u> |
| <u>7.1</u> | <u>2019 NSDUH Segments with Truncated Sample Sizes</u> | <u>37</u> |

This page intentionally left blank

1. Introduction

A coordinated sample design was developed for the 2014 through 2017 National Surveys on Drug Use and Health (NSDUHs) and was extended to the 2018 through 2022 NSDUHs.¹ The multiyear design consists of a deeply stratified, multistage area probability design. The 2019 sample design is thoroughly documented in the 2019 NSDUH sample design report (Center for Behavioral Health Statistics and Quality, 2020a). The goal of this report is to further document the 2019 NSDUH sample experiences, including a comparison of actual sample yields to state and quarter targets, a comparison of achieved and expected design effects (DEFFs) and relative standard errors, and documentation of any issues encountered during sample implementation.

This report is organized as follows. Chapter 2 summarizes the 2019 sample design. Chapters 3, 4, and 5 document the sample experiences at the third, fourth, and fifth stages of sample selection, respectively. Chapter 6 includes a comparison of the observed precisions with the expected precisions and a comparison of median and mean DEFFs. Finally, one issue encountered during sample implementation is described in Chapter 7.

¹ This report presents information from the 2019 NSDUH. Prior to 2002, the survey was called the National Household Survey on Drug Abuse (NHSDA).

This page intentionally left blank

2. Overview of the 2019 Sample Design

2.1 Target Population

The respondent universe for the 2019 National Survey on Drug Use and Health (NSDUH) was the civilian, noninstitutionalized population aged 12 years or older residing in the United States. Consistent with the NSDUH designs since 1991, the 2019 NSDUH universe included residents of noninstitutional group quarters (e.g., shelters, rooming houses, dormitories, and group homes), residents of Alaska and Hawaii, and civilians residing on military bases in the United States. People excluded from the 2019 universe included those with no fixed household address (e.g., homeless or transient people not in shelters) and residents of institutional group quarters, such as jails and hospitals.

2.2 Design Overview

The Substance Abuse and Mental Health Services Administration (SAMHSA) implemented major changes in the way that NSDUH would be conducted beginning in 1999 and continuing through subsequent years. The survey is conducted using computer-assisted interviewing methods and provides state estimates based on target sample sizes per state. Furthermore, the NSDUH sample was redesigned in 2014 to allow for a more cost-efficient sample allocation to the largest states while maintaining adequate sample sizes in smaller states to support reliable state and substate estimates based on the small area estimation (SAE) methodology. Reliable direct state estimates are also possible (in any state) by pooling multiple years of data. The target national sample size of 67,507 is distributed across five age groups as follows: 25 percent for youths aged 12 to 17, 25 percent for young adults aged 18 to 25, 15 percent for adults aged 26 to 34, 20 percent for adults aged 35 to 49, and 15 percent for adults aged 50 or older. This large sample size allows SAMHSA to continue reporting precise estimates for demographic subgroups at the national level without needing to oversample specially targeted demographics, as was required prior to 1999. The achieved sample for the 2019 NSDUH was 67,625 people.

Beginning with the 2002 NSDUH and continuing through the 2019 NSDUH, survey respondents were given a \$30 incentive for participation. As expected, the incentive had the effect of increasing response rates, thereby requiring fewer selected households than previous surveys. In recent years, however, response rates have been declining, which has required the number of selected households to increase. Beginning in 2014 and continuing through 2019, this increase was partially offset by selecting fewer youths aged 12 to 17, requiring fewer selected households per completed interview. That is, with more proportional sampling by age group, fewer households are needed to support the oversample of youths aged 12 to 17.

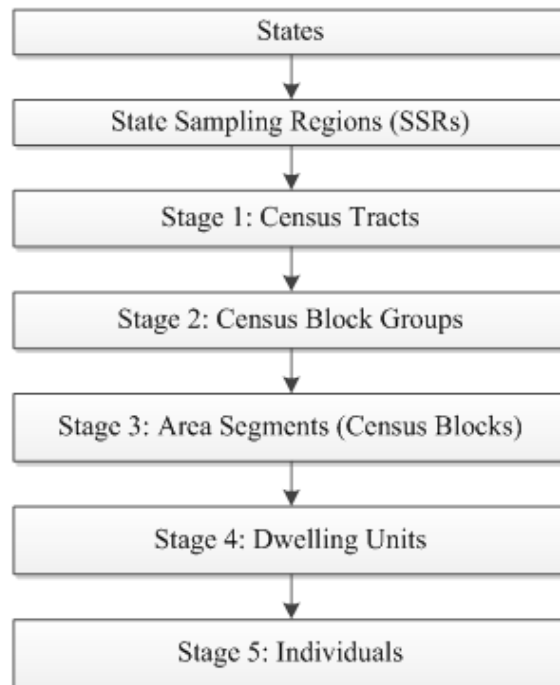
An additional design change was made in 2002 and continued through 2013. A new pair sampling strategy was implemented that increased the number of pairs selected in dwelling units (DUs) with older people on the roster (Chromy & Penne, 2002). With the increase in the number of pairs came a moderate decrease in the response rate for older people. In 2014, changes to the sample design with respect to age group and state necessitated a review of the pair sampling strategy. As a result, slightly fewer pairs were selected for the 2014 through 2019 NSDUHs.

2.2.1 Multiyear Coordinated Design

A coordinated sample design was developed for the 2014 through 2017 NSDUHs. A large reserve sample of area clusters or segments was selected at the time the 2014 through 2017 NSDUH sample was selected. This reserve sample is being used to field the 2018 through 2022 NSDUHs. Thus, the 2018 through 2022 NSDUH designs simply continue the coordinated design.

[Exhibit 1](#) summarizes the 2014 through 2022 NSDUH multistage design. The coordinated design facilitates 50 percent overlap in third-stage units (area segments) within each successive 2-year period from 2014 through 2022. This designed sample overlap may slightly increase the precision of estimates of year-to-year trends when there is a small positive correlation in successive survey years due to an overlapped segment being somewhat homogeneous. The 50 percent overlap of segments significantly reduces segment listing costs because only one half of the segments need to be listed for the 2015 through 2022 surveys.

Exhibit 1. Summary of the 2014 through 2022 NSDUH Design



The 2019 design provides for estimates by state in all 50 states plus the District of Columbia. States may therefore be viewed as the first level of stratification and as a reporting variable. In the 2005 through 2013 NSDUH design, the sample was divided into 8 "large" states and 43 "small" states (including the District of Columbia), with the large and small sample states designed to yield 3,600 and 900 respondents per state, respectively. Beginning in 2014 and continuing through 2019, the survey's sample was designed to yield the following:

- 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.

To accommodate state and local policymakers' need for substate estimates in Kauai County, Hawaii, the sample was designed to yield a minimum of 200 completed interviews in this county over a 3-year period. This allows for Kauai County to be included as a separate entity in the production of substate estimates that are produced biennially and typically based on 3 years of data. To achieve this goal while maintaining precision at the state level, Kauai County was treated separately from the remainder of Hawaii for sample allocation and sample size management purposes. The target annual sample in Hawaii consisted of 67 completed interviews in Kauai County and 900 completed interviews in the remainder of the state, for an expected total of approximately 967 completed interviews each year.

In all states, the sample sizes were sufficient to support reliable estimates based on the SAE methodology while maintaining efficiency for national estimates (RTI International, 2012). Direct estimates and estimates of change are typically produced by pooling multiple years of data to increase precision. For example, 2008-2009 data are pooled, 2017-2018 data are pooled, then the two pooled estimates are compared for an estimate of long-term change (Center for Behavioral Health Statistics and Quality [CBHSQ], 2019-2020).

Prior to selecting the sample for the 2014 through 2022 NSDUHs, state sampling regions (SSRs) were formed within each state. These SSRs will be used for the duration of the 2014 through 2022 NSDUHs. Based on a composite size measure,² each state was geographically partitioned into roughly equal-sized regions according to population. In other words, regions were formed such that each area yielded, in expectation, roughly the same number of interviews within each state during each quarterly data collection period. This partitioning divided the United States into 750 SSRs.

Similar to the 2005 through 2013 NSDUHs, the first stage of selection for the 2014 through 2022 NSDUHs was census tracts.³ This stage was included to contain sample segments within a single census tract to the extent possible.⁴ Segments that cross census tract boundaries make merging to external data sources difficult.

The first stage of selection began with the construction of an area sample frame that contained one record for each census tract in the United States. If necessary, census tracts were aggregated within SSRs until each primary sampling unit (PSU; one or more census tracts) met the minimum size requirement. In California, Florida, Georgia, Illinois, Michigan, New Jersey, New York, North Carolina, Ohio, Pennsylvania, Texas, and Virginia, this minimum size

² The composite size measure is defined as the population weighted by the state sampling rate in each age group.

³ A census tract is a small, relatively permanent statistical subdivision of a county or equivalent entity that contains between 1,200 and 8,000 people, with an optimum size of 4,000 people (U.S. Census Bureau, Redistricting Data Office, 2009).

⁴ Some census tracts had to be aggregated in order to meet the minimum DU requirement.

requirement was 250 DUs⁵ in urban areas and 200 DUs in rural areas.⁶ In the remaining states and the District of Columbia, the minimum requirement was 150 DUs in urban areas and 100 DUs in rural areas.

Before selecting PSUs, additional implicit stratification was achieved by sorting the first-stage sampling units by a CBSA/SES⁷ (core-based statistical area/socioeconomic status) indicator⁸ and by the percentage of the population that is non-Hispanic and white.⁹ From this well-ordered sample frame, 48 PSUs per SSR were sequentially selected with probabilities proportionate to a composite size measure and with minimum replacement¹⁰ (Chromy, 1979).

For the second stage of selection, adjacent census block groups were aggregated within selected PSUs as necessary to meet the minimum DU requirements (150 or 250 DUs in urban areas and 100 or 200 DUs in rural areas according to state). After the resulting second-stage sampling units (SSUs) were formed, they were sorted in the order they were formed (i.e., geographically), and one SSU was selected per sampled PSU with probability proportionate to a composite size measure and with minimum replacement (Chromy, 1979). Compared with 2013 and prior years, the selection of census block groups in the 2014 through 2022 NSDUH samples is an additional stage of selection to facilitate possible transitioning to an address-based sampling (ABS) design in the future.

The SSUs were generally larger than practical for building frames of housing units through field enumeration. Therefore, one smaller geographic region was selected within each sampled SSU. For this third stage of sampling, each selected SSU was partitioned into compact

⁵ DU counts were obtained from the 2010 census data supplemented with revised population counts from Claritas, a market research firm (see <https://www.claritas.com/> .

⁶ The basis for the differing minimum DU requirement in urban and rural areas is that it is more difficult to meet the requirement in rural areas, 100 DUs are sufficient to support one field test and two main study samples in the smaller states, and 200 DUs are sufficient to support three samples in the larger sample states.

⁷ CBSAs include metropolitan and micropolitan statistical areas as defined by the Office of Management and Budget (2009). Metropolitan statistical areas contain at least one urbanized area with 50,000 or more people and may include adjacent territory with a high degree of social and economic integration with the core as measured by commuting. Micropolitan statistical areas have an urban core with at least 10,000 but fewer than 50,000 people, plus adjacent territory that is socioeconomically tied to the core by commuting. Both metropolitan and micropolitan statistical areas are defined in terms of whole counties (or equivalent entities).

⁸ Four categories are defined as (1) CBSA/low SES, (2) CBSA/high SES, (3) non-CBSA/low SES, and (4) non-CBSA/high SES. To define SES, census tract-level median rents and property values obtained from the 2006 to 2010 American Community Survey (ACS) data were given a rank (1,...,5) based on state and CBSA quintiles. The rent and value ranks then were averaged, weighted by the percentages of renter- and owner-occupied DUs, respectively. If the resulting score fell in the lower 25th percentile by state and CBSA, the area was considered "low SES"; otherwise, it was considered "high SES."

⁹ Although the large sample size eliminates the need for the oversampling of specially targeted demographic subgroups as was required prior to the 1999 National Household Survey on Drug Abuse (NHSDA), sorting by a CBSA/SES indicator and by the percentage of the population that is non-Hispanic and white ensures dispersion of the sample with respect to SES and race/ethnicity. Implicit stratification also has the potential to make sampling error lower by reducing the selection of neighboring and possibly similar segments than if the selection was done completely at random.

¹⁰ Probability minimum replacement sampling is a type of probability proportional to size sampling that allows for a sampling unit with a relatively large measure of size to be selected more than once. More importantly, neighboring units on the sorted sampling frame are unlikely to be jointly selected for the sample. This reduces the potential for variance when neighboring units are similar with respect to the outcomes being measured.

clusters¹¹ of DUs by aggregating adjacent census blocks.¹² Consistent with the terminology used in previous NSDUHs, these geographic clusters of blocks are referred to as "segments." A sample DU in NSDUH refers to either a housing unit or a group quarters listing unit, such as a dormitory room or a shelter bed. Similar to PSUs and SSUs, segments were formed to contain a minimum of 150 or 250 DUs in urban areas and 100 or 200 DUs in rural areas according to state. This minimum DU requirement will support the overlapping sample design and any special supplemental samples or field tests that SAMHSA may wish to conduct.

Prior to selection, the segments were sorted in the order they were formed (i.e., geographically), and one segment was selected within each sampled SSU using Chromy's method of sequential random sampling (with probability proportionate to a composite size measure and with minimum replacement) (Chromy, 1979). The 48 selected segments within each SSR were randomly assigned to a survey year and quarter of data collection. Although only 20 segments per SSR were needed to support the 2014 through 2017 NSDUHs, an additional 28 segments were selected to serve as replacements when segment DUs are depleted, to support any supplemental studies embedded within NSDUH, and to extend the sample to the next decennial census, if desired. These 28 segments constitute the "reserve" sample and are being used to support the 2018 through 2022 NSDUHs.

An equal probability subsample of eight segments per SSR is used for each NSDUH year. These eight segments are randomly assigned to quarters and to two panels within each quarter. For 2019, the first panel segments (panel F) were used for the 2018 and 2019 surveys, constituting the overlap sample. The second panel segments (panel G) were used for the 2019 survey and are being used again for the 2020 survey.

2.2.2 Sample Frame

Beginning in 2014, three changes related to the sample frame were implemented. First, although the sampling frame for the 2005 through 2013 NSDUHs was constructed using 2000 census data, the 2014 through 2022 sample frame was built using 2010 census data supplemented with 2013 population projections from Claritas.¹³ Furthermore, because the Census Bureau's long-form data were no longer available, census tract-level median rents and property values were obtained from the 2006 to 2010 ACS. These data were used to form the CBSA/SES indicator that was used in the implicit stratification of the first-stage sampling units (census tracts or PSUs). This change was confirmed to improve coverage and therefore require smaller poststratification adjustments in weighting (CBHSQ, 2015b).

¹¹ Although the entire cluster is compact, the final sample of DUs represents a noncompact cluster. Noncompact clusters (selection from a list) differ from compact clusters in that not all units within the cluster are included in the sample. Although compact cluster designs are less costly, a noncompact cluster design was used because it provides for greater heterogeneity of dwellings within the sample. Also, social interaction (contagion) among neighboring dwellings is sometimes introduced with compact clusters (Kish, 1965).

¹² A census block is a small statistical area bounded by visible features (streets, roads, streams, railroad tracks, etc.) and nonvisible boundaries (e.g., city, town, and county limits). A block group is a cluster of census blocks within the same census tract and generally contains between 300 and 6,000 people (U.S. Census Bureau, Redistricting Data Office, 2009).

¹³ See footnote 5 for information about Claritas.

Next, the number and distribution of SSRs were revised in 2014. In the 2005 through 2013 design, the 8 large states were each partitioned into 48 SSRs and the 43 small states were each partitioned into 12 SSRs, for a total of 900 SSRs. Beginning in 2014, the sampling frame was stratified into 750 SSRs with the number of SSRs varying by state. In each of the eight large states (California, Florida, Illinois, Michigan, New York, Ohio, Pennsylvania, and Texas), the total number of SSRs was reduced. In four of the small states (Georgia, New Jersey, North Carolina, and Virginia), the total number of SSRs was increased, while there was no change in the number of SSRs in the remaining small states. Thus, the change in the number and distribution of SSRs affected only 12 states.

In general, the new SSR distribution in the affected states resulted in increased efficiency in the highly populated areas and efficiency losses in the less populated areas. In the highly populated areas, some efficiency was gained because the SSRs and segment locations were more compact and the work could be completed by fewer field interviewers (FIs). The concentrated locations reduced travel for the FIs, provided sufficient options for case assignments, and provided the option for FIs to work more hours, if desired. In some less populated areas, the decrease in SSRs created some inefficiencies because the SSRs were larger and the segment locations were not as central. In these areas, it was often challenging for an FI to cover all of the work in an SSR due to the varying locations of segments and the location of an FI's home. Depending on the quarter, some SSRs experienced inefficiencies due to increased FI travel for the initial assignments and fewer FIs to help with data collection cleanup. Not all of the FIs were willing or available to travel longer distances, creating some inefficiencies in case assignments. In these areas, the use of borrowed FIs (BFIs) and sometimes traveling FIs (TFIs) was required.

In Illinois, Michigan, Ohio, and Pennsylvania, the total number of SSRs was cut in half, and the average segment size was increased. As a result, these states experienced both gains and losses in efficiency. In the highly populated areas, these states experienced gains due not only to increased yield and clustering, but also to staffing selection. Field management for 2014 was able to reduce the number of field staff members in the highly populated areas of these states. With location being equal, the most proficient, efficient, and dedicated field staff members were retained. In the less populated areas of these states, increased SSR sizes and varying segment locations caused inefficiencies. Some segments had no nearby FI and had to be worked by a BFI or TFI, thus increasing travel costs. Overall, the greater yields in these states resulted in gains in efficiency, but also kept the field staff members working in their local segments longer while waiting for that assignment to be finished and before going to another area to work or clean up.

Overall, the new SSR distribution resulted in a reduction of hours and miles per interview; however, miscellaneous travel expenses increased. The SSR distribution also affected scheduling because it was initially not known what the yield would be in an area or when an FI would be available to move on to another assignment. With the higher yield, FIs took longer to work their initial assignments and were delayed while working a travel assignment. Compared with previous designs when field staff members completed their assignments several weeks before the end of the quarter, FIs were working late into each quarter. Without a short break at the end of the quarter, getting off to a strong start at the beginning of the following quarter was more challenging. That is, without the downtime at the end of the quarter that they were accustomed to, field staff were not as motivated when working their new assignments during the following quarter.

The issues related to the new SSR distribution continued into 2019. In response to these issues, management of the FI case assignments and staffing levels was modified based on the 2014 through 2018 field experience. Staffing levels were increased in many states, and new local staff were hired with the expectation of fewer hours per week and required travel in rural areas to accommodate quarterly fluctuations in location assignments. Because management of the new SSR distribution required more travel to cover remote areas, a team of 14 dedicated TFIs was maintained, and field supervisors increased the utilization of FIs who could travel from adjacent states. Field supervisors proactively made quarterly assignments to account for working cases later into the quarter, and weekly production benchmarks were adjusted to reflect realistic expectations based on the 2018 fieldwork.

The third change, which was mentioned previously, was the addition of a sample selection stage by selecting census block groups from selected PSUs. The purpose of this change was to facilitate the possible transition to ABS. The introduction of census block groups as a sampling stage was transparent in the area sampling results and was expected to have little impact on the person-level analysis weight.

2.2.3 Sample Selection at Fourth and Fifth Stages

After sample segments for the 2019 NSDUH were selected, specially trained field household listers visited the areas and compiled complete lists of all eligible DUs within the sample segment boundaries. These lists served as the frames for the fourth stage of sample selection.

The primary objective of the fourth stage of sample selection (listing units) was to select the minimum number of DUs needed in each segment to meet the targeted sample sizes for all age groups. Thus, listing unit sample sizes for the segment were determined using the age group with the largest sampling rate, which is referred to as the "driving" age group. Using 2010 census data adjusted to more recent projections from Claritas, state- and age-specific sampling rates were computed. These rates then were adjusted by (a) the segment's probability of selection; (b) the subsegmentation inflation factor,¹⁴ if any; (c) the probability of selecting a person in the age group (equal to the maximum, or 0.99, for the driving age group); and (d) an adjustment for the "maximum of two" rule.¹⁵ In addition to these factors, historical data from the 2017, 2018, and 2019 NSDUHs were used to compute predicted screening and interviewing response rate adjustments.¹⁶ The final adjusted sampling rate then was multiplied by the actual number of DUs found in the field during counting and listing activities. The product represents the segment's listing unit sample size.

¹⁴ Segments found to be very large in the field were partitioned into *subsegments*. Then one subsegment was chosen at random with probability proportional to the size to be fielded. In some cases, a second-level subsegmenting was required if the census totals used in the initial subsegmenting were off and the selected subsegment was still too large for listing. The subsegmentation inflation factor accounts for reducing the size of the segment.

¹⁵ Brewer's Selection Algorithm never allows for greater than two people per household to be chosen (Brewer, 1963, 1975; Cochran, 1977, pp. 261-263). Thus, sampling rates are adjusted to satisfy this constraint.

¹⁶ Data from quarters 1 and 2 of the 2019 NSDUH were used to compute screening and interview response rate adjustments for the quarter 3 and 4 samples.

Some constraints were put on the listing unit sample sizes. First, to ensure adequate samples for supplemental studies, the listing unit sample size could not exceed 100 per segment¹⁷ or half of the actual listing unit count. Next, for cost efficiency (i.e., to make traveling to and from the segment worth the expense), a minimum of five listing units were selected if five unused listing units remained in the segment.

Using a random start point and interval-based (systematic) selection, the actual listing units were selected from the segment frame. In 2015 and beyond, DUs that are selected from the overlap segments in the prior year are flagged as "used" and are not eligible for selection in the "current" year (i.e., two separate samples are selected with the complement of the prior year's sample serving as the DU frame in the "current" year). Individuals may be selected in consecutive years if they move and their new residence is selected the year after their original DU was sampled. No mechanism is currently in place for identifying duplicate people in a given year, but this number should be small given the restriction on DUs that were sampled in the previous year.

After DU selections were made, an FI visited each selected DU to obtain a roster of all people residing in the DU. Using the roster information obtained from an eligible member of the selected DU, 0, 1, or 2 people were selected for the survey. Sampling rates were preset by age group and state. Roster information was entered directly into the electronic screening instrument, which automatically implemented this fifth stage of selection based on the state and age group sampling parameters.

One advantage of using an electronic screening instrument in NSDUH is the ability to impose a more complicated person-level selection algorithm on the fifth stage of the NSDUH design. Similar to the 1999 through 2013 designs, one feature that was included in the 2014 through 2019 design is that any two survey-eligible people within a DU have some chance of being selected (i.e., all survey-eligible pairs of people had some nonzero chance of being selected). This design feature is of interest to NSDUH researchers because, for example, it allows analysts to examine how the drug use propensity of one individual in a family relates to the drug use propensity of another family member residing in the same DU (e.g., the relationship of drug use between a parent and his or her child). The pair sampling algorithm in NSDUH is based on the Chromy and Penne (2002) adaptation of the Brewer (1963, 1975) method for selecting samples of size two. Chromy and Penne (2002) also introduced a pair sampling parameter λ that governs the number of pairs selected. A simulation analysis was conducted to select the pair sampling parameter for the 2014 through 2022 NSDUHs (see the 2019 NSDUH sample design report; CBHSQ, 2020a).

¹⁷ Because of declining response rates, larger initial listing unit sample sizes were often required. A review of the data prompted an increase of the maximum listing unit sample size limit to 120 per segment for the 2020 NSDUH. See Chapter 7 for more details.

As in previous years, during the 2019 data collection period, if an FI encountered any new or missed DU on the premises of a sampled DU (e.g., a garage apartment), the new or missed dwelling was selected into the 2019 NSDUH. However, unlike the 2005 through 2013 NSDUHs, the half-open interval (HOI) procedure¹⁸ was not implemented. An evaluation of 2010 NSDUH data found that the HOI procedure accounted for only 0.2 percent of the total DUs on the supplemented NSDUH frame (Iannacchione, McMichael, Shook-Sa, & Morton, 2012). Further, an analysis of cases added to the sample through the HOI procedure found that these respondents did not have an appreciable impact on the estimates (Cunningham et al., 2009). Excluding the HOI procedure decreases the burden on FIs and simplifies the screening process. This decrease in burden outweighs the small increase in coverage resulting from implementation of the HOI procedure. To minimize bias associated with large numbers of missed DUs, FIs were instructed to call their supervisors if they noticed large differences between the segment listing and what they encountered in the field. If large differences were noticed, then special "bust" procedures were implemented (see the 2019 NSDUH sample design report; CBHSQ, 2020a).

2.2.4 Creation of Variance Strata and Variance PSUs

The nature of the stratified, clustered sampling design requires that the design structure be taken into consideration when computing variances of survey estimates. Variance strata and variance PSUs¹⁹ were created to capture explicit stratification and to identify clustering. For the 2014 through 2022 NSDUHs, variance strata are defined at the SSR level, and each SSR is assigned to a different stratum every quarter in a pseudorandom fashion. Similar to the 2005 through 2013 definition of variance strata, the 2014 through 2022 definition also has the effect of increasing the number of degrees of freedom (*df*) for state-level estimates while preserving the number of degrees of freedom for national estimates (750). Each of the smallest sample states is in 48 different strata (12 SSRs × 4 quarters); therefore, 48 degrees of freedom are available for state estimates in these states. At the other extreme, the largest sample state, California, is in 144 strata (36 SSRs × 4 quarters) and therefore has 144 degrees of freedom for estimation. Two variance PSUs per year were defined within each variance stratum. Each variance PSU consists of four segments, one for each quarter of data collection. The first variance PSU consists of those segments that are "phasing out" or will not be used in the next survey year. The second variance PSU consists of those segments that are "phasing in" or will be fielded again the following year, thus constituting the 50 percent overlap between survey years.

Because census tracts, block groups, and segments are nested within variance PSUs, the variance contributions of all three sampling stages are covered by the variance estimation variables. Also, because one segment is selected per sampled PSU and SSU, the selection of census tracts and block groups at the first stages of selection may reduce variance by controlling the sample distribution and minimizing the chance of selecting neighboring and possibly similar segments within the same census tract or block group. All weighted statistical analyses for which variance estimates are needed should use the variance stratum and variance PSU variables to

¹⁸ In summary, the rules for the HOI technique state that, if a DU is selected and an FI observes any new or missed DUs between the selected DU and the DU appearing immediately after the selection on the counting and listing form, all new or missed dwellings falling in this interval will be selected. If a large number of new or missed DUs are encountered (greater than 10), a sample of the new or missing DUs is selected, and the sample weight is adjusted accordingly.

¹⁹ Variance PSUs were previously called "variance replicates."

identify nesting. Stratification reduces variances, while clustering increases them. Ignoring these design elements may produce standard errors that lead to false-positive or false-negative test outcomes. Variance estimates can be computed using a clustered data analysis software package such as SUDAAN[®] (CBHSQ, 2020a; RTI International, 2013).

3. Segment (Third-Stage) Sample Experience

As mentioned in Chapter 2, the third stage of selection for the 2014 through 2022 National Surveys on Drug Use and Health (NSDUHs) was area segments. To form segments within sampled second-stage sampling units (SSUs), adjacent census blocks were collapsed until the total number of dwelling units (DUs) within the area met the minimum requirement. In California, Florida, Georgia, Illinois, Michigan, New Jersey, New York, North Carolina, Ohio, Pennsylvania, Texas, and Virginia, this minimum size requirement was 250 DUs in urban areas and 200 DUs in rural areas. In the remaining states and the District of Columbia, the minimum requirement was 150 DUs in urban areas and 100 DUs in rural areas.

To control the geographic distribution of the sample, segments were sorted in the order they were formed (geographically within SSUs), and one segment was selected per sampled SSU using the probability proportional to composite size with minimal replacement sequential sampling method. As a result, 48 segments per state sampling region (SSR) were chosen for a total of 36,000 segments. Although only 20 segments per SSR or 15,000 segments total were needed to support the 2014 through 2017 NSDUHs, an additional 28 "reserve" segments were selected to serve as replacements when segment DUs are depleted, to support any supplemental studies embedded within NSDUH, or to extend the sample to the next decennial census. Out of the reserve sample, 20 segments per SSR are being used to support the 2018 through 2022 NSDUHs.

The 48 sampled segments per SSR were randomly assigned to survey years by drawing equal probability subsamples of four segments. The first subsample of segments was assigned to the 2014 NSDUH and constituted the panel of segments to be used for 2014 only. The second subsample of segments was assigned to the 2014 NSDUH and was used again in the 2015 survey. The third subsample of segments was assigned to the 2015 NSDUH and was used again in the 2016 survey. This process continued for the remaining 36 sampled segments. Within each subsample, segments were assigned to survey quarters 1 through 4 in the order that they were selected.

Using the survey year and quarter assignments, a segment identification number (SEGID) then was assigned. The first two digits of the SEGID are the state abbreviation, the second two digits are for the SSR within the state, and the last two digits are called the "segment suffix," with the next-to-last digit being the panel identifier and the last digit being the original quarter assignment.

Because segments were selected with minimal replacement, some segments may have been selected more than once. Among the 30,000 segments selected for the 9-year study, 28,608 (95.4 percent) were unique. [Table 3.1](#) lists the duplicated segments in the 2019 NSDUH sample. Panels B, D, and E segments in [Table 3.1](#) are not in the 2019 sample but are duplicates of segments in the 2019 sample. Because segments are randomly assigned to panels, duplicates may exist within or across panels. For example, AK01B1 and AK01F1 are in different panels, while DC05F3 and DC05F4 are in the same panel. The original segment (e.g., AK01B1 or DC05F3) is field enumerated, and the resulting DU frame is used each time the segment is fielded (e.g., for AK01B1 and AK01F1 and for DC05F3 and DC05F4). DUs that were selected in previous fieldings are ineligible for selection in subsequent fieldings. Within survey years, duplicate

segments may be assigned to the same or different quarters. If a segment is used twice in the same quarter, a double sample is selected from the first segment, then randomly split for analysis purposes.

Table 3.1 Duplicated Segments in the 2019 NSDUH Sample

| Original Segment | 2019 NSDUH Duplicate Segment |
|---------------------|------------------------------|
| AK01B1* | AK01F1 |
| DC02F3 | DC02G4 |
| DC05F3 | DC05F4 |
| DC06B2* | DC06F1 |
| DC06F3 ^a | DC06F4 |
| DC07E4 | DC07F4 |
| HI01F2 | HI01F3 |
| HI02B2* | HI02F3 |
| ND09E1* | ND09F1 |
| ND09B4* | ND09F4 |
| RI07D3* | RI07F2 |
| SC05F4 ^a | SC05G3 |
| SD02F2 | SD02F3 |
| VT04F2 | VT04F3 |
| VT04F2 | VT04G3 |
| VT07F2 ^a | VT07F3 |
| VT07F2 ^a | VT07G1 |
| VT07F2 ^a | VT07G2 |

*Segment is not in the 2019 NSDUH sample but is a duplicate of a segment in the 2019 NSDUH sample.

^aOriginal segment is also a duplicate of an earlier segment.

During field enumeration, a small number of segments are switched with another segment in the same SSR and panel due to difficult conditions during the winter months. In general, quarter 1 segments are switched with quarter 2 segments, and quarter 3 segments are switched with quarter 4 segments. In 2019, one quarter 1 segment (SD12G1) was switched with a quarter 3 segment (SD12G3) because the quarter 2 segment also had winter access problems. [Table 3.2](#) lists the quarter switches in the 2019 NSDUH sample.

Table 3.2 2019 NSDUH Segment Quarter Switches in the Field

| Quarters 1 and 2 | | Quarters 3 and 4 | |
|-------------------------------|---------------------------------|-------------------------------|---------------------------------|
| Original Segment in Quarter 1 | Switched Segment from Quarter 2 | Original Segment in Quarter 3 | Switched Segment from Quarter 4 |
| AK02G1 | AK02G2 | AK03F3 | AK03F4 |
| AK04G1 | AK04G2 | AK10F3 | AK10F4 |
| AK11F1 | AK11F2 | AR09G3 | AR09G4 |
| AK11G1 | AK11G2 | CO09G3 | CO09G4 |
| AL04G1 | AL04G2 | CO11F3 | CO11F4 |
| AZ11G1 | AZ11G2 | CO12F3 | CO12F4 |
| CO12G1 | CO12G2 | ID02F3 | ID02F4 |
| ID05F1 | ID05F2 | ID05F3 | ID05F4 |
| MO12G1 | MO12G2 | MO01G3 | MO01G4 |
| MS08G1 | MS08G2 | MT02F3 | MT02F4 |
| MS09G1 | MS09G2 | MT09F3 | MT09F4 |
| MT01F1 | MT01F2 | MT10G3 | MT10G4 |
| MT01G1 | MT01G2 | MT12F3 | MT12F4 |
| MT04F1 | MT04F2 | MT12G3 | MT12G4 |
| MT09F1 | MT09F2 | ND07F3 | ND07F4 |
| ND01G1 | ND01G2 | ND11F3 | ND11F4 |
| ND03F1 | ND03F2 | NM01F3 | NM01F4 |
| ND07G1 | ND07G2 | NM03F3 | NM03F4 |
| ND10F1 | ND10F2 | NM05F3 | NM05F4 |
| ND11F1 | ND11F2 | NM10F3 | NM10F4 |
| ND12F1 | ND12F2 | NM11G3 | NM11G4 |
| NM06F1 | NM06F2 | NV04F3 | NV04F4 |
| NV01G1 | NV01G2 | SD01F3 | SD01F4 |
| NV03F1 | NV03F2 | SD02F3 | SD02F4 |
| NV05F1 | NV05F2 | SD08F3 | SD08F4 |
| OR02F1 | OR02F2 | TN09G3 | TN09G4 |
| OR03F1 | OR03F2 | UT10F3 | UT10F4 |
| SD10F1 | SD10F2 | WA12F3 | WA12F4 |
| SD10G1 | SD10G2 | WY04F3 | WY04F4 |
| SD11G1 | SD11G2 | WY04G3 | WY04G4 |
| SD12G1 | SD12G3 | WY05F3 | WY05F4 |
| TX02F1 | TX02F2 | WY07G3 | WY07G4 |
| UT05G1 | UT05G2 | WY12F3 | WY12F4 |
| WA11F1 | WA11F2 | | |
| WY05F1 | WY05F2 | | |
| WY05G1 | WY05G2 | | |
| WY07F1 | WY07F2 | | |
| WY10F1 | WY10F2 | | |
| WY10G1 | WY10G2 | | |

This page intentionally left blank

4. Dwelling Unit (Fourth-Stage) Sample Experience

The process by which the dwelling unit (DU) frame is constructed is called "counting and listing." In summary, a certified lister visits the selected area and lists a detailed and accurate address (or description, if no address is available) for each DU within the segment boundaries. Using a paper-based format, lists of DUs constructed during counting and listing are entered into a database and serve as the frame from which the fourth-stage sample is drawn.

As described in Section 2.2.3, after the DU frame was constructed, the next step was to determine the minimum number of DUs to select for each segment to meet the targeted sample sizes for all age groups. This sample size determination was performed on a quarterly basis to take advantage of both segment differences and, if necessary, make adjustments to design parameters (e.g., to use the most recent response rate experience). [Table 4.1](#) provides the number of DUs that were enumerated during the counting and listing process and the number of DUs that were sampled. After accounting for anticipated screening and interview response rates using historical National Survey on Drug Use and Health (NSDUH) data, an average of 40.8 sample dwelling units (SDUs) were selected per segment. The number of SDUs per segment varied by state according to the state's sample size, number of segments, and anticipated response rates.

Table 4.1 2019 NSDUH Segment and Dwelling Unit Summary

| State | Total Segments | Listed DUs | SDUs | SDUs per Segment | Added DUs | Percent Increase in DUs | Total Selected DUs |
|-------------------------|----------------|------------|---------|------------------|-----------|-------------------------|--------------------|
| Total Population | 6,000 | 1,491,538 | 244,805 | 40.80 | 958 | 0.39 | 245,763 |
| Alabama | 96 | 23,049 | 3,270 | 34.06 | 17 | 0.52 | 3,287 |
| Alaska | 96 | 20,340 | 3,583 | 37.32 | 39 | 1.09 | 3,622 |
| Arizona | 96 | 23,129 | 3,471 | 36.16 | 3 | 0.09 | 3,474 |
| Arkansas | 96 | 21,520 | 3,027 | 31.53 | 4 | 0.13 | 3,031 |
| California | 288 | 78,033 | 15,078 | 52.35 | 27 | 0.18 | 15,105 |
| Colorado | 96 | 21,423 | 3,409 | 35.51 | 8 | 0.23 | 3,417 |
| Connecticut | 96 | 22,925 | 3,392 | 35.33 | 9 | 0.27 | 3,401 |
| Delaware | 96 | 23,264 | 5,025 | 52.34 | 14 | 0.28 | 5,039 |
| District of Columbia | 96 | 30,130 | 7,746 | 80.69 | 22 | 0.28 | 7,768 |
| Florida | 240 | 71,218 | 11,776 | 49.07 | 29 | 0.25 | 11,805 |
| Georgia | 120 | 31,095 | 4,366 | 36.38 | 4 | 0.09 | 4,370 |
| Hawaii | 96 | 22,984 | 3,683 | 38.36 | 25 | 0.68 | 3,708 |
| Idaho | 96 | 19,915 | 2,625 | 27.34 | 10 | 0.38 | 2,635 |
| Illinois | 192 | 54,602 | 10,592 | 55.17 | 33 | 0.31 | 10,625 |
| Indiana | 96 | 22,350 | 3,486 | 36.31 | 17 | 0.49 | 3,503 |
| Iowa | 96 | 21,506 | 3,408 | 35.50 | 15 | 0.44 | 3,423 |
| Kansas | 96 | 22,014 | 3,003 | 31.28 | 8 | 0.27 | 3,011 |
| Kentucky | 96 | 23,907 | 3,067 | 31.95 | 3 | 0.10 | 3,070 |
| Louisiana | 96 | 22,505 | 2,846 | 29.65 | 3 | 0.11 | 2,849 |
| Maine | 96 | 20,476 | 4,444 | 46.29 | 21 | 0.47 | 4,465 |

(continued)

Table 4.1 2019 NSDUH Segment and Dwelling Unit Summary (continued)

| State | Total Segments | Listed DUs | SDUs | SDUs per Segment | Added DUs | Percent Increase in DUs | Total Selected DUs |
|----------------|----------------|------------|--------|------------------|-----------|-------------------------|--------------------|
| Maryland | 96 | 25,156 | 3,318 | 34.56 | 10 | 0.30 | 3,328 |
| Massachusetts | 96 | 22,359 | 4,170 | 43.44 | 21 | 0.50 | 4,191 |
| Michigan | 192 | 53,261 | 7,914 | 41.22 | 26 | 0.33 | 7,940 |
| Minnesota | 96 | 21,666 | 3,506 | 36.52 | 12 | 0.34 | 3,518 |
| Mississippi | 96 | 22,653 | 2,880 | 30.00 | 4 | 0.14 | 2,884 |
| Missouri | 96 | 22,232 | 3,354 | 34.94 | 3 | 0.09 | 3,357 |
| Montana | 96 | 19,738 | 3,710 | 38.65 | 37 | 1.00 | 3,747 |
| Nebraska | 96 | 22,311 | 3,078 | 32.06 | 8 | 0.26 | 3,086 |
| Nevada | 96 | 24,048 | 3,369 | 35.09 | 7 | 0.21 | 3,376 |
| New Hampshire | 96 | 20,676 | 3,916 | 40.79 | 61 | 1.56 | 3,977 |
| New Jersey | 120 | 33,704 | 5,307 | 44.23 | 10 | 0.19 | 5,317 |
| New Mexico | 96 | 20,871 | 2,951 | 30.74 | 3 | 0.10 | 2,954 |
| New York | 240 | 66,264 | 14,307 | 59.61 | 87 | 0.61 | 14,394 |
| North Carolina | 120 | 34,505 | 5,690 | 47.42 | 11 | 0.19 | 5,701 |
| North Dakota | 96 | 20,154 | 3,899 | 40.61 | 12 | 0.31 | 3,911 |
| Ohio | 192 | 51,588 | 8,265 | 43.05 | 30 | 0.36 | 8,295 |
| Oklahoma | 96 | 21,731 | 2,726 | 28.40 | 3 | 0.11 | 2,729 |
| Oregon | 96 | 21,898 | 3,206 | 33.40 | 44 | 1.37 | 3,250 |
| Pennsylvania | 192 | 53,864 | 9,334 | 48.61 | 41 | 0.44 | 9,375 |
| Rhode Island | 96 | 22,492 | 4,194 | 43.69 | 32 | 0.76 | 4,226 |
| South Carolina | 96 | 21,344 | 3,463 | 36.07 | 12 | 0.35 | 3,475 |
| South Dakota | 96 | 19,162 | 3,433 | 35.76 | 4 | 0.12 | 3,437 |
| Tennessee | 96 | 21,575 | 2,914 | 30.35 | 1 | 0.03 | 2,915 |
| Texas | 240 | 68,383 | 8,892 | 37.05 | 24 | 0.27 | 8,916 |
| Utah | 96 | 21,373 | 1,959 | 20.41 | 13 | 0.66 | 1,972 |
| Vermont | 96 | 23,729 | 4,007 | 41.74 | 45 | 1.12 | 4,052 |
| Virginia | 120 | 33,611 | 5,273 | 43.94 | 22 | 0.42 | 5,295 |
| Washington | 96 | 21,354 | 3,775 | 39.32 | 7 | 0.19 | 3,782 |
| West Virginia | 96 | 22,678 | 4,446 | 46.31 | 11 | 0.25 | 4,457 |
| Wisconsin | 96 | 20,622 | 3,231 | 33.66 | 19 | 0.59 | 3,250 |
| Wyoming | 96 | 20,151 | 3,021 | 31.47 | 27 | 0.89 | 3,048 |

DU = dwelling unit; SDU = sample dwelling unit.

To compensate for quarterly variations in response rates and yields, a sample partitioning procedure was implemented in all quarters. The entire sample of DUs still would be selected, but only certain percentages of the total would be released into the field. An initial percentage would be released in all segments at the beginning of the quarter. Based on interquarter work projections, additional percentages would be released 1 month into the quarter as needed and if field staff could handle the added workload. Each partitioning of the sample is a valid sample and helps manage the sample sizes by state without jeopardizing the validity of the study. Incidentally, a reserve DU sample of 20 percent also was selected within each selected segment, over and above the required quarterly sample, to allow for supplemental releases within each quarter. These releases usually occur as a result of lower than expected response rates but are also released for other reasons, including a large percentage of sample in controlled access areas and in college dormitories that are vacant during the summer months. In previous years, additional sample has also been released to compensate for sample lost to natural disasters and other emergency situations (e.g., following Hurricanes Katrina and Rita). Sample releases are made at the state level and do not target any particular age group. In each quarter, the DU sample was allocated to states in the following release percentages:

Release 1: 67 percent of entire sample (80/120, main sample + 20 percent reserve);
Release 2: 4 percent of entire sample (5/120, main sample + 20 percent reserve);
Release 3: 4 percent of entire sample (5/120, main sample + 20 percent reserve);
Release 4: 8 percent of entire sample (10/120, main sample + 20 percent reserve);
Release 5: 8 percent of entire sample (10/120, main sample + 20 percent reserve); and
Release 6: 8 percent of entire sample (10/120, main sample + 20 percent reserve).

As described in the 2019 NSDUH sample design report (Center for Behavioral Health Statistics and Quality [CBHSQ], 2020a), a weight adjustment is applied to all DUs within a segment to account for the partial release of sample. The DU release adjustment is equal to the inverse of the percentage of the sample that is released into the field. A summary of the quarterly sample sizes and percentages released is provided in [Table 4.2](#). If the release plan was implemented without changes, a percentage equal to 100/120 or 83.3 percent would be expected.

To ensure that nearly all DUs had a chance of selection and to minimize bias associated with incomplete frames, a check for missed DUs was implemented at most sampled DUs.²⁰ During the screening interview, the field interviewer (FI) asked the screening respondent about other units on the property of the sampled DU (e.g., a garage apartment). When found on the property of a sampled DU, the unlisted units became part of the sample (added DUs) and were considered "linked" to that DU. If the number of added DUs linked to any particular sample DU did not exceed 5, and if the number for the entire segment was less than or equal to 10, FIs were instructed to consider these DUs as part of their assignments. If either of these limits was exceeded, special subsampling procedures were implemented (see CBHSQ, 2020a).

In addition to checking for missed DUs at each sampled DU, FIs were instructed to call their supervisors if they noticed large differences in the segment listing and what they encountered in the field. If an FI identified 150 or more missed DUs in a segment or 50 or more missed DUs following any DU, special "bust" procedures were implemented to minimize bias associated with large numbers of missed DUs. The bust procedures involve selecting a subsample of the missed DUs and adding them to the FI's assignment; these procedures are described in more detail in CBHSQ (2020a). The total number of added DUs identified during the screening interview or added through the bust procedures is summarized in [Table 4.1](#). Overall, a 0.39 percent increase in sample was realized through the check for missed DUs. Larger increases (1 percent or more) in sample were realized in the states of Alaska, Montana, New Hampshire, Oregon, and Vermont. In Alaska, a large number of DUs were difficult or impossible to see during the listing stage (e.g., DUs on large properties or multiple units within a single structure). In rural Montana, several missed housing units were found on land behind a main housing unit. In New Hampshire, a large number of garage apartments were found, and single-family homes were converted into multiple units. In Oregon, several group quarters under construction at the time of listing were found to have a much larger number of units than were originally listed. Finally, in Vermont, there was a lot of new construction. The added DU information in [Table 4.1](#) will be used in the sample size calculations for future NSDUHs.

²⁰ The screening respondent was not asked about other units on the property or within the sampled DU in apartment buildings and other multiunit structures.

Table 4.2 2019 NSDUH Quarterly Dwelling Unit Sample Sizes and Percentages Released

| Region/State | Quarter 1 | | | Quarter 2 | | |
|-------------------------|------------|------------|---------|------------|------------|---------|
| | # Selected | # Released | Percent | # Selected | # Released | Percent |
| Total Population | 65,532 | 54,609 | 83.33 | 71,773 | 66,278 | 92.34 |
| Northeast | 14,372 | 11,985 | 83.39 | 15,810 | 14,897 | 94.23 |
| Connecticut | 936 | 779 | 83.23 | 943 | 941 | 99.79 |
| Maine | 1,098 | 916 | 83.42 | 1,274 | 1,274 | 100.00 |
| Massachusetts | 1,080 | 893 | 82.69 | 1,096 | 1,096 | 100.00 |
| New Hampshire | 920 | 766 | 83.26 | 1,047 | 1,046 | 99.90 |
| New Jersey | 1,605 | 1,340 | 83.49 | 1,801 | 1,501 | 83.34 |
| New York | 3,986 | 3,329 | 83.52 | 4,413 | 4,228 | 95.81 |
| Pennsylvania | 2,456 | 2,050 | 83.47 | 2,748 | 2,633 | 95.82 |
| Rhode Island | 1,116 | 930 | 83.33 | 1,231 | 1,128 | 91.63 |
| Vermont | 1,175 | 982 | 83.57 | 1,257 | 1,050 | 83.53 |
| Midwest | 14,924 | 12,451 | 83.43 | 16,215 | 15,752 | 97.14 |
| Illinois | 2,353 | 1,961 | 83.34 | 2,679 | 2,678 | 99.96 |
| Indiana | 1,014 | 846 | 83.43 | 938 | 900 | 95.95 |
| Iowa | 1,053 | 878 | 83.38 | 1,101 | 1,054 | 95.73 |
| Kansas | 781 | 649 | 83.10 | 855 | 855 | 100.00 |
| Michigan | 2,209 | 1,841 | 83.34 | 2,301 | 2,301 | 100.00 |
| Minnesota | 820 | 683 | 83.29 | 998 | 994 | 99.60 |
| Missouri | 823 | 688 | 83.60 | 940 | 940 | 100.00 |
| Nebraska | 792 | 663 | 83.71 | 874 | 873 | 99.89 |
| North Dakota | 1,028 | 866 | 84.24 | 1,162 | 1,162 | 100.00 |
| Ohio | 2,256 | 1,884 | 83.51 | 2,498 | 2,290 | 91.67 |
| South Dakota | 820 | 684 | 83.41 | 905 | 905 | 100.00 |
| Wisconsin | 975 | 808 | 82.87 | 964 | 800 | 82.99 |
| South | 21,511 | 17,899 | 83.21 | 23,470 | 21,522 | 91.70 |
| Alabama | 793 | 661 | 83.35 | 898 | 898 | 100.00 |
| Arkansas | 804 | 670 | 83.33 | 801 | 801 | 100.00 |
| Delaware | 1,222 | 1,017 | 83.22 | 1,460 | 1,396 | 95.62 |
| District of Columbia | 1,997 | 1,662 | 83.22 | 2,110 | 1,934 | 91.66 |
| Florida | 3,569 | 2,977 | 83.41 | 3,886 | 3,083 | 79.34 |
| Georgia | 1,314 | 1,086 | 82.65 | 1,401 | 1,165 | 83.15 |
| Kentucky | 846 | 700 | 82.74 | 920 | 842 | 91.52 |
| Louisiana | 885 | 735 | 83.05 | 866 | 648 | 74.83 |
| Maryland | 823 | 681 | 82.75 | 964 | 927 | 96.16 |
| Mississippi | 694 | 576 | 83.00 | 690 | 690 | 100.00 |
| North Carolina | 1,320 | 1,101 | 83.41 | 1,541 | 1,541 | 100.00 |
| Oklahoma | 914 | 760 | 83.15 | 928 | 657 | 70.80 |
| South Carolina | 894 | 742 | 83.00 | 902 | 902 | 100.00 |
| Tennessee | 688 | 572 | 83.14 | 759 | 759 | 100.00 |
| Texas | 2,280 | 1,900 | 83.33 | 2,528 | 2,528 | 100.00 |
| Virginia | 1,405 | 1,174 | 83.56 | 1,574 | 1,509 | 95.87 |
| West Virginia | 1,063 | 885 | 83.25 | 1,242 | 1,242 | 100.00 |
| West | 14,725 | 12,274 | 83.35 | 16,278 | 14,107 | 86.66 |
| Alaska | 906 | 761 | 84.00 | 1,049 | 1,005 | 95.81 |
| Arizona | 944 | 786 | 83.26 | 1,037 | 781 | 75.31 |
| California | 4,186 | 3,495 | 83.49 | 4,738 | 3,945 | 83.26 |
| Colorado | 794 | 660 | 83.12 | 927 | 888 | 95.79 |
| Hawaii | 1,065 | 889 | 83.47 | 1,071 | 927 | 86.55 |
| Idaho | 740 | 623 | 84.19 | 838 | 697 | 83.17 |
| Montana | 1,078 | 896 | 83.12 | 1,161 | 1,161 | 100.00 |
| Nevada | 935 | 781 | 83.53 | 988 | 823 | 83.30 |
| New Mexico | 797 | 663 | 83.19 | 856 | 712 | 83.18 |
| Oregon | 1,047 | 857 | 81.85 | 1,084 | 812 | 74.91 |
| Utah | 531 | 440 | 82.86 | 562 | 540 | 96.09 |
| Washington | 887 | 740 | 83.43 | 1,051 | 1,051 | 100.00 |
| Wyoming | 815 | 683 | 83.80 | 916 | 765 | 83.52 |

(continued)

Table 4.2 2019 NSDUH Quarterly Dwelling Unit Sample Sizes and Percentages Released (continued)

| Region/State | Quarter 3 | | | Quarter 4 | | |
|-------------------------|------------|------------|---------|------------|------------|---------|
| | # Selected | # Released | Percent | # Selected | # Released | Percent |
| Total Population | 74,026 | 64,393 | 86.99 | 68,246 | 59,525 | 87.22 |
| Northeast | 15,840 | 13,910 | 87.82 | 14,356 | 12,279 | 85.53 |
| Connecticut | 1,153 | 1,056 | 91.59 | 925 | 616 | 66.59 |
| Maine | 1,223 | 1,223 | 100.00 | 1,243 | 1,031 | 82.94 |
| Massachusetts | 1,134 | 1,134 | 100.00 | 1,089 | 1,047 | 96.14 |
| New Hampshire | 1,253 | 1,253 | 100.00 | 1,135 | 851 | 74.98 |
| New Jersey | 1,749 | 1,308 | 74.79 | 1,630 | 1,158 | 71.04 |
| New York | 4,059 | 3,713 | 91.48 | 3,656 | 3,037 | 83.07 |
| Pennsylvania | 2,747 | 2,169 | 78.96 | 2,485 | 2,482 | 99.88 |
| Rhode Island | 1,250 | 1,097 | 87.76 | 1,080 | 1,039 | 96.20 |
| Vermont | 1,272 | 957 | 75.24 | 1,113 | 1,018 | 91.46 |
| Midwest | 17,009 | 14,414 | 84.74 | 16,132 | 14,552 | 90.21 |
| Illinois | 2,903 | 2,899 | 99.86 | 3,185 | 3,054 | 95.89 |
| Indiana | 963 | 888 | 92.21 | 884 | 852 | 96.38 |
| Iowa | 1,154 | 864 | 74.87 | 925 | 612 | 66.16 |
| Kansas | 888 | 703 | 79.17 | 873 | 796 | 91.18 |
| Michigan | 2,349 | 1,771 | 75.39 | 2,183 | 2,001 | 91.66 |
| Minnesota | 999 | 918 | 91.89 | 995 | 911 | 91.56 |
| Missouri | 1,073 | 941 | 87.70 | 995 | 785 | 78.89 |
| Nebraska | 991 | 864 | 87.18 | 961 | 678 | 70.55 |
| North Dakota | 1,244 | 1,031 | 82.88 | 1,012 | 840 | 83.00 |
| Ohio | 2,639 | 1,871 | 70.90 | 2,316 | 2,220 | 95.85 |
| South Dakota | 950 | 911 | 95.89 | 933 | 933 | 100.00 |
| Wisconsin | 856 | 753 | 87.97 | 870 | 870 | 100.00 |
| South | 24,428 | 21,703 | 88.84 | 22,778 | 19,601 | 86.05 |
| Alabama | 958 | 841 | 87.79 | 988 | 870 | 88.06 |
| Arkansas | 836 | 836 | 100.00 | 720 | 720 | 100.00 |
| Delaware | 1,834 | 1,379 | 75.19 | 1,562 | 1,233 | 78.94 |
| District of Columbia | 2,238 | 2,238 | 100.00 | 2,295 | 1,912 | 83.31 |
| Florida | 4,044 | 3,193 | 78.96 | 3,565 | 2,523 | 70.77 |
| Georgia | 1,475 | 1,226 | 83.12 | 1,338 | 889 | 66.44 |
| Kentucky | 919 | 687 | 74.76 | 838 | 838 | 100.00 |
| Louisiana | 872 | 727 | 83.37 | 736 | 736 | 100.00 |
| Maryland | 952 | 952 | 100.00 | 956 | 758 | 79.29 |
| Mississippi | 748 | 748 | 100.00 | 866 | 866 | 100.00 |
| North Carolina | 1,618 | 1,618 | 100.00 | 1,715 | 1,430 | 83.38 |
| Oklahoma | 843 | 702 | 83.27 | 728 | 607 | 83.38 |
| South Carolina | 992 | 992 | 100.00 | 827 | 827 | 100.00 |
| Tennessee | 816 | 816 | 100.00 | 799 | 767 | 95.99 |
| Texas | 2,506 | 2,195 | 87.59 | 2,269 | 2,269 | 100.00 |
| Virginia | 1,495 | 1,432 | 95.79 | 1,324 | 1,158 | 87.46 |
| West Virginia | 1,282 | 1,121 | 87.44 | 1,252 | 1,198 | 95.69 |
| West | 16,749 | 14,366 | 85.77 | 14,980 | 13,093 | 87.40 |
| Alaska | 1,137 | 993 | 87.34 | 1,042 | 824 | 79.08 |
| Arizona | 1,221 | 960 | 78.62 | 1,031 | 944 | 91.56 |
| California | 4,768 | 3,977 | 83.41 | 4,393 | 3,661 | 83.34 |
| Colorado | 960 | 960 | 100.00 | 901 | 901 | 100.00 |
| Hawaii | 1,053 | 987 | 93.73 | 923 | 880 | 95.34 |
| Idaho | 740 | 592 | 80.00 | 713 | 713 | 100.00 |
| Montana | 1,218 | 903 | 74.14 | 1,065 | 750 | 70.42 |
| Nevada | 925 | 884 | 95.57 | 881 | 881 | 100.00 |
| New Mexico | 928 | 733 | 78.99 | 843 | 843 | 100.00 |
| Oregon | 1,136 | 941 | 82.83 | 896 | 596 | 66.52 |
| Utah | 527 | 507 | 96.20 | 494 | 472 | 95.55 |
| Washington | 1,204 | 1,151 | 95.60 | 1,003 | 833 | 83.05 |
| Wyoming | 932 | 778 | 83.48 | 795 | 795 | 100.00 |

This page intentionally left blank

5. Person (Fifth-Stage) Sample Experience

Compared with previous designs, the 2014 through 2022 National Survey on Drug Use and Health (NSDUH) design places more sample in the 26 or older age groups to estimate substance use and related mental health measures more accurately among the aging drug-using population. As noted previously, the target national sample size of 67,507 was distributed across five age groups as follows: 25 percent for youths aged 12 to 17, 25 percent for young adults aged 18 to 25, 15 percent for adults aged 26 to 34, 20 percent for adults aged 35 to 49, and 15 percent for adults aged 50 or older. Further, the sample was designed to yield desired sample sizes in each state as described in Section 2.2.1. [Table 5.1](#) displays the desired and achieved sample yields by age group and state. In general, the sample allocation and sample size management procedures were effective at achieving the numerous sample size targets.

Table 5.1 2019 NSDUH Yields, by Age Group and State

| Age Group and State | Targeted | Achieved | Percent Difference |
|----------------------|----------|----------|--------------------|
| Total | 67,507 | 67,625 | 0.17 |
| 12-17 | 16,877 | 16,894 | 0.10 |
| 18-25 | 16,877 | 16,665 | -1.26 |
| 26-34 | 10,126 | 10,051 | -0.74 |
| 35-49 | 13,501 | 13,579 | 0.58 |
| 50+ | 10,126 | 10,436 | 3.06 |
| Alabama | 960 | 978 | 1.88 |
| Alaska | 960 | 988 | 2.92 |
| Arizona | 960 | 932 | -2.92 |
| Arkansas | 960 | 936 | -2.50 |
| California | 4,560 | 4,677 | 2.57 |
| Colorado | 960 | 932 | -2.92 |
| Connecticut | 960 | 985 | 2.60 |
| Delaware | 960 | 935 | -2.60 |
| District of Columbia | 960 | 961 | 0.10 |
| Florida | 3,300 | 3,331 | 0.94 |
| Georgia | 1,500 | 1,519 | 1.27 |
| Hawaii | 967 | 1,007 | 4.14 |
| Kauai County | 67 | 67 | 0.00 |
| Remainder of Hawaii | 900 | 940 | 4.44 |
| Idaho | 960 | 960 | 0.00 |
| Illinois | 2,400 | 2,361 | -1.63 |
| Indiana | 960 | 974 | 1.46 |
| Iowa | 960 | 968 | 0.83 |
| Kansas | 960 | 951 | -0.94 |
| Kentucky | 960 | 978 | 1.88 |
| Louisiana | 960 | 936 | -2.50 |
| Maine | 960 | 961 | 0.10 |

(continued)

Table 5.1 2019 NSDUH Yields, by Age Group and State (continued)

| Age Group and State | Targeted | Achieved | Percent Difference |
|---------------------|----------|----------|--------------------|
| Maryland | 960 | 959 | -0.10 |
| Massachusetts | 960 | 955 | -0.52 |
| Michigan | 2,400 | 2,497 | 4.04 |
| Minnesota | 960 | 962 | 0.21 |
| Mississippi | 960 | 1,041 | 8.44 |
| Missouri | 960 | 968 | 0.83 |
| Montana | 960 | 963 | 0.31 |
| Nebraska | 960 | 999 | 4.06 |
| Nevada | 960 | 978 | 1.88 |
| New Hampshire | 960 | 961 | 0.10 |
| New Jersey | 1,500 | 1,490 | -0.67 |
| New Mexico | 960 | 917 | -4.48 |
| New York | 3,300 | 3,260 | -1.21 |
| North Carolina | 1,500 | 1,572 | 4.80 |
| North Dakota | 960 | 938 | -2.29 |
| Ohio | 2,400 | 2,446 | 1.92 |
| Oklahoma | 960 | 952 | -0.83 |
| Oregon | 960 | 954 | -0.63 |
| Pennsylvania | 2,400 | 2,386 | -0.58 |
| Rhode Island | 960 | 930 | -3.13 |
| South Carolina | 960 | 901 | -6.15 |
| South Dakota | 960 | 935 | -2.60 |
| Tennessee | 960 | 987 | 2.81 |
| Texas | 3,300 | 3,238 | -1.88 |
| Utah | 960 | 947 | -1.35 |
| Vermont | 960 | 944 | -1.67 |
| Virginia | 1,500 | 1,474 | -1.73 |
| Washington | 960 | 945 | -1.56 |
| West Virginia | 960 | 968 | 0.83 |
| Wisconsin | 960 | 958 | -0.21 |
| Wyoming | 960 | 930 | -3.13 |

Some unique challenges arose with the increase in sample allocated to the 26 or older age groups. Respondents in these age groups are more challenging to find at home and persuade to take the time to participate in the study. In 2019, the unweighted interview response rate was 71.49 percent for individuals aged 12 to 17, 66.21 percent for individuals aged 18 to 25, and 64.63 percent for individuals aged 26 or older (Center for Behavioral Health Statistics and Quality [CBHSQ], 2020b). Within the 26 or older age group, individuals aged 50 or older had the lowest unweighted interview response rate (64.33 percent). The higher percentage of respondents within the 26 or older age group required more visits to attempt completion of pending cases and for refusal conversion efforts. Anecdotally, individuals aged 50 or older are more likely to live in controlled access situations, making it more challenging to complete screening and interviewing for these cases.

In addition to the shift in sample to the older age groups, a new pair sampling parameter was selected for the 2014 through 2022 NSDUHs. The pair sampling algorithm in NSDUH is

based on the Chromy and Penne (2002) adaptation of the Brewer (1963, 1975) method for selecting samples of size 2 as a means of selecting samples of 0, 1, or 2 people within a selected dwelling unit containing at least 1 eligible person. Chromy and Penne's adaptation includes a pair sampling parameter, λ , that governs the number of pairs selected. Simulation analyses resulted in the selection of $\lambda = 0.50$ for the 2002 to 2013 NSDUH sample designs because this selection increased the number of pairs by about 20 percent (relative to the selection of $\lambda = 0.00$) with only a moderate impact on the response rates by age group.

For the 2014 through 2022 NSDUHs, simulation analyses based on the 2012 screening data, modified to reflect the 2014 through 2022 age group sample proportions, were conducted, and $\lambda = 0.25$ was selected (CBHSQ, 2018). As a result, fewer pairs were projected to be selected in the 2014 through 2022 NSDUHs than were selected in the 2002 through 2013 NSDUHs. However, as a result of increasing the older adult sample, a λ value of 0.25 yielded a large projected number of adolescent-adult pairs in 2014 through 2022 when compared with earlier years.

Prior to the 2019 NSDUH, pair projection models were updated using 2017 screening data. [Tables 5.2](#) and [5.3](#) provide projected and observed pair selection counts and response rates, respectively, by age group pairs for the three age groups: 12 to 17, 18 to 25, and 26 or older. Observed selection counts in 2019 were larger than their projected counterparts overall and in most age group pairs. This is partially because the observed counts are based on an overall sample of 67,625 interviews and the projected counts were normalized to yield 67,507 interviews. Further, response rates were lower than anticipated, requiring more selections to achieve the desired sample.

Table 5.2 2019 NSDUH Projected and Observed Pair Selection Counts, by Age Group Pairs (Three Age Groups: 12 to 17, 18 to 25, and 26 or Older)

| Age Group Pair | Projected Count ($\lambda = 0.25$) | Observed Count ($\lambda = 0.25$) |
|----------------|--------------------------------------|-------------------------------------|
| 12+, 12+ | 28,889 | 29,818 |
| 12-17, 12-17 | 3,245 | 3,400 |
| 12-17, 18-25 | 2,670 | 2,779 |
| 12-17, 26+ | 8,343 | 8,514 |
| 18-25, 18-25 | 4,220 | 4,533 |
| 18-25, 26+ | 4,911 | 5,152 |
| 26+, 26+ | 5,499 | 5,440 |

Note: Projected pair selection counts are based on 2017 NSDUH screening data.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2017 and 2019.

Table 5.3 2019 NSDUH Projected and Observed Pair Response Rates, by Age Group Pairs (Three Age Groups: 12 to 17, 18 to 25, and 26 or Older)

| Age Group Pair | Projected Response Rate ($\lambda = 0.25$) | Observed Response Rate ($\lambda = 0.25$) |
|----------------|---|--|
| 12+, 12+ | 61.4 | 57.5 |
| 12-17, 12-17 | 72.6 | 69.6 |
| 12-17, 18-25 | 65.6 | 61.4 |
| 12-17, 26+ | 65.9 | 61.9 |
| 18-25, 18-25 | 59.2 | 53.9 |
| 18-25, 26+ | 55.8 | 51.7 |
| 26+, 26+ | 52.7 | 49.5 |

Note: Projected pair response rates are based on 2017 NSDUH screener data.

Note: Observed response rates are based on questionnaire age.

Note: A pair response requires both members of the age group pair to respond.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2017 and 2019.

Departures from the planned sample sizes occur for several reasons, including sampling variability, access problems (e.g., in gated communities or college dormitories), and departures from expected response rates. [Table 5.4](#) provides weighted screening and interview response rates by state for the 2019 NSDUH.²¹ These rates will be used to fine-tune the sample size calculations for the 2020 NSDUH.

²¹ The NSDUH screening response rate is calculated using the American Association for Public Opinion Research (AAPOR) Response Rate 1 (RR1), and the interview response rate is calculated using AAPOR's Response Rate 2 (RR2) (AAPOR, 2016, p. 61).

Table 5.4 2019 NSDUH Weighted Screening and Interview Response Rates, by State

| State | Total Selected DUs | Total Eligible DUs | Weighted DU Eligibility Rate | Total Completed Screeners | Weighted DU Screening Response Rate | Total Selected | Total Respondents | Weighted Interview Response Rate | Weighted Overall Response Rate |
|----------------------|--------------------|--------------------|------------------------------|---------------------------|-------------------------------------|----------------|-------------------|----------------------------------|--------------------------------|
| Overall | 245,763 | 210,244 | 85.46 | 148,023 | 70.50 | 101,509 | 67,625 | 64.92 | 45.77 |
| Alabama | 3,287 | 2,616 | 79.23 | 2,213 | 84.87 | 1,424 | 978 | 66.95 | 56.82 |
| Alaska | 3,622 | 2,858 | 78.15 | 2,066 | 72.18 | 1,468 | 988 | 66.36 | 47.90 |
| Arizona | 3,474 | 2,943 | 80.40 | 1,978 | 67.41 | 1,408 | 932 | 65.42 | 44.09 |
| Arkansas | 3,031 | 2,374 | 77.97 | 2,001 | 83.99 | 1,273 | 936 | 71.78 | 60.29 |
| California | 15,105 | 14,044 | 92.85 | 8,774 | 61.66 | 7,274 | 4,677 | 60.87 | 37.53 |
| Colorado | 3,417 | 2,905 | 81.02 | 1,995 | 68.81 | 1,450 | 932 | 64.29 | 44.24 |
| Connecticut | 3,401 | 3,087 | 90.89 | 2,080 | 67.13 | 1,580 | 985 | 59.36 | 39.85 |
| Delaware | 5,039 | 4,227 | 75.30 | 2,619 | 61.54 | 1,560 | 935 | 59.35 | 36.52 |
| District of Columbia | 7,768 | 6,704 | 87.10 | 3,400 | 48.17 | 1,244 | 961 | 74.87 | 36.07 |
| Florida | 11,805 | 9,736 | 78.23 | 7,032 | 70.76 | 4,696 | 3,331 | 67.05 | 47.44 |
| Georgia | 4,370 | 3,759 | 85.92 | 2,947 | 78.40 | 2,062 | 1,519 | 72.28 | 56.67 |
| Hawaii | 3,708 | 3,190 | 85.38 | 2,076 | 65.02 | 1,540 | 1,007 | 65.07 | 42.31 |
| Idaho | 2,635 | 2,258 | 83.70 | 1,774 | 78.86 | 1,350 | 960 | 70.54 | 55.63 |
| Illinois | 10,625 | 9,412 | 88.78 | 5,602 | 58.84 | 4,157 | 2,361 | 54.31 | 31.96 |
| Indiana | 3,503 | 2,972 | 84.96 | 2,060 | 69.26 | 1,445 | 974 | 66.00 | 45.71 |
| Iowa | 3,423 | 2,943 | 86.07 | 2,286 | 77.44 | 1,422 | 968 | 68.30 | 52.89 |
| Kansas | 3,011 | 2,500 | 82.55 | 1,947 | 77.84 | 1,403 | 951 | 67.44 | 52.50 |
| Kentucky | 3,070 | 2,570 | 83.42 | 2,005 | 78.07 | 1,438 | 978 | 66.96 | 52.27 |
| Louisiana | 2,849 | 2,292 | 80.30 | 1,967 | 85.79 | 1,329 | 936 | 67.54 | 57.94 |
| Maine | 4,465 | 3,455 | 75.95 | 2,723 | 78.51 | 1,487 | 961 | 63.23 | 49.64 |
| Maryland | 3,328 | 2,991 | 89.78 | 2,000 | 67.01 | 1,362 | 959 | 69.88 | 46.83 |
| Massachusetts | 4,191 | 3,825 | 90.94 | 2,434 | 64.04 | 1,633 | 955 | 55.84 | 35.76 |
| Michigan | 7,940 | 6,736 | 84.31 | 5,231 | 77.62 | 3,496 | 2,497 | 70.19 | 54.49 |
| Minnesota | 3,518 | 3,117 | 88.78 | 2,299 | 73.14 | 1,410 | 962 | 68.88 | 50.38 |
| Mississippi | 2,884 | 2,393 | 82.38 | 1,976 | 83.25 | 1,514 | 1,041 | 67.39 | 56.10 |
| Missouri | 3,357 | 2,768 | 80.07 | 2,184 | 78.81 | 1,422 | 968 | 66.50 | 52.41 |
| Montana | 3,747 | 3,117 | 82.00 | 2,437 | 78.35 | 1,501 | 963 | 65.51 | 51.33 |
| Nebraska | 3,086 | 2,654 | 86.34 | 2,013 | 75.75 | 1,371 | 999 | 73.34 | 55.55 |

(continued)

Table 5.4 2019 NSDUH Weighted Screening and Interview Response Rates by State (continued)

| State | Total Selected DUs | Total Eligible DUs | Weighted DU Eligibility Rate | Total Completed Screeners | Weighted DU Screening Response Rate | Total Selected | Total Respondents | Weighted Interview Response Rate | Weighted Overall Response Rate |
|----------------|--------------------|--------------------|------------------------------|---------------------------|-------------------------------------|----------------|-------------------|----------------------------------|--------------------------------|
| Nevada | 3,376 | 3,094 | 91.62 | 1,967 | 63.68 | 1,401 | 978 | 66.40 | 42.28 |
| New Hampshire | 3,977 | 3,324 | 81.72 | 2,422 | 72.61 | 1,537 | 961 | 63.83 | 46.34 |
| New Jersey | 5,317 | 4,802 | 90.24 | 3,227 | 67.46 | 2,461 | 1,490 | 59.57 | 40.18 |
| New Mexico | 2,954 | 2,380 | 80.44 | 1,758 | 73.95 | 1,309 | 917 | 68.44 | 50.61 |
| New York | 14,394 | 12,697 | 87.92 | 7,419 | 57.76 | 5,485 | 3,260 | 56.12 | 32.42 |
| North Carolina | 5,701 | 4,794 | 83.73 | 3,234 | 67.54 | 2,181 | 1,572 | 70.71 | 47.76 |
| North Dakota | 3,911 | 3,200 | 81.84 | 2,391 | 75.04 | 1,544 | 938 | 59.30 | 44.50 |
| Ohio | 8,295 | 7,188 | 86.21 | 5,206 | 72.53 | 3,626 | 2,446 | 66.82 | 48.47 |
| Oklahoma | 2,729 | 2,325 | 85.10 | 1,801 | 77.40 | 1,393 | 952 | 68.31 | 52.87 |
| Oregon | 3,250 | 2,943 | 90.32 | 2,286 | 77.34 | 1,417 | 954 | 66.80 | 51.67 |
| Pennsylvania | 9,375 | 8,287 | 88.29 | 5,684 | 68.75 | 3,627 | 2,386 | 65.37 | 44.95 |
| Rhode Island | 4,226 | 3,613 | 81.63 | 2,317 | 64.01 | 1,432 | 930 | 62.09 | 39.75 |
| South Carolina | 3,475 | 2,796 | 80.02 | 1,996 | 71.48 | 1,257 | 901 | 71.55 | 51.15 |
| South Dakota | 3,437 | 2,747 | 79.30 | 2,182 | 79.51 | 1,376 | 935 | 69.51 | 55.26 |
| Tennessee | 2,915 | 2,451 | 84.29 | 1,915 | 78.21 | 1,419 | 987 | 67.93 | 53.13 |
| Texas | 8,916 | 7,523 | 84.33 | 5,652 | 74.87 | 4,666 | 3,238 | 66.61 | 49.87 |
| Utah | 1,972 | 1,710 | 86.71 | 1,271 | 74.58 | 1,335 | 947 | 70.23 | 52.38 |
| Vermont | 4,052 | 3,172 | 77.79 | 2,445 | 77.01 | 1,461 | 944 | 64.36 | 49.56 |
| Virginia | 5,295 | 4,531 | 85.58 | 3,269 | 71.44 | 2,127 | 1,474 | 67.04 | 47.90 |
| Washington | 3,782 | 3,306 | 87.05 | 2,558 | 77.58 | 1,475 | 945 | 63.17 | 49.01 |
| West Virginia | 4,457 | 3,671 | 82.39 | 2,785 | 75.87 | 1,610 | 968 | 58.99 | 44.76 |
| Wisconsin | 3,250 | 2,768 | 83.66 | 2,196 | 79.19 | 1,400 | 958 | 68.06 | 53.90 |
| Wyoming | 3,048 | 2,476 | 81.26 | 1,923 | 78.11 | 1,251 | 930 | 76.14 | 59.48 |

DU = dwelling unit.

6. Sampling Error

6.1 Computation of Relative Standard Errors and Design Effects

Several objectives were set for calculating relative standard errors (RSEs) and design effects (DEFFs) for the 2019 National Survey on Drug Use and Health (NSDUH). One objective was to provide a mechanism for comparing the expected precision of the 2019 design with the precision actually obtained. A second objective was to have a record of the magnitudes of the DEFFs for a future redesign of the survey.

The RSE of a domain d prevalence estimate \hat{p}_d is the standard error (SE) of the estimate divided by the estimate, that is,

$$RSE(\hat{p}_d) = SE(\hat{p}_d) / \hat{p}_d . \quad (1)$$

The DEFF for a prevalence estimate is its variance divided by the variance that would be observed if simple random sampling (SRS) had been used:

$$DEFF(\hat{p}_d) = \frac{\text{VAR}(\hat{p}_d)}{\text{VAR}_{\text{SRS}}(\hat{p}_d)} . \quad (2)$$

Hence, the SE of the estimated prevalence can be approximated as follows:

$$SE(\hat{p}_d) \cong [DEFF(d)\hat{p}_d(1 - \hat{p}_d) / n_d]^{1/2} , \quad (3)$$

where n_d and $DEFF(d)$ are the sample size and median (or mean, as the case may be) DEFF of domain d , respectively.

As noted previously, the DEFF is the ratio of the design-based variance estimate divided by the variance estimate that would have been obtained from an SRS of the same size. Therefore, the DEFF summarizes the effects of stratification, clustering, and unequal weighting on the variance of a complex sample design. Because clustering and unequal weighting are expected to increase the variance and generally dominate the stratification effect, the DEFF is expected to be greater than 1 in most instances. However, median (or mean) DEFFs were sometimes less than 1 for prevalence rates near 0.

Note that the DEFF is based on the with-replacement (wr) variance estimate as obtained from SUDAAN, which properly accounts for clustering, stratification, and unequal weighting, assuming the first-stage sampling units were selected with replacement. They were not in fact selected with replacement, but the assumption simplifies variance estimation immensely while only slightly and conservatively biasing the results upward (RTI International, 2013). Prior to the 2000 survey, a more complex method of variance estimation was used; however, it was decided that only the standard SUDAAN wr SE, based on the primary sampling unit, would be used for the sake of simpler interpretation and for easier computation of the SE of functions of estimates, such as differences and ratios. A description of the previous method of variance estimation can

be found in the 1999 National Household Survey on Drug Abuse (NHSDA) sampling error report (Wheeless, Gordek, & Singh, 2001).

Also note that, prior to 2004, the SEs were applied directly from SUDAAN to only a subset of tables. Since then, the process has changed so that the decision about which method of calculation would be used for the SEs of estimated totals was made at the estimate level (e.g., the cell level) rather than at the marginal table level. In this way, the estimated totals would have consistent values for their variances throughout all of the reported tables. A specific set of domains used as covariates in the poststratification step of the NSDUH weighting process were designated as the "controlled" domains. The SEs reported for estimates from these domains would be based on the original method. Estimates from all other domains would include the SEs directly from the SUDAAN calculation. A more detailed discussion of the 2004 change in SE reporting can be found in Section 3.2.1 of Chapter 3 in the 2019 NSDUH's methodological summary and definitions report (Center for Behavioral Health Statistics and Quality [CBHSQ], 2020c).

DEFFs associated with prevalence estimates below 0.00005 or greater than or equal to 0.99995 (an ad hoc rule representing 0 or 1 in practice) or prevalence estimates exhibiting low precision were not used for determining the medians. To identify estimates with low precision, the suppression rule used in earlier years was applied. Specifically, DEFFs or the corresponding prevalence estimates were not included if the corresponding RSE of $-\ln(\hat{p})$ satisfies

$$RSE[-\ln(\hat{p})] > 0.175 \text{ when } \hat{p} \leq 0.5$$

or

$$RSE[-\ln(1 - \hat{p})] > 0.175 \text{ when } \hat{p} > 0.5.$$

Another way to identify estimates with potentially low precision is to find estimates where the nominal sample size is under 100 or the effective sample size is under 68. The effective sample size is defined as follows:

$$\text{Effective } n = \frac{n}{deff} = \frac{\hat{p}(1 - \hat{p})}{[SE(\hat{p})]^2}.$$

This equation is part of the standard suppression rule that is used in the reporting of NSDUH estimates. See the 2019 NSDUH statistical inference report for more information (CBHSQ, in press).

It may be noted that, for a given sample size, the RSE of $-\ln(1 - \hat{p})$ increases as \hat{p} decreases, and for a given \hat{p} , it increases as the sample size decreases. Although the RSE of $-\ln(1 - \hat{p})$ is not symmetric about $\hat{p} = 0.5$, it makes logical sense for precision requirements to be identical for \hat{p} and $1 - \hat{p}$. Therefore, it is convenient to use the convention that the suppression rule for $\hat{p} < 0.5$ also applies for $\hat{p} > 0.5$ by replacing \hat{p} with $1 - \hat{p}$.

6.2 Derivation of the Approximation for the RSE of a Prevalence Estimate's Negative Natural Logarithm

Define the first-order Taylor series of a twice differentiable function, $f(\hat{\theta})$, about a point, θ , where $\hat{\theta}$ is a nearly (asymptotically unbiased) estimator for θ , as

$$f(\hat{\theta}) \cong f(\theta) + \left[df(\hat{\theta}) / d\hat{\theta} \right]_{\hat{\theta}=\theta} (\hat{\theta} - \theta),$$

then $\left[f(\hat{\theta}) - f(\theta) \right] \cong \left[df(\hat{\theta}) / d\hat{\theta} \right]_{\hat{\theta}=\theta} (\hat{\theta} - \theta)$. If $E(\hat{\theta}) = \theta$, then

$$MSE \left[f(\hat{\theta}) \right] = E \left[f(\hat{\theta}) - f(\theta) \right]^2 \cong \left[df(\hat{\theta}) / d\hat{\theta} \right]_{\hat{\theta}=\theta}^2 Var(\hat{\theta}),$$

where $Var(\hat{\theta}) = E(\hat{\theta} - \theta)^2$.

Let

$$\hat{\theta} = \hat{p} \quad f(\hat{\theta}) = -\ln(\hat{p}) \quad d \ln(\hat{p}) / d\hat{p} = -1/\hat{p},$$

then the approximation of the mean squared error (MSE) would be

$$MSE \left[-\ln(\hat{p}) \right] \cong MSE(\hat{p}) \div (-\hat{p})^2 = \left[RSE(\hat{p}) \right]^2,$$

and the approximation of the relative MSE could be shown as

$$RelMSE \left[-\ln(\hat{p}) \right] \cong \left[RSE(\hat{p}) \right]^2 \div \left[-\ln(\hat{p}) \right]^2.$$

Taking the square root of both sides of the equation leads to the approximation of $RSE[-\ln(\hat{p})]$ as

$$RSE[-\ln(\hat{p})] \cong RSE(\hat{p}) \div [-\ln(\hat{p})].$$

The derivation of $RSE[-\ln(1 - \hat{p})]$ follows a similar set of steps.

6.3 Comparison of Observed Precision with Expected Precision

In this section, benchmarks from the 2019 NSDUH design process are compared with the estimated achieved precision of important outcome measures. These benchmarks are the predicted precision that the statisticians anticipated during the design of the survey. Using 2017 NSDUH data, predicted precision requirements for the 2019 designs were specified in terms of targeted RSEs and minimum sample sizes. To obtain the targeted RSEs, RSEs were computed for 25 measures of interest for specific domains of interest. These 25 key NSDUH outcomes that the sample design optimization for the 2019 NSDUH was based on included

recency-of-use estimates for both illicit and licit drugs, dependence on alcohol and illicit drug use, treatment for substance abuse, and mental health issues. Specifically, the following outcomes were used for 2019 (variable names on the NSDUH data files are in parentheses):

- alcohol use in the past month (ALCMON),
- binge alcohol use in the past month (BNGDRKMON),
- marijuana use in the past month (MRJMON),
- cigarette use in the past month (CIGMON),
- misuse of a pain reliever in the past month (PNRNMMON),
- alcohol use disorder in the past year (ABODALC),
- illicit drug use disorder in the past year (UDPYILL),
- alcohol use disorder or illicit drug use disorder in the past year (UDPYILAL),
- specialty substance use treatment in the past year (TXYRSPILAL),
- serious mental illness (SMI) in the past year (SMIYR_U), and
- major depressive episode (MDE) in the past year (AMDEYR).

[Table 6.1](#) shows a comparison of the projected and observed RSEs for the 25 outcomes from the 2019 sample design report's specified domain breakdowns (CBHSQ, 2020a).

6.3.1 Sample and Precision Requirements

Initial sample requirements for the 2019 NSDUH were defined in terms of the following:

- target sample sizes of 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; and
- allocation to age groups as follows: 25 percent for youths aged 12 to 17, 25 percent for young adults aged 18 to 25, 15 percent for adults aged 26 to 34, 20 percent for adults aged 35 to 49, and 15 percent for adults aged 50 or older.

The 1999 sample was the first to reflect the objective of the Substance Abuse and Mental Health Services Administration to develop more reliable national estimates and representative state-level estimates using small area estimation (SAE) and direct estimation procedures. To achieve this objective in 2019, the targeted sample size by state was set to be at least 960 completed interviews. In 13 states, the target was set at greater than 960 completed interviews. The larger overall sample made it possible to get adequate precision for Hispanic and non-Hispanic black or African-American populations without any targeted oversampling of high concentration areas of these populations or any oversampling through screening for these populations.

Table 6.1 Comparisons of Projected and Observed Relative Standard Errors and Sample Sizes for Key Outcome Measures, by Demographic Domain, 2019 NSDUH

| Data File Variable Name | Measure | Domain | Prevalence | Projected RSE | Observed RSE | Relative Change in RSE ¹ | Expected Sample Size | Observed Sample Size | Relative Change in Sample Size ² |
|-------------------------|---|-----------------|------------|---------------|--------------|-------------------------------------|----------------------|----------------------|---|
| ALCMON | Past Month Alcohol Use | 12+ | 0.5077 | 0.0071 | 0.0068 | -0.0469 | 67,507 | 67,625 | 0.0017 |
| ALCMON | Past Month Alcohol Use | 12-20 | 0.1851 | 0.0163 | 0.0243 | 0.4926 | 23,125 | 23,158 | 0.0014 |
| ALCMON | Past Month Alcohol Use | 50+ | 0.4919 | 0.0124 | 0.0130 | 0.0523 | 10,126 | 10,509 | 0.0378 |
| ALCMON | Past Month Alcohol Use | API, 12+ | 0.3781 | 0.0419 | 0.0374 | -0.1064 | 3,909 | 3,432 | -0.1220 |
| ALCMON | Past Month Alcohol Use | AIAN, 12+ | 0.3210 | 0.0755 | 0.0980 | 0.2974 | 750 | 804 | 0.0720 |
| ALCMON | Past Month Alcohol Use | Pregnant, 12-44 | 0.0956 | 0.2017 | 0.1458 | -0.2769 | 758 | 759 | 0.0013 |
| BNGDRKMON | Past Month Binge Alcohol Use | 18-25 | 0.3427 | 0.0136 | 0.0158 | 0.1608 | 16,877 | 16,516 | -0.0214 |
| BNGDRKMON | Past Month Binge Alcohol Use | 12+ | 0.2392 | 0.0116 | 0.0109 | -0.0640 | 67,507 | 67,625 | 0.0017 |
| MRJMON | Past Month Marijuana Use | 12+ | 0.1148 | 0.0180 | 0.0166 | -0.0802 | 67,507 | 67,625 | 0.0017 |
| MRJMON | Past Month Marijuana Use | 12-17 | 0.0738 | 0.0389 | 0.0354 | -0.0906 | 16,877 | 16,858 | -0.0011 |
| MRJMON | Past Month Marijuana Use | 18-25 | 0.2297 | 0.0193 | 0.0191 | -0.0093 | 16,877 | 16,516 | -0.0214 |
| MRJMON | Past Month Marijuana Use | 50+ | 0.0647 | 0.0546 | 0.0460 | -0.1582 | 10,126 | 10,509 | 0.0378 |
| MRJMON | Past Month Marijuana Use | API, 12+ | 0.0482 | 0.0996 | 0.0896 | -0.1001 | 3,909 | 3,432 | -0.1220 |
| MRJMON | Past Month Marijuana Use | AIAN, 12+ | 0.1202 | 0.1008 | 0.1517 | 0.5050 | 750 | 804 | 0.0720 |
| MRJMON | Past Month Marijuana Use | Pregnant, 12-44 | 0.0542 | 0.2359 | 0.1867 | -0.2084 | 758 | 759 | 0.0013 |
| CIGMON | Past Month Cigarette Use | 12-17 | 0.0230 | 0.0584 | 0.0653 | 0.1189 | 16,877 | 16,858 | -0.0011 |
| CIGMON | Past Month Cigarette Use | 12+ | 0.1667 | 0.0153 | 0.0139 | -0.0909 | 67,507 | 67,625 | 0.0017 |
| PNRNMMON | Past Month Pain Reliever Misuse | 18-25 | 0.0124 | 0.0718 | 0.0839 | 0.1680 | 16,877 | 16,516 | -0.0214 |
| PNRNMMON | Past Month Pain Reliever Misuse | 12+ | 0.0102 | 0.0513 | 0.0579 | 0.1290 | 67,507 | 67,625 | 0.0017 |
| ABODALC | Past Year Alcohol Use Disorder | 12+ | 0.0527 | 0.0244 | 0.0232 | -0.0483 | 67,507 | 67,625 | 0.0017 |
| UDPYILL | Past Year Illicit Drug Use Disorder | 12+ | 0.0302 | 0.0311 | 0.0301 | -0.0315 | 67,507 | 67,625 | 0.0017 |
| UDPYILAL | Past Year Substance Use Disorder | 50+ | 0.0411 | 0.0595 | 0.0525 | -0.1179 | 10,126 | 10,509 | 0.0378 |
| TXYRSPILAL | Past Year Specialty Substance Use Treatment | 12+ | 0.0096 | 0.0596 | 0.0563 | -0.0559 | 67,507 | 67,625 | 0.0017 |
| SMIYR_U | Past Year SMI | 18+ | 0.0524 | 0.0288 | 0.0250 | -0.1326 | 50,630 | 50,767 | 0.0027 |
| AMDEYR | Past Year MDE | 18+ | 0.0783 | 0.0229 | 0.0199 | -0.1317 | 50,630 | 50,134 | -0.0098 |

AIAN = American Indian or Alaska Native (NEWRACE2 = 3); API = Asian or Other Pacific Islander (NEWRACE2 = 4 or 5); MDE = major depressive episode; Pregnant 12-44 = (PREG2=1); RSE = relative standard error; SMI = serious mental illness.

Note: Projected RSEs were determined using 2014 through 2022 state and age sample allocations in a variance component model. All model components were updated using 2017 NSDUH data.

¹ Relative Change in RSE = $\{[RSE(observed) - RSE(projected)] / RSE(projected)\}$.

² Relative Change in Sample Size = $\{[2018 Sample Size - Expected Sample Size] / (Expected Sample Size)\}$.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2017 and 2019.

Unlike previous NSDUHs, no specific precision requirements were set for the 2014 through 2022 NSDUHs. Instead, they were designed to achieve acceptable precision for various subpopulations of interest, which accounted for the allocation of people per state and the requirement to support direct estimation in some large sample states and SAE in the remaining states. Using the state and age group distribution, estimates and RSEs were modeled for 25 key outcome measures and domains of interest.

6.3.2 Observed versus Expected Precision

In [Table 6.1](#), the expected and observed RSEs and sample sizes for the 2019 NSDUH are presented for the 25 key outcome measures. All but four of the observed RSEs relatively differed by no more than roughly 21 percent. Although this might be considered quite a large difference in terms of percentages, the point changes from the expected RSEs were generally quite small. Also, the percentage changes followed no particular direction. Out of 25 observed precisions, 17 were less than the expected precisions described in the 2019 sample design plan, and out of the 4 observed RSEs greater than roughly 21 percent, 1 was less than the expected precision (CBHSQ, 2018).

6.4 Comparison of Median and Mean Design Effects

The mean DEFF is more sensitive to outliers and is generally larger than the median DEFF. [Table 6.2](#) compares the median and mean of 53 DEFFs for three age groups and over all ages in the 2019 NSDUH design. Comparisons are also provided for the four race/ethnicity categories, although they were not used as stratification variables when selecting individuals within households. [Table 6.3](#) provides the same median and mean DEFFs for the 2017 through 2019 NSDUHs. In a continuation of what was seen for the 2014 NSDUH, the current design used in 2019 is more efficient overall and within all race/ethnicity groups when compared with the design used prior to 2014 (CBHSQ, 2015a).

Table 6.2 Comparison of Median and Mean Design Effects of 53 Outcomes: 2019

| Outcome | Median Design Effect | Mean Design Effect | Difference (Mean – Median) | Percentage Difference ¹ |
|---|----------------------|--------------------|----------------------------|------------------------------------|
| Total | 2.35 | 2.40 | 0.06 | 2.42 |
| Age (Years) | | | | |
| 12 to 17 | 1.71 | 1.83 | 0.11 | 6.65 |
| 18 to 25 | 1.84 | 1.81 | -0.03 | -1.62 |
| 26+ | 1.78 | 1.82 | 0.04 | 2.21 |
| Race/Ethnicity | | | | |
| White, Not Hispanic or Latino | 2.21 | 2.20 | -0.01 | -0.47 |
| Black or African American, Not Hispanic or Latino | 2.13 | 2.30 | 0.17 | 7.95 |
| Hispanic or Latino | 2.34 | 2.38 | 0.04 | 1.65 |
| Other or Multiple, Not Hispanic or Latino | 2.59 | 2.57 | -0.01 | -0.41 |

¹ Computed as $100 * (Mean - Median) / Median$.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2019.

Table 6.3 Median and Mean Design Effects of 53 Outcomes: 2017 through 2019

| Outcome | Median Design Effect | | | Mean Design Effect | | |
|---|----------------------|------|------|--------------------|------|------|
| | 2017 | 2018 | 2019 | 2017 | 2018 | 2019 |
| Total | 2.25 | 2.42 | 2.35 | 2.31 | 2.47 | 2.40 |
| Age (Years) | | | | | | |
| 12 to 17 | 1.59 | 1.72 | 1.71 | 1.56 | 1.70 | 1.83 |
| 18 to 25 | 1.80 | 1.74 | 1.84 | 1.83 | 1.76 | 1.81 |
| 26+ | 1.75 | 1.78 | 1.78 | 1.78 | 1.88 | 1.82 |
| Race/Ethnicity | | | | | | |
| White, Not Hispanic or Latino | 2.22 | 2.23 | 2.21 | 2.13 | 2.20 | 2.20 |
| Black or African American, Not Hispanic or Latino | 2.11 | 2.24 | 2.13 | 2.26 | 2.29 | 2.30 |
| Hispanic or Latino | 2.39 | 2.53 | 2.34 | 2.40 | 2.56 | 2.38 |
| Other or Multiple, Not Hispanic or Latino | 2.29 | 2.51 | 2.59 | 2.43 | 2.74 | 2.57 |

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2017 to 2019.

The median and mean DEFF estimates were based on estimates from the following four types of substance use and mental health categories: (a) *eight illicit drug use categories*: any illicit drug use, marijuana, cocaine, crack, inhalants, hallucinogens, misuse of any psychotherapeutics, and misuse of pain relievers; (b) *seven licit drug use categories*: tobacco, cigarettes, smokeless tobacco, cigars, alcohol, heavy drinking, and binge drinking; (c) *six treatment or abuse categories*: abuse of drugs or alcohol, dependence on drugs or alcohol, treatment received for illicit drug use, treatment received for alcohol use, treatment received for either alcohol use or illicit drug use, and treatment received for both alcohol use and illicit drug use; and (d) *six mental health categories*: any mental illness, SMI, mental health services, MDE, suicidal thoughts, and serious psychological distress. Estimates used from the illicit and licit categories included one from each of three recency-of-use classes: ever used, used in past year, and used in past month. An exception was made for estimates of heavy drinking and binge drinking, which are past month variables. The treatment or abuse and mental health variables are for the past year.

The median and the mean DEFFs were calculated from the above estimates for the total population, by age and by race/ethnicity. As seen from [Table 6.2](#), the mean DEFF turned out to be larger than the median DEFF in five of the eight domains. The differences between the mean and median DEFFs were 2.42 percent or less for six of the eight comparison groups with the largest differences being 6.65 percent for the 12 to 17 age group and 7.95 percent for the non-Hispanic or Latino black or African American race/ethnicity group.

This page intentionally left blank

7. Issue Encountered

In the 2019 National Survey on Drug Use and Health (NSDUH), one issue with the sample implementation was encountered. As mentioned in Section 2.2.3, to ensure adequate samples for supplemental studies, the maximum dwelling unit (DU) sample size was set at 100 per segment or half of the actual listing unit count, whichever was smaller. When the maximum allowable sample size was smaller than the ideal initial sample allocation, then the segment's DU sample size was said to be truncated, and the excess allocation was reallocated to other segments in the same state. This process was iterative until the projected sample approximately matched the state's target sample. With declining response rates, more sample was allocated initially, and a higher percentage of segment samples was truncated than in previous years. In some states, the projected sample size did not merge to the target sample size after several iterations of the sample allocation programs. [Table 7.1](#) lists the segments with sample sizes truncated to the maximum DU limits from the 2019 quarterly sample allocations.

Table 7.1 2019 NSDUH Segments with Truncated Sample Sizes


| Quarter | Number of Truncated Segments | Truncated Segment List |
|---------|------------------------------|--|
| 1 | 57 | AR02G1, AZ10G1, CA21G1, DC02G1, DC04F1, DC04G1, DC06G1, DC07F1, DC07G1, DC09F1, DC11F1, DC11G1, DC12G1, DE01G1, DE12F1, FL04G1, FL08F1, FL11G1, FL13F1, FL15G1, FL16G1, FL25F1, FL28F1, GA13G1, HI03G1, IL09G1, ME12G1, MI23F1, MI24F1, MO09F1, MT04G1, MT07G1, NC08F1, ND06G1, ND09F1, ND11G1, ND12G1, NH01G1, NH03F1, NV06G1, NV07G1, NY05F1, NY06G1, NY11F1, NY16F1, NY17G1, NY24F1, OH23G1, OK02G1, OK11G1, PA21G1, RI11G1, VA04G1, VA06F1, VT09G1, VT10F1, WY08F1 |
| 2 | 93 | AZ01F2, AZ03F2, CA02F2, CA04F2, CA07G2, CA09F2, CA09G2, CA16G2, CA20G2, CA22G2, CA25G2, CO04G2, DC01G2, DC02F2, DC02G2, DC03F2, DC03G2, DC04F2, DC05F2, DC06F2, DC06G2, DC07F2, DC07G2, DC08F2, DC08G2, DC09F2, DC09G2, DC10F2, DC11F2, DC11G2, DC12F2, DC12G2, DE07G2, DE12F2, DE12G2, FL02F2, FL16G2, FL20F2, FL23G2, FL24G2, FL29F2, HI11G2, IA10G2, IA11F2, IL09F2, IL17G2, LA02G2, LA06G2, ME01F2, ME08F2, ME08G2, MI12F2, MI15G2, MI23F2, MT01G1, MT02F2, MT02G2, MT04G2, MT06G2, MT09F1, NC01F2, NC07F2, ND01F2, ND07F2, ND09F2, ND12G2, NH01G2, NJ09G2, NJ14F2, NJ14G2, NJ15G2, NV05G2, NV07F2, NY05F2, NY06G2, NY08G2, NY09G2, NY17F2, NY17G2, NY18F2, NY25F2, OR01F2, OR12F2, PA17F2, PA22F2, RI12F2, VA06G2, VA09F2, VT11G2, WA11F1, WV07F2, WV09F2 |
| 3 | 100 | AZ04G3, AZ11F3, AZ11G3, CA03G3, CA09F3, CA10F3, CA11F3, CA17G3, CA21G3, CA30G3, CA33F3, CA33G3, CA34G3, CO01F3, CO01G3, CO11G3, DC02F3, DC02G3, DC03G3, DC04F3, DC04G3, DC05F3, DC05G3, DC06G3, DC07F3, DC07G3, DC08F3, DC08G3, DC09F3, DC09G3, DC10F3, DC10G3, DC11F3, DC11G3, DC12F3, DC12G3, DE01F3, DE01G3, DE03G3, DE04F3, DE04G3, DE05G3, DE10G3, DE11F3, DE12F3, DE12G3, FL02G3, FL06F3, FL07F3, FL07G3, FL11F3, FL19F3, FL22F3, FL22G3, FL27F3, FL27G3, FL29G3, GA12F3, HI07G3, HI11G3, HI12F3, ID05F4, IL07F3, IL15G3, IN10F3, LA01G3, MD05G3, ME03F3, ME12G3, MT06F3, MT09F4, MT12F4, NC08F3, ND07F4, ND09G3, ND11F4, ND11G3, NE04G3, NH01F3, NH12F3, NM03F4, NV10F3, NY17F3, NY17G3, NY30G3, OH03G3, PA08G3, RI07F3, RI10G3, SC11G3, SD05G3, SD10G3, TX19G3, TX26G3, VA03F3, VA13F3, VT02F3, VT12F3, VT12G3, WY12F4 |
| 4 | 65 | AK11F4, AK12G4, AZ09F4, AZ12F4, CA01F4, CA11F4, CA15F4, CA21F4, CO04G4, DC01F4, DC01G4, DC02G4, DC03F4, DC03G4, DC04F4, DC04G4, DC05F4, DC05G4, DC06F4, DC06G4, DC07F4, DC07G4, DC08F4, DC08G4, DC09F4, DC09G4, DC10F4, DC10G4, DC11G4, DC12G4, DE01G4, DE03F4, DE08G4, DE09G4, DE12G4, FL06F4, FL13G4, FL15G4, FL23F4, FL28G4, FL30G4, GA15F4, IL09G4, IL23F4, KS04F4, MA11G4, ME11G4, MI24F4, MN09F4, MT01G4, NJ09F4, NY02G4, NY03F4, NY07F4, NY14F4, NY14G4, NY17F4, NY18F4, OH19F4, OK11G4, SC03G4, SD01G4, VT12G4, WA06G4, WI02G4 |


In recent years, the average number of DUs per segment sample increased dramatically in the District of Columbia. A review of the data from quarter 3 of 2019 found that most of the segment samples in the District of Columbia were truncated; therefore, the district was not projected to yield its target sample size. California, Delaware, and Florida also had relatively large numbers of truncated segment samples in quarter 3 of 2019. Nationally, the majority of segments with truncated samples were in urban areas.

As a result of these findings, the Substance Abuse and Mental Health Services Administration approved the decision to increase the maximum DU sample size to 120 beginning in quarter 1 of the 2020 NSDUH. With the majority of affected segments clustered in a handful of states, the change is expected to have very little overall impact on data collection. It will not affect estimation because the average number of completed interviews per segment does not change.

References

- American Association for Public Opinion Research. (2016). *Standard definitions: Final dispositions of case codes and outcome rates for surveys* (9th ed.). Retrieved from https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf ↗
- Brewer, K. R. W. (1963). A model of systematic sampling with unequal probabilities. *Australian Journal of Statistics*, 5(1), 5-13. <https://doi.org/10.1111/j.1467-842x.1963.tb00132.x> ↗
- Brewer, K. R. W. (1975). A simple procedure for sampling π pswor. *Australian Journal of Statistics*, 17(3), 166-172. <https://doi.org/10.1111/j.1467-842x.1975.tb00954.x> ↗
- Center for Behavioral Health Statistics and Quality. (2015a). *2014 National Survey on Drug Use and Health: Methodological resource book (Section 14, Sample experience report)*. Retrieved from <https://www.samhsa.gov/data/>
- Center for Behavioral Health Statistics and Quality. (2015b). *National Survey on Drug Use and Health: 2014 sample redesign impact assessment, final 12-month report*. Retrieved from <https://www.samhsa.gov/data/>
- Center for Behavioral Health Statistics and Quality. (2018). *2019 National Survey on Drug Use and Health: Sample design plan*. Rockville, MD: Substance Abuse and Mental Health Services Administration.
- Center for Behavioral Health Statistics and Quality. (2019-2020). *State estimates of substance use from the 2017-2018 National Surveys on Drug Use and Health*. Retrieved from <https://www.samhsa.gov/data/>
- Center for Behavioral Health Statistics and Quality. (2020a). *2019 National Survey on Drug Use and Health: Methodological resource book (Section 2, Sample design report)*. Retrieved from <https://www.samhsa.gov/data/>
- Center for Behavioral Health Statistics and Quality. (2020b). *2019 National Survey on Drug Use and Health: Methodological Resource Book (Section 8, Data collection final report)*. Retrieved from <https://www.samhsa.gov/data/>
- Center for Behavioral Health Statistics and Quality. (2020c). *2019 National Survey on Drug Use and Health: Methodological summary and definitions*. Retrieved from <https://www.samhsa.gov/data/>
- Center for Behavioral Health Statistics and Quality. (in press). *2019 National Survey on Drug Use and Health: Methodological resource book (Section 13, Statistical inference report)*. Retrieved from <https://www.samhsa.gov/data/>

Chromy, J. R. (1979). Sequential sample selection methods. In *Proceedings of the 1979 American Statistical Association, American Statistical Association, Survey Research Methods Section, Washington, DC* (pp. 401-406). Retrieved from <http://www.asasrms.org/Proceedings/index.html> 

Chromy, J. R., & Penne, M. (2002). Pair sampling in household surveys. In *Proceedings of the 2002 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, New York, NY* (pp. 552-554). Retrieved from <http://www.asasrms.org/Proceedings/index.html> 

Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). New York, NY: John Wiley & Sons.

Cunningham, D., Peytchev, A., Stanforth, P., Cohen, L., Johnson, I., McKamey, A., Meyer, M., Morton, K., Snodgrass, J., & McHenry, G. (2009). *Simplifying the field interviewer job and increasing interviewer retention: Ideas and recommendations* (prepared for the Substance Abuse and Mental Health Services Administration under Contract No. 283-2004-00022). Research Triangle Park, NC: RTI International.

Iannacchione, V., McMichael, J., Shook-Sa, B., & Morton, K. (2012). *A proposed hybrid sampling frame for the National Survey on Drug Use and Health* (prepared for the Substance Abuse and Mental Health Services Administration under Contract No. 283-2004-00022). Research Triangle Park, NC: RTI International.

Kish, L. (1965). *Survey sampling* (pp. 313-315). New York, NY: John Wiley & Sons.

Office of Management and Budget. (2009, December 1). *OMB Bulletin No. 10-02: Update of statistical area definitions and guidance on their uses*. Retrieved from <https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/bulletins/2010/b10-02.pdf>

RTI International. (2012). *National Survey on Drug Use and Health: Sample redesign issues and methodological studies* (RTI/0209009.486.001 and 0211838.108.006.005, prepared for the Substance Abuse and Mental Health Services Administration under Contract Nos. 283-2004-00022 and HHSS283200800004C). Research Triangle Park, NC: Author.

RTI International. (2013). *SUDAAN[®], Release 11.0.1 [computer software]*. Research Triangle Park, NC: Author.

U.S. Census Bureau, Redistricting Data Office. (2009, April). Appendix A: Geographic terms and concepts. In *2008 Redistricting data prototype (Public Law 94-171) summary file: 2010 Census of Population and Housing: Technical documentation (TD/PL2008(P))*. Retrieved from https://www.census.gov/support/2008PLPrototype_files/pl2008.pdf

Wheless, S. C., Gordek, H., & Singh, A. C. (2001). *The 1999 National Household Survey on Drug Abuse: Sampling error report* (prepared for the Substance Abuse and Mental Health Services Administration under Contract No. 283-98-9008). Research Triangle Park, NC: RTI International.

List of Contributors

This National Survey on Drug Use and Health methodological report was prepared by the Substance Abuse and Mental Health Services Administration (SAMHSA), Center for Behavioral Health Statistics and Quality, and RTI International (a registered trademark and a trade name of Research Triangle Institute). Work by RTI was performed under Contract No. HHSS283201700002C, Deliverable No. 18. Kathryn Piscopo served as the government project officer and as the contracting officer representative, and David Hunter served as the RTI project director.

Significant contributors to this report at SAMHSA include Yang Cheng and Marlon G. Daniel.

Significant contributors to this report at RTI include Peilan C. Martin, Dexter McNutt, and Katherine B. Morton. Other contributors at RTI include Harper Gordek, Donna K. Hewitt, David Hunter (project director), Phillip S. Kott, and Rebecca K. Thomson.

Also at RTI, Margaret A. Johnson and Richard S. Straw copyedited the report, and Teresa F. Bass coordinated its web production.

This page intentionally left blank

This page intentionally left blank



**SAMHSA's mission is to reduce the impact of substance abuse and mental illness on America's communities.
1-877-SAMHSA-7 (1-877-726-4727) | 1-800-487-4889 (TDD) | www.samhsa.gov**