NATIONAL SURVEY ON DRUG USE AND HEALTH

SAMPLE REDESIGN ISSUES AND METHODOLOGICAL STUDIES

Contract Nos. 283-2004-00022 and HHSS283200800004C
RTI Project Nos. 0209009.486.001 and 0211838.108.006.005

Prepared for:

Substance Abuse and Mental Health Services Administration
Rockville, Maryland 20857

Prepared by:

RTI International
Research Triangle Park, North Carolina 27709

April 2012

NATIONAL SURVEY ON DRUG USE AND HEALTH

SAMPLE REDESIGN ISSUES AND METHODOLOGICAL STUDIES

Contract Nos. 283-2004-00022 and HHSS283200800004C
RTI Project Nos. 0209009.486.001 and 0211838.108.006.005

Prepared for:

Substance Abuse and Mental Health Services Administration
Rockville, Maryland 20857

Prepared by:

RTI International
Research Triangle Park, North Carolina 27709

April 2012

Acknowledgments

This report was prepared for the Substance Abuse and Mental Health Services Administration (SAMHSA), Center for Behavioral Health Statistics and Quality (CBHSQ), formerly the Office of Applied Studies (OAS), under Contract Nos. 283-2004-00022 and HHSS283200800004C, by staff at RTI International (a trade name of Research Triangle Institute).

Authors of this report listed alphabetically include James R. Chromy, David B. Cunningham, Douglas Currivan, Elizabeth Dean, Ralph E. Folsom, Rebecca A. Granger, David C. Heller, Donna Hewitt, Erica L. Hirsch, Ilona Johnson, Phillip S. Kott, Larry A. Kroutil, Patty LeBaron, Allison C. McKamey, Martin Meyer, Katherine B. Morton, Scott Payne, Thomas G. Virag (Project Director), and Kevin Wang.

In addition to the listed authors, significant contributors include Grant Bettinger, Joyce Clay-Brooks, Lauren Cohen, Valerie Garner, Richard Hair, Jeffrey Lyons, Mary Ellen Marsden, Peilan Chen Martin, Gretchen McHenry, Carol Offen, Lisa Packer, Hyunjoo Park, Michelle Pattie, Heather Ringeisen, Roxanne Snaauw, Richard S. Straw, Pamela Tuck, and Tammie D. Woerner.

Table of Contents

Chapter **Page**

Table of Contents (continued)

Chapter **Page**

Table of Contents (continued)

Table of Contents (continued)

Chapter **Page**

Appendix

List of Tables

List of Tables (continued)

Table **Page**

List of Tables (continued)

Table **Page**

List of Tables (continued)

Table **Page**

List of Tables (continued)

Table **Page**

List of Figures

Figure **Page**

1. Introduction and Summary

The National Survey on Drug Use and Health (NSDUH), sponsored by the Substance Abuse and Mental Health Services Administration (SAMHSA), is a national survey of the U.S. civilian, noninstitutionalized population aged 12 or older. The conduct of NSDUH is paramount in meeting a critical objective of SAMHSA's mission to provide information on the incidence and prevalence of substance abuse and mental illness in the United States and selected areas as required by Section 505 of the Public Health Service Act (42 U.S.C. 290aa-4). To continue producing timely and relevant data, SAMHSA's Center for Behavioral Health Statistics and Quality (CBHSQ) must update NSDUH periodically to reflect changing substance use and mental health behavior and data needs. CBHSQ is planning to implement a redesigned NSDUH sample beginning with the 2014 survey and a redesigned questionnaire beginning with the 2015 survey. The overall redesign will seek to achieve three main goals: (1) to bring NSDUH costs in line with anticipated budget levels, (2) to revise the questionnaire to address changing policy and data needs, and (3) to modify the survey methodology to improve the quality of the estimates and the efficiency of data collection and processing. In preparation for the redesign, SAMHSA tasked RTI International[1] with examining the feasibility and impacts of a variety of sample design changes on survey costs and data precision. With sample design (which includes data collection) driving much of the cost of the survey, these issues will be important in helping SAMHSA reach its goal of implementing a more efficient design that meets the data needs of its users while maintaining the integrity of the results.

The Sample Issues Methods Study examined the following issues:

- optimal cluster size,
- optimal sample distribution with respect to demographic and geographic groups,
- pros and cons of collecting data on a more continuous basis (with little or no downtime within data collection periods),
- feasibility of conducting the survey every other year,
- implications of expanding the target population to include children under the age of 12, and
- feasibility of a flexible design.

Each of these issues was examined in a separate report. This report consolidates the six previous Sample Issues Methods Study reports into a single document, with Chapters 2 through 7 each examining one of the above sample issues. Three appendices provide additional information for the oversampling optimization discussion. A short summary of the findings from Chapters 2 to 7 follows.

Cluster Optimization: Chapter 2. Since 1999, NSDUH's design has provided for State-level estimates in all 50 States and the District of Columbia by requiring minimum sample sizes in each State. The design targets 300 persons in each of three age groups (12 to 17, 18 to 25, and

---

[1] RTI International is a trade name of Research Triangle Institute.

26 or older) or 900 persons total within each small sample State. In the large sample States,[2] 1,200 persons per age group or 3,600 persons total are required. Based on an analysis of cost variance tradeoffs, an average cluster size of 9.375 was considered near optimal. Thus, each annual sample is selected from 48 area segments in each small sample State and 192 segments in each large sample State, or a total of 7,200 area segments.

Using updated segment and person-level cost estimates, a formal optimization indicated that a cluster size of around 29 persons would be optimal for the redesigned NSDUH. The larger cluster size affords the ability to produce nearly as precise estimates at a greatly reduced cost. However, the optimization does not consider SAMHSA's desire to select the sample from a sufficient number of clusters to be able to produce reliable State- and substate-level estimates using small area estimation (SAE) or direct methods. To accommodate this secondary objective, SAMHSA may consider requiring a minimum number of clusters per State.

Oversampling Optimization: Chapter 3. The oversampling optimization investigation was conducted in several stages. A preliminary report presented estimates and variances for youths aged 12 to 17, the 50 or older age group, several racial/ethnic groups, and pregnant women to identify candidates for oversampling. The report then provided several solutions to a formal age group optimization among seven age groups (12 to 14, 15 to 17, 18 to 25, 26 to 34, 35 to 49, 50 to 64, and 65 or older). For the first solution, all sample size requirements were relaxed, including the requirement that the national sample size be 67,500 interviews. This solution showed that the SAMHSA precision requirements could be met with a much smaller national sample. The second solution required that the overall sample size be maintained at 67,500, but relaxed the requirement that the sample be equally allocated to the 12 to 17, 18 to 25, and 26 or older age groups. This solution suggested that the optimal age group distribution is 22, 27, and 51 percent to the three major age groups, respectively. The final solution required that the total sample of 67,500 be distributed equally to the 12 to 17, 18 to 25, and 26 or older age groups and showed the optimal distribution to the seven age groups under this scenario.

Subsequently, SAMHSA refined the design specifications, and several alternative sample designs were evaluated. These designs included a national design of 40,000 persons, a hybrid design that preserves minimal sample sizes for some States, and a 50-State design with the current sample size. For each of these designs, three age group allocations were evaluated: the current age group allocation and two allocations in which the sample is distributed to the 12 to 17, 18 to 25, and 26 or older age groups in proportions of 25, 25, and 50 percent, respectively. In addition, the current cluster size of 9.375 and a larger cluster size of 28.125 were evaluated for each design option. The design alternatives were evaluated in six dimensions: sampling frame structure, cost, precision of estimates, field management issues, State and substate estimates, and response rates.

The national design with a sample of 40,000 offers the greatest cost reductions, in large part to the reduction in field staff required. Hybrid designs still produce major cost reductions while preserving minimal sample sizes for some State estimates. On the other hand, the precision analysis favors a 50-State design with a smaller cluster size. State and substate estimates are not

[2] The large sample States are California, Florida, Illinois, Michigan, New York, Ohio, Pennsylvania, and Texas.

practical under the national design, and State estimates will require pooling data over more years in the hybrid design. Finally, unweighted overall response rates are expected to decrease for the age allocations that shift the sample distribution toward older age groups.

Several additional 50-State design alternatives with a State allocation different from the current design were later evaluated and are included in Appendix B. For each of these options, referred to as Options A, B, and C, several cluster size options were evaluated in terms of the six dimensions above. In addition, the interviewer experience implications were evaluated for Options A and B.

In summary, in all of the 50-State design alternatives, precision improvements are expected for all demographic subgroups except youths aged 12 to 17. Similar to the earlier finding, the greatest cost savings will occur by lowering the sample size and increasing the cluster size. Finally, the new design options with cluster sizes of 9.375 will support State and substate SAE on the same schedule as the current design. The larger cluster sizes would be preferred for SAE if the 50 percent overlapping sample design were to be eliminated, such as with an address-based sampling (ABS) frame.

Lastly, because of concern over not being able to produce a sufficiently large sample in the Hawaiian island of Kauai for previous NSDUH substate reports, options were considered to adjust the allocation to achieve a minimum 3-year sample of 200 persons aged 12 or older in Hawaii's four more populous counties (see Appendix C).

Continuous Design Assessment: Chapter 4. The annual NSDUH sample of 67,500 persons is allocated approximately equally to four calendar quarters. Sample is released at the beginning of each quarter, and the majority of the interviews are completed within the first 4 to 6 weeks of the quarter. Although there is little evidence to suggest seasonal differences in substance use and mental health measures at the national level, there may be subgroup-level differences. Concerns over the front-heavy data collection periods and the possibility of temporal differences in prevalence estimates prompted the desire to evaluate more continuous design options.

Options of releasing data every 2, 4, and 6 weeks were considered. Although these options would possibly more accurately portray annual drug use and mental health measures, the reduced data collection periods could have negative impacts on response rates, costs, staffing, and data quality. In addition, there would be little downtime for training new and veteran field interviewers (FIs). On the other hand, more frequent sample releases could be beneficial for implementing questionnaire changes, responding to major events, and achieving sample targets. Releasing data more frequently and allowing overlapping data collection periods would present additional challenges.

A less drastic continuous design option involves collecting data every 8 weeks. This option is similar to the current design, but provides better coverage of the calendar year. The compressed time period would pose some problems for overcoming field challenges, such as adverse weather conditions and controlled access situations. However, it would afford more time for training and field management activities.

Finally, the continuous designs of other large surveys were considered for NSDUH. The Monitoring the Future study uses a single 6-month data collection period. Although a single, longer data collection period allows more time for overcoming field challenges, the data would not be representative of the full year. The American Community Survey releases sample on a monthly basis and allows 3 months for completion. This complex model would pose challenges for sampling, field management, weighting, and analysis.

Biennial Design Assessment: Chapter 5. Collecting and analyzing NSDUH data every other year rather than annually would result in large cost savings. However, there may be data quality and timeliness implications under a biennial design. Also, there would be some loss of efficiency compared with an annual design because of startup and shutdown costs under a biennial design.

Under the current NSDUH design, national estimates are produced annually and trends can be established relatively quickly when either a new baseline is established or new questions or estimates are introduced. By requiring minimum sample sizes in each State, the design also provides for State-level estimates in all 50 States and the District of Columbia. Using SAE procedures, annual State estimates are produced based on 2-year moving averages, and substate estimates are produced based on 3 years of data. Direct State estimates also are produced based on 3 or more years of pooled data. Under a biennial design, analyses requiring pooling across 2 or more years of data will no longer be timely and in some cases will become impractical. In addition, NSDUH data will no longer be useful for quickly responding to substantive issues or conducting special analyses on the impact of national events.

The current annual design offers continuous employment of field and professional staff. Under a biennial design, field and professional staff would not be needed in the "off" years; therefore, the survey would be significantly less expensive to conduct. On the other hand, the inability to retain field staff could contribute to an inconsistency in interviewer experience levels, which could affect response rates and respondent reporting of sensitive items.

To mitigate some of the concerns with a biennial design, two alternatives were considered: reducing the annual sample by one half (the half-sample design) and conducting the survey in half of the States one year and the other half the next year (the half-State design). As with the biennial design, both alternatives offer significant cost savings over the current design. However, there are still concerns with each alternative. The half-sample design eliminates data quality concerns by maintaining consistency of field and professional staff. However, the precision of annual estimates would be reduced, and analyses requiring pooling across multiple years of NSDUH data would require pooling over twice as many years to maintain equivalent precision. As a result, these estimates would be less timely and may not be useful.

With the half-State design, estimates would be published every other year and results would be difficult to interpret because not all of the data would be collected at the same time. Furthermore, the mix of FIs would change from year to year, which would likely have data quality implications. Finally, the half-State design would be limited in the ability to respond to emerging issues or conduct special analyses following major national events because not every State would be included in the sample each year.

Expanded Target Population Assessment: Including Children Under Age 12: Chapter 6. There are compelling reasons to expand the NSDUH target population to include children under the age of 12. Most notably, some children under the age of 12 do use drugs. Also, the undercoverage of past year initiates aged 11 or younger in the current design affects the mean age of first use estimate.

Although expanding the target population to include children under age 12 would have analytic benefit, there are several issues to consider. Issues related to sampling, nonresponse, the questionnaire, data quality, and costs were examined. First, expanding the target population to include children under age 12 is operationally feasible, but would come with an increase in costs. The expanded sample could be accommodated by increasing the number of household screenings or by selecting up to three persons per household. Based on experience with selecting pairs, the latter may have a negative impact on interview response rates. Response rates may be further affected by an increase in parental refusals. The parental refusal rate currently increases as the age of the youth respondent decreases. Other concerns noted include difficulty understanding some of the complex questions or wording in the NSDUH questionnaire; cognitive ability to understand respondent rights, such as privacy, confidentiality, and the voluntary nature of the study; and burden associated with the duration of the instrument administration.

Flexible Design: Chapter 7. A flexible design is desirable in the event that SAMHSA is required to reduce or supplement the NSDUH sample to address a particular need, change from a 50-State design to a national design, or make other significant changes. The flexible design chapter summarizes aspects of the current design that provide flexibility and describes features that could be incorporated to increase the flexibility of the redesigned NSDUH. Depending on the objective, flexible design options for the current and next NSDUH design include transitioning to a national or hybrid design (a design that preserves minimum State sample sizes at some level by supplementing a national sample), changing cluster sizes, using ABS, switching to more continuous or biennial data collection, and expanding the target population. Although they provide flexibility, the impact of these options on other areas of the survey design should be considered, such as loss of data comparability (i.e., trends), decreased precision, and changes in operational procedures that may have cost implications.

2. Cluster Optimization

2.1 Introduction

Since 1999, the design of the National Survey on Drug Use and Health (NSDUH) has provided for State-level estimates in all 50 States and the District of Columbia by requiring minimum sample sizes in each State. The design targeted 300 persons in each of three age groups (12 to 17, 18 to 25, and 26 or older) equally allocated to four quarters within each small sample State. Based on an analysis of the cost variance tradeoffs, an average cluster size of 3.125 persons in each of the three age groups (or an average of 9.375 persons over the three age groups combined) was considered near optimal.[3] When applied to the small States, a quarterly sample of 75 persons per quarter per age group could be obtained from 24 clusters or area segments. For unbiased variance estimation purposes, at least two observations are required per stratum (Chromy, 1981); maximum geographic stratification was obtained by defining 12 strata (State sampling [SS] regions) with two area segments each per quarter. For each of the other three quarters, 2 additional segments were selected, yielding 8 area segments per stratum, or 96 area segments per small sample State. This approach supported the target sample size for the small States of 300 persons per age group, or a total of 900 for the year. In the large sample States, 4 times as large a sample was required. Optimum cluster size configuration and maximum stratification, given the need for unbiased variance estimation, were maintained by simply quadrupling the number of strata (SS regions) to 48 per large sample State, yielding a sample of 300 persons per age group per quarter, 1,200 persons per age group over four quarters, and 3,600 persons per year over all three age groups. In summary, each annual sample is selected from 900 SS regions and 7,200 area segments nationally (Morton, Martin, Hirsch, & Chromy, 2008b).

Using recent experience, segment and person-level cost estimates could be refined. In addition, a formal cluster optimization was considered necessary for informing the NSDUH redesign. Thus, this chapter describes the results of an analysis that was conducted to determine the optimal number and size of clusters for NSDUH.

Sections 2.2 and 2.3 describe how the cost and variance inputs for the optimization model(s) were computed. Section 2.4 presents the results of some univariate calculations and the associated impact on cost and precision. Finally, Section 2.5 describes the results of a formal optimization process that considers several drug use measures and age groups simultaneously. A summary of the chapter's conclusions appears in Section 2.6.

---

[3] An informal optimization indicated that a slightly larger cluster size can be used. However, the cluster size of 3.125 persons per age group both supported the goals of the 50-State design and was similar to the size of primary sampling units (PSUs) used in previous National Household Surveys on Drug Abuse (NHSDAs).

2.2 Estimating Segment and Person-Level Costs

The cost figures[4] included in this NSDUH cluster optimization model are based on 2006 calendar year cost data. Using best knowledge and logical assumptions, task-level costs were divided into (a) variable costs per geographic cluster, (b) variable costs per interview, and (c) fixed costs. Only variable costs were included in the analysis. This included a portion of the costs from the instrumentation development and training and field materials (Task 2), sample design (Task 3), sample selection (Task 4), and data management and processing (Task 7), as well as the majority of costs from the field preparations (Task 5) and screening and interviewing (Task 6).[5]

2.3 Overview of Variance Modeling

Parametric variance models allow sampling statisticians to represent the variance of key estimates as a function of sample design parameters. Data from the 2007 NSDUH were used to estimate some of the key population parameters needed for these models. The required parameters include the following:

- variance components,
- unequal weighting effects (UWEs),
- averages and coefficients of variation for respondents per dwelling unit,
- averages and coefficients of variation for respondents per segment, and
- prevalence estimates, by age group.

2.3.1 Variance Components

The notation for the variance components associated with stratum, segment, dwelling unit, and person is shown in Table 2.1. The variance components were estimated using the method of moments in SAS PROC NESTED with equal weights.[6] They were computed for nine substance use and treatment variables of interest, controlling for age group. Variance components for the nine selected measures are shown in Table 2.2.

---

[4] The characteristics of census tracts played no direct role in determining these cost estimates. However, because the first stage of selection is census tracts and sample segments are contained within a single census tract to the extent possible, sample segments associated with densely populated census tracts will in theory have lower per screening and interview unit costs than more sparsely populated census tracts because of less travel time between travel points.

[5] It was assumed that the fixed costs and the per unit variable field costs will be essentially the same as under the current design with some allowance for salary increases and inflation. No adjustment was included for changes in the questionnaire structure, weighting, imputation, and other methods.

[6] Ignoring the weights in PROC NESTED (i.e., setting the weights equal to 1 assumes that the clustering in the sample mimics the clustering in the population. Based on prior experience, this assumption is not unreasonable for variance-modeling purposes. There is, however, an increase in variance caused by employing weights in the estimation of proportions. This increase is captured by the expression UWE_d in the variance models in Section 2.3.5.

Table 2.1 Variance Component Notation

Level	Symbol
State Sampling Region	σ^2_{SSR}
Segment	$\sigma^2_{segment}$
Dwelling Unit	σ^2_{DU}
Person	σ^2_{person}
Total	$\sigma^2 = \sigma^2_{SSR} + \sigma^2_{segment} + \sigma^2_{DU} + \sigma^2_{person}$

Table 2.2 Variance Components of Residuals as a Percentage of Total Variance for Selected Measures among Persons Aged 12 or Older: 2007 NSDUH

		Variance Component as a Percentage of Total Variance			
Variable	**Mean**	σ^2_{SSR}	$\sigma^2_{segment}$	σ^2_{DU}	σ^2_{person}
Past Month Cigarette Use	0.2424	1.3907	2.2335	25.5099	70.8660
Past Month Alcohol Use	0.5114	2.4960	2.8418	20.8137	73.8484
Past Month Use of Any Illicit Drug	0.0801	0.9811	1.3498	16.6156	81.0535
Past Month Use of Any Illicit Drug but Marijuana	0.0374	0.3394	0.7368	7.9423	90.9815
Past Month Cocaine Use	0.0084	0.1268	0.4461	10.4990	88.9281
Past Year, Dependent on Illicit Drugs	0.0189	0.2914	0.2783	10.0255	89.4048
Past Year, Dependent on Alcohol[1]	0.0286	0.1647	0.0000	6.4544	93.3810
Past Year, Received Treatment for Illicit Drug Use	0.0087	0.0549	0.0000	9.8298	90.1153
Past Year, Received Treatment for Alcohol Use[2]	0.0054	0.0943	0.1496	11.5096	88.2465

[1] Dependent on alcohol, but not dependent on illicit drugs.

[2] Received treatment for alcohol use, but not for illicit drug use.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007.

2.3.2 Unequal Weighting Effects

The UWE for a domain, *d*, is

$$UWE_d = \frac{n_d \sum_{i \in d} w_i^2}{\left(\sum_{i \in d} w_i\right)^2} \approx 1 + CV_{w_d}^2$$

where

$$CV_{w_d} = \sqrt{\frac{\frac{1}{n_d - 1}\left(\sum_{i \in d} w_i^2 - (\sum_{i \in d} w_i / n_d)^2\right)}{(\sum_{i \in d} w_i / n_d)^2}}$$ = coefficient of variation (CV) of the domain weights.

Table 2.3 shows the UWEs for the 2007 NSDUH national estimates, by five age groups.

Table 2.3 Unequal Weighting Effects for the National Estimates: 2007 NSDUH

Age Group	Unequal Weighting Effect
12 or Older	3.58627
12 to 17	1.73779
18 to 25	1.75686
26 to 34	1.75555
35 to 49	1.62846
50 or Older	1.66734

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007.

2.3.3 Averages and Coefficients of Variation for Respondents per Dwelling Unit and Respondents per Segment

The parametric models also require the average cluster size and the CV of the cluster size at the segment and dwelling unit levels, as shown in Table 2.4.

Table 2.4 Average Cluster Sizes and Coefficients of Variation: 2007 NSDUH

Age Group Domain (*d*)	Weighted *n*	*n*	Respondents/DU		Respondents/Segment	
			$\bar{m}_{d,DU}$	$CV_{m_{d,DU}}$	$\bar{m}_{d,seg}$	$CV_{m_{d,seg}}$
12 to 17	25,241,088	22,433	0.1586	2.7205	3.1157	0.8168
18 to 25	32,730,854	22,187	0.1568	2.7844	3.0815	1.0617
26 to 34	35,300,404	6,908	0.0488	4.8620	0.9594	1.3283
35 to 49	64,843,855	9,976	0.0705	3.8593	1.3856	1.0113
50 or Older	89,729,008	6,366	0.0450	5.0413	0.8842	1.2866
26 or Older	189,873,266	23,250	0.1643	2.5399	3.2292	0.7627
12 or Older	247,845,208	67,870	0.4797	1.5180	9.4264	0.5804

CV = coefficients of variation; DU = dwelling unit.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007.

2.3.4 Prevalence Estimates, by Age Group

Table 2.5 shows the variation in the key estimates for five age groups: 12 to 17, 18 to 25, 26 to 34, 35 to 49, and 50 or older. These are the age groups for which sampling rates are specified (within State), as well as the age groups used as main effect poststratification controls in the weight calibration process. The age group-specific estimates are used in the variance models as described in the next section.

Table 2.5 Prevalence Rates, by Age Group: Weighted Results, 2007 NSDUH

Variable	12 to 17	18 to 25	26 to 34	35 to 49	50 or Older
Past Month Cigarette Use	0.098	0.362	0.334	0.282	0.174
Past Month Alcohol Use	0.159	0.612	0.626	0.594	0.469
Past Month Use of Any Illicit Drug	0.095	0.197	0.109	0.072	0.028
Past Month Use of Any Illicit Drug but Marijuana	0.047	0.081	0.050	0.035	0.016
Past Month Cocaine Use	0.004	0.017	0.014	0.010	0.003
Past Year, Dependent on Illicit Drugs	0.023	0.054	0.025	0.017	0.004
Past Year, Dependent on Alcohol[1]	0.015	0.057	0.038	0.034	0.014
Past Year, Received Treatment for Illicit Drug Use	0.009	0.014	0.011	0.011	0.004
Past Year, Received Treatment for Alcohol Use[2]	0.002	0.010	0.006	0.007	0.003

[1] Dependent on alcohol, but not dependent on illicit drugs.

[2] Received treatment for alcohol use, but not for illicit drug use.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007.

2.3.5 Variance Models

The goal is to determine the optimal number of clusters and respondents needed to minimize the costs associated with NSDUH data collection, subject to multiple precision requirements established by the Substance Abuse and Mental Health Services Administration (SAMHSA). The precision requirements are that the expected relative standard error (RSE) on a prevalence of 10 percent not exceed the following:

- 3.00 percent for total population statistics, and
- 5.00 percent for statistics in three age group domains: 12 to 17, 18 to 25, and 26 or older.

For this reason, the cluster optimization will focus on the following age domains:

- 12 to 17,
- 18 to 25,
- 26 or older, and
- 12 or older.

Three variance models, prepared for NSDUH national estimates by age group, are described as follows.

Model 1: For the first two age-specific domains (12 to 17 and 18 to 25) designated by the subscript *a*,

$$Var(\hat{p}_a) = \frac{p_a(1-p_a)}{n_a} UWE_a \{\sigma^2_{seg}\,\overline{m}_{a,seg}\,(1 + CV^2_{m_{a,seg}}) + \sigma^2_{DU}\overline{m}_{a,DU}(1 + CV^2_{m_{a,DU}}) + \sigma^2_{person}\},$$

where

$\sigma^2_{segment}$, σ^2_{DU}, and σ^2_{person} = variance components as a percentage of the total variance;

$\overline{m}_{a,seg} = n_a / n_{seg}$ average number of respondents per segment in age group a; and

$\overline{m}_{a,DU}$ = average number of respondents per selected dwelling unit in age group a.

For the purposes of this chapter's discussion, it is convenient to express the variance as the following function of the numbers of segments and respondents:

$$Var(\hat{p}_a) = \frac{Var_{a,seg}}{n_{seg}} + \frac{Var_{a,person}}{n_a},$$

where

$Var_{a,seg}$ = segment component of variance, and
$Var_{a,person}$ = person component of variance.

Model 1 becomes

$$Var(\hat{p}_a) = \frac{p_a(1-p_a)UWE_a\sigma^2_{seg}(1 + CV^2_{m_{a,seg}})}{n_{seg}} + \frac{p_a(1-p_a)UWE_a\left\{\sigma^2_{DU}\overline{m}_{a,DU}(1+CV^2_{m_{a,DU}})+\sigma^2_{person}\right\}}{n_a}.$$

Model 2: For all persons aged 26 or older (three age groups combined),

$$Var(\hat{p}_{26+}) = \frac{\sum_{a=3}^{5} W_a^2 p_a(1-p_a)UWE_a\sigma^2_{seg}(1 + CV^2_{m_{a,seg}})}{n_{seg}} + \frac{\sum_{a=3}^{5} W_a^2 p_a(1-p_a)UWE_a\left\{\sigma^2_{DU}\overline{m}_{a,DU}(1+CV^2_{m_{a,DU}})+\sigma^2_{person}\right\}/f_a}{n_{26+}},$$

where

$f_a = \frac{n_a}{n_{26+}}$ and $W_a = \frac{N_a}{N_{26+}}$ = associated population fraction for age group a.

Model 3: For all persons aged 12 or older (five age groups combined),

$$Var(\hat{p}_{12+}) = \frac{\sum_{a=1}^{5} W_a^2 p_a(1-p_a)UWE_a\sigma^2_{seg}(1 + CV^2_{m_{a,seg}})}{n_{seg}} + \frac{\sum_{a=1}^{5} W_a^2 p_a(1-p_a)UWE_a\left\{\sigma^2_{DU}\overline{m}_{a,DU}(1+CV^2_{m_{a,DU}})+\sigma^2_{person}\right\}/f_a}{n_{12+}},$$

where

$$f_a = \frac{n_a}{n_{12+}}, \text{ and}$$

$$W_a = \frac{N_a}{N_{12+}}.$$

2.4 Some Univariate Calculations

Using Model 3, the segment and person-level components of variance were estimated for each of the nine measures shown in Table 2.5 for all persons aged 12 or older. Then nine options were considered for the number and size of the segments. These options cover a range of segment counts and sizes to demonstrate the impact on cost and precision. Table 2.6 shows the estimated cost reduction and maximum impact on precision for each of these options when compared with the current design. The RSE standardized to a prevalence of 10 percent has been computed separately for each measure. Using this information, the table reports the maximum increase in standardized relative standard error (SRSE) across the nine measures.

Table 2.6 Impact of Alternative Designs on Cost and Precision

Option	Segments	Persons per Segment	Total Persons	Variable Costs	Cost Reduction Compared with Current Design	Maximum Increase in SRSE[1]
Current	7,200	9.375	67,500	$29,504,025		
1	3,600	18.75	67,500	$25,644,825	$3,859,200	0.04%
2	3,800	18.75	71,250	$27,069,538	$2,434,488	-0.02%
3	3,000	24	72,000	$26,454,000	$3,050,025	-0.01%
4	4,800	14	67,200	$26,834,400	$2,669,625	0.03%
5	4,000	15	60,000	$23,653,000	$5,851,025	0.16%
6	4,000	12	48,000	$19,780,000	$9,724,025	0.44%
7	3,600	12	43,200	$17,802,000	$11,702,025	0.58%
8	1,453	28.6	41,556	$14,969,750	$14,534,275	0.73%
9	2,471	28.6	70,671	$25,457,848	$4,046,177	0.03%

SRSE = standardized relative standard error; the SRSE reflects the relative standard error (RSE) for a true value of 0.10.

[1]Maximum Increase in SRSE is computed over the nine variables listed in Table 2.5.

Note that Model 3 incorporates UWE parameters and the current sample age distribution from the current design; therefore, this analysis focuses on increasing the segment sample size and reducing the number of segments. Consider Option 9 in Table 2.6. By increasing the segment size to 28.6 persons, decreasing the number of segments to 2,471, and increasing the total sample by approximately 3,170 persons, the maximum increase in the SRSE over the nine variables studied is 0.03 percent.[7] The Option 9 design will save $4 million. The corresponding decrease in the number of segments in small sample States will be from a current 96 to about 31.

[7] The current design specifications require an SRSE of 3 percent or less for national estimates for all nine variables. It should be noted that even the largest maximum increase in SRSE in Table 2.6 does not correspond to a projected SRSE of greater than 3 percent.

This analysis ignores potential impacts on substate estimates because of the sparse distribution of sample areas for small area estimation (SAE) purposes.

2.5 Simultaneous Solution

In addition to describing the impact on cost and variance for several cluster size options, a formal optimization for NSDUH was completed to provide a simultaneous solution for key prevalence measures across major age groups (12 to 17, 18 to 25, 26 or older, and 12 or older). A solution to this multiple constraint optimization was achieved using Chromy's algorithm (Chromy, 1987).

In summary, given specified variance bounds (V*), the algorithm simultaneously solves for n_{seg} and $n_a, a = 1, 2, ..., 5$ in Models 1 through 3 (see Section 2.4) while minimizing cost. In other words, the algorithm minimizes

$$C = C_{seg} n_{seg} + \sum_{a=1}^{5} C_a n_a$$

subject to

(1) $Var(\hat{p}) \le V^*$,
(2) $n_{seg} \ge 0$, and
(3) $n_a \ge 0$ for $a = 1, 2, ..., 5$.

The process is iterative until the optimum solution is reached. In general, the solution will involve one or more age-specific measures for which $Var(\hat{p}) = V^*$. This age-specific measure is called the "driver" measure.

Models 1 through 3 were used to derive the segment and person-level components of variance for each major age group. To assess the accuracy of the models, the mean and standard error (SE) for selected measures were calculated by the model and by SUDAAN® (RTI International, 2004).

Table 2.7 shows the summary statistics for the percentage difference in the means and SEs of selected prevalence estimates calculated by the model and by SUDAAN. Although the means calculated by both methods were the same, the model produced an SE that is slightly low; the median difference was 1.0 percent below SUDAAN's SE.

As shown in Table 2.8, two solutions were prepared for the multiple constraint optimization. Tables 2.9 and 2.10 show the expected RSEs by age group for Solutions 1 and 2, respectively. In these tables, the specified SRSEs are the variance bounds for a prevalence of 10 percent by age group. The specified RSEs are the corresponding RSEs given the mean p. The achieved RSEs in these tables are the RSEs resulting from the model $(= SE[p]/p)$, and the achieved SRSEs are those RSEs standardized to a prevalence of 10 percent.

Table 2.7 Summary Statistics for the Percentage Difference in Mean and Standard Error of Selected Measures and Age Groups When Calculated by the Model and by SUDAAN

Statistic	Relative Difference in Mean	Relative Difference in Standard Error
Mean	0.000	0.002
Min	0.000	-0.094
Q1	0.000	-0.044
Median	0.000	-0.010
Q3	0.000	0.021
Max	0.000	0.266

Note: The relative difference was calculated as follows: [*Model Estimate – SUDAAN Estimate*] / *SUDAAN Estimate*.

Table 2.8 Results of Formal Optimization

Result	Current	Solution 1	Solution 2
Number of Segments	7,200	1,453	2,471
Number of Persons	67,500	41,552	70,698
Average Cluster Size	9.3750	28.5974	28.6111

For Solution 1, the variance bounds were set at the precision requirements specified by SAMHSA. That is, the SRSE bounds were set at 3 percent for the total population estimates and 5 percent for the age group-specific estimates, as shown in Table 2.9. The optimal number of clusters is 1,453, and the optimal number of persons is 41,552, for an optimal cluster size of 28.6. Solution 1 is driven by past month alcohol use among persons aged 12 or older.

For Solution 2, the specified bounds were tightened so that they would be closer to the achieved SRSEs from the 2007 NSDUH. Similar to Solution 1, the optimal cluster size for this solution is 28.6. However, the optimal number of clusters is 2,471, and the optimal number of persons is 70,698 (Table 2.8). As in Solution 1, Solution 2 is driven by past month alcohol use among persons aged 12 or older (Table 2.10).

Note that Options 8 and 9 in Table 2.6 correspond to Solutions 1 and 2 in the simultaneous solution.[8] By increasing the cluster size to 28.6 and reducing the number of clusters to approximately 2,500 (Option 9), a large cost savings is realized with minimal loss of precision.

The simultaneous solutions assume the same proportional distribution of the sample among States. A national sample that did not require minimum sample sizes at the State level will produce a much lower UWE and therefore requires a smaller overall sample size. The optimum cluster size, however, will not be different from what is shown.

Although the cost-variance analysis suggests that a larger sample per segment will be more nearly optimal, the current design requires that the sample be drawn from a sufficient

[8] The differences in the total number of persons are due to the rounding of the average cluster size.

Table 2.9 Expected Relative Standard Errors, by Age Group: Solution 1 (Cluster Optimization)

Domain	Variable	Specified		Achieved	
		SRSE	**RSE**	**SRSE**	**RSE**
12 to 17	Past Month Cigarette Use	0.05000	0.05047	0.03981	0.04018
	Past Month Alcohol Use	0.05000	0.03829	0.04071	0.03117
	Past Month Use of Any Illicit Drug	0.05000	0.05131	0.03752	0.03850
	Past Month Use of Any Illicit Drug but Marijuana	0.05000	0.07523	0.03576	0.05381
	Past Month Cocaine Use	0.05000	0.25967	0.03524	0.18304
	Past Year, Dependent on Illicit Drugs	0.05000	0.10896	0.03479	0.07581
	Past Year, Dependent on Alcohol but Not Illicit Drugs	0.05000	0.13328	0.03393	0.09046
	Past Year Received Treatment for Illicit Drug Use	0.05000	0.17395	0.03414	0.11877
	Past Year Received Treatment for Alcohol Use but Not Illicit Drug Use	0.05000	0.42344	0.03459	0.29294
18 to 25	Past Month Cigarette Use	0.05000	0.02215	0.04154	0.01840
	Past Month Alcohol Use	0.05000	0.01327	0.04274	0.01134
	Past Month Use of Any Illicit Drug	0.05000	0.03360	0.03871	0.02601
	Past Month Use of Any Illicit Drug but Marijuana	0.05000	0.05624	0.03652	0.04107
	Past Month Cocaine Use	0.05000	0.12491	0.03582	0.08948
	Past Year, Dependent on Illicit Drugs	0.05000	0.06947	0.03524	0.04897
	Past Year, Dependent on Alcohol but Not Illicit Drugs	0.05000	0.06773	0.03416	0.04628
	Past Year Received Treatment for Illicit Drug Use	0.05000	0.13749	0.03439	0.09457
	Past Year Received Treatment for Alcohol Use but Not Illicit Drug Use	0.05000	0.16714	0.03496	0.11688
26 or Older	Past Month Cigarette Use	0.05000	0.02958	0.03603	0.02131
	Past Month Alcohol Use	0.05000	0.01535	0.03760	0.01155
	Past Month Use of Any Illicit Drug	0.05000	0.06726	0.03313	0.04456
	Past Month Use of Any Illicit Drug but Marijuana	0.05000	0.09698	0.03311	0.06422
	Past Month Cocaine Use	0.05000	0.19344	0.03169	0.12260
	Past Year, Dependent on Illicit Drugs	0.05000	0.14966	0.03144	0.09409
	Past Year, Dependent on Alcohol but Not Illicit Drugs	0.05000	0.10323	0.03281	0.06773
	Past Year Received Treatment for Illicit Drug Use	0.05000	0.18937	0.03253	0.12321
	Past Year Received Treatment for Alcohol Use but Not Illicit Drug Use	0.05000	0.23315	0.03349	0.15618
12 or Older	Past Month Cigarette Use	0.03000	0.01768	0.02919	0.01720
	Past Month Alcohol Use	**0.03000**	**0.00977**	**0.03000**	**0.00977**
	Past Month Use of Any Illicit Drug	0.03000	0.03388	0.02384	0.02693
	Past Month Use of Any Illicit Drug but Marijuana	0.03000	0.05073	0.02391	0.04042
	Past Month Cocaine Use	0.03000	0.10883	0.02426	0.08801
	Past Year, Dependent on Illicit Drugs	0.03000	0.07205	0.02160	0.05187
	Past Year, Dependent on Alcohol but Not Illicit Drugs	0.03000	0.05831	0.02496	0.04852
	Past Year Received Treatment for Illicit Drug Use	0.03000	0.10657	0.02478	0.08804
	Past Year Received Treatment for Alcohol Use but Not Illicit Drug Use	0.03000	0.13632	0.02633	0.11963

RSE = relative standard error; SRSE = standardized relative standard error.

Note: Bolded line indicates driver.

Table 2.10 Expected Relative Standard Errors, by Age Group: Solution 2 (Cluster Optimization)

Domain	Variable	Specified		Achieved	
		SRSE	RSE	SRSE	RSE
12 to 17	Past Month Cigarette Use	0.03150	0.03180	0.03052	0.03081
	Past Month Alcohol Use	0.03150	0.02412	0.03121	0.02390
	Past Month Use of Any Illicit Drug	0.03150	0.03233	0.02876	0.02952
	Past Month Use of Any Illicit Drug but Marijuana	0.03150	0.04739	0.02742	0.04125
	Past Month Cocaine Use	0.03150	0.16359	0.02702	0.14033
	Past Year, Dependent on Illicit Drugs	0.03150	0.06864	0.02667	0.05812
	Past Year, Dependent on Alcohol but Not Illicit Drugs	0.03150	0.08397	0.02602	0.06935
	Past Year Received Treatment for Illicit Drug Use	0.03150	0.10959	0.02617	0.09105
	Past Year Received Treatment for Alcohol Use but Not Illicit Drug Use	0.03150	0.26677	0.02652	0.22458
18 to 25	Past Month Cigarette Use	0.03350	0.01484	0.03185	0.01411
	Past Month Alcohol Use	0.03350	0.00889	0.03277	0.00870
	Past Month Use of Any Illicit Drug	0.03350	0.02251	0.02967	0.01994
	Past Month Use of Any Illicit Drug but Marijuana	0.03350	0.03768	0.02800	0.03149
	Past Month Cocaine Use	0.03350	0.08369	0.02746	0.06860
	Past Year, Dependent on Illicit Drugs	0.03350	0.04654	0.02702	0.03754
	Past Year, Dependent on Alcohol but Not Illicit Drugs	0.03350	0.04538	0.02619	0.03548
	Past Year Received Treatment for Illicit Drug Use	0.03350	0.09212	0.02637	0.07250
	Past Year Received Treatment for Alcohol Use but Not Illicit Drug Use	0.03350	0.11199	0.02680	0.08960
26 or Older	Past Month Cigarette Use	0.02900	0.01716	0.02762	0.01634
	Past Month Alcohol Use	0.02900	0.00891	0.02883	0.00885
	Past Month Use of Any Illicit Drug	0.02900	0.03901	0.02540	0.03416
	Past Month Use of Any Illicit Drug but Marijuana	0.02900	0.05625	0.02538	0.04923
	Past Month Cocaine Use	0.02900	0.11219	0.02429	0.09399
	Past Year, Dependent on Illicit Drugs	0.02900	0.08680	0.02410	0.07214
	Past Year, Dependent on Alcohol but Not Illicit Drugs	0.02900	0.05987	0.02515	0.05193
	Past Year Received Treatment for Illicit Drug Use	0.02900	0.10983	0.02494	0.09446
	Past Year Received Treatment for Alcohol Use but Not Illicit Drug Use	0.02900	0.13523	0.02568	0.11973
12 or Older	Past Month Cigarette Use	0.02300	0.01356	0.02238	0.01319
	Past Month Alcohol Use	**0.02300**	**0.00749**	**0.02300**	**0.00749**
	Past Month Use of Any Illicit Drug	0.02300	0.02598	0.01828	0.02065
	Past Month Use of Any Illicit Drug but Marijuana	0.02300	0.03889	0.01833	0.03099
	Past Month Cocaine Use	0.02300	0.08343	0.01860	0.06747
	Past Year, Dependent on Illicit Drugs	0.02300	0.05524	0.01656	0.03976
	Past Year, Dependent on Alcohol but Not Illicit Drugs	0.02300	0.04470	0.01914	0.03719
	Past Year Received Treatment for Illicit Drug Use	0.02300	0.08171	0.01900	0.06750
	Past Year Received Treatment for Alcohol Use but Not Illicit Drug Use	0.02300	0.10451	0.02018	0.09172

RSE = relative standard error; SRSE = standardized relative standard error.

Note: Bolded line indicates driver.

number of clusters within each State in order to produce small area estimates. For substate estimates using SAE, the more segments that are available in a State, the more likely that smaller areas of interest can be covered. A NSDUH design change that better accommodates small area estimates will need to consider having a sufficient number of clusters in each State to produce these estimates.

2.6 Conclusions

This analysis of the cost-variance tradeoffs suggests that a larger sample per segment than what is currently being used for NSDUH will be nearly optimal for producing national estimates. In the formal optimization, which solved across several drug use, dependence, and treatment variables for multiple age groups simultaneously, it was determined that the optimal cluster size was around 29 persons per segment. Univariate calculations confirmed that this cluster size will provide as nearly as precise estimates at a greatly reduced cost (approximate savings of $4 million). To accommodate State-level estimates, a minimum number of clusters per State may need to be considered.

3. Oversampling Optimization

3.1 Introduction

Since 1999, the National Survey on Drug Use and Health (NSDUH) has been designed to target 67,500 interviews annually that are equally allocated across three age groups: 12 to 17, 18 to 25, and 26 or older. The large sample size allows the Substance Abuse and Mental Health Services Administration (SAMHSA) to report demographic subgroups at the national level with adequate precision without the need to oversample specially targeted demographics as was required prior to 1999 (Morton et al., 2008b). The allocation of the NSDUH sample to States also has been fixed since 1999 with eight large sample States[9] targeting 3,600 completed interviews per year and the remaining 42 States and the District of Columbia targeting 900 interviews annually. This chapter focuses on the optimal sample distribution with respect to demographic and geographic groups in NSDUH. Both the preliminary and final oversampling optimization analyses are included in this chapter.

The sample optimization work that has been completed in recent years has focused on the 26 or older age group. The preliminary oversampling analysis expanded the optimization work to additional ages and other demographic groups. Sections 3.2 through 3.4 present findings from the preliminary oversampling analysis. Specifically, Section 3.2 investigates estimates and variances for youths aged 12 to 17, the 50 or older age group, several racial/ethnic groups, and pregnant women to identify candidates for oversampling in the redesigned NSDUH. In Section 3.3, the variance inputs for the age group optimization models are described. Then, Section 3.4 describes several solutions to the formal optimization process.

Chapter 2 addresses cluster size optimization based on cost and variance models. It concludes that much larger cluster sizes (up to about 29) will reduce costs while maintaining equivalent precision for national estimates. Nine specific designs are presented in Table 2.6 showing cost savings and the impacts on standardized relative standard errors (SRSEs).

Using results from the preliminary oversampling optimization and the cluster size optimization, staff at SAMHSA subsequently refined the specifications for completing the oversampling issues component. Section 3.5 describes several alternative sample designs based on SAMHSA's specifications that are associated with key outcome measures. The alternative sample designs then are evaluated in Sections 3.6 through 3.12. A summary of the chapter's conclusions appears in Section 3.13. Also, the variance of a proportion p under the competing designs is treated in Appendix A.

Several additional options, referred to as Options A, B, and C, were later evaluated and are included in Appendix B. Finally, Appendix C discusses two options for guaranteeing a 3-year minimum sample size of 200 in each of Hawaii's four major counties.

---

[9] The large sample States are California, Florida, Illinois, Michigan, New York, Ohio, Pennsylvania, and Texas.

3.2 Estimates by Age and Other Demographic Subgroups

The first task in the oversampling optimization study was to identify demographic subgroups who may be candidates for oversampling. Estimates and variances were examined for various age groups (Tables 3.1a and 3.1b), for some racial/ethnic groups (Tables 3.2a and 3.2b), and for pregnant women (Tables 3.3a and 3.3b). Finer race groups also were considered; however, these groups may be of limited analytical interest because tables by finer race groups are no longer being prepared as part of the standard set of detailed tables produced each year.

As shown in Table 3.1a, there were some large differences in drug use rates among youths aged 12 to 14 and youths aged 15 to 17. In addition, within the 50 or older age group, drug use rates dropped off dramatically after the age of 65. For these reasons, the following seven age groups were selected for the formal optimization: 12 to 14, 15 to 17, 18 to 25, 26 to 34, 35 to 49, 50 to 64, and 65 or older.

Although there were some differences in drug use rates by race/ethnicity, the differences were not as substantial as those by age group (Table 3.2a). Section 3.5 discusses a preliminary plan for verifying that estimates by race and ethnicity have sufficient precision under the optimal age group allocation.

Table 3.3a shows that there were large differences in drug use estimates between pregnant and not pregnant women aged 12 to 44. In previous years, estimates for pregnant women required pooling multiple years of NSDUH data in order to report estimates with sufficient precision. As discussed in Section 3.5, SAMHSA was presented the option to determine whether this demographic group was of great enough interest to warrant oversampling.

3.3 Overview of Variance Modeling for Age Group Optimization

Parametric variance models allow the sampling statistician to represent the variance of key estimates as a function of sample design parameters. The 2007 NSDUH data were used to estimate some of the key population parameters needed for these models.

The required parameters include the following:

- variance components,
- unequal weighting effects (UWEs),
- averages and coefficients of variation for respondents per segment,
- averages and coefficients of variation for respondents per dwelling unit, and
- prevalence estimates by age group.

Table 3.1a Prevalence Rates, by Age Group: Weighted Results, Based on 2007 NSDUH

Variable	12-14	15-17	18-25	26-34	35-49
Past Month Cigarette Use	0.0320	0.1610	0.3616	0.3342	0.2825
Past Month Alcohol Use	0.0568	0.2562	0.6121	0.6258	0.5940
Past Month Use of Any Illicit Drug	0.0447	0.1434	0.1975	0.1095	0.0715
Past Month Use of Any Illicit Drug But Marijuana	0.0305	0.0622	0.0807	0.0498	0.0354
Past Month Cocaine Use	0.0010	0.0070	0.0175	0.0141	0.0103
Past Year, Dependent on Illicit Drugs	0.0070	0.0379	0.0544	0.0245	0.0166
Past Year, Dependent on Alcohol[1]	0.0052	0.0250	0.0571	0.0377	0.0342
Past Year Received Treatment for Illicit Drugs	0.0026	0.0152	0.0145	0.0106	0.0115
Past Year Received Treatment for Alcohol Use[2]	0.0007	0.0024	0.0098	0.0061	0.0073

Variable	50-54	55-59	60-64	50-64	65+
Past Month Cigarette Use	0.2673	0.2204	0.1884	0.2315	0.0900
Past Month Alcohol Use	0.5702	0.5200	0.4756	0.5294	0.3805
Past Month Use of Any Illicit Drug	0.0572	0.0408	0.0189	0.0420	0.0066
Past Month Use of Any Illicit Drug But Marijuana	0.0264	0.0222	0.0145	0.0220	0.0061
Past Month Cocaine Use	0.0052	0.0037	*	0.0034	0.0014
Past Year, Dependent on Illicit Drugs	0.0076	0.0089	*	0.0062	0.0014
Past Year, Dependent on Alcohol[1]	0.0232	0.0238	0.0167	0.0218	0.0031
Past Year Received Treatment for Illicit Drugs	0.0114	0.0035	0.0024	0.0064	*
Past Year Received Treatment for Alcohol Use[2]	0.0056	0.0064	0.0012	0.0048	0.0006

*Low precision; no estimate reported.

[1]Dependent on Alcohol but Not Dependent on Illicit Drugs.

[2]Received Treatment for Alcohol Use but Not for Illicit Drug Use.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007.

Table 3.1b Standard Errors of Prevalence Rates, by Age Group, Based on 2007 NSDUH

Variable	12-14	15-17	18-25	26-34	35-49
Past Month Cigarette Use	0.0022	0.0045	0.0048	0.0079	0.0058
Past Month Alcohol Use	0.0028	0.0055	0.0052	0.0083	0.0067
Past Month Use of Any Illicit Drug	0.0026	0.0046	0.0039	0.0052	0.0033
Past Month Use of Any Illicit Drug But Marijuana	0.0022	0.0030	0.0025	0.0035	0.0024
Past Month Cocaine Use	0.0004	0.0009	0.0012	0.0020	0.0015
Past Year, Dependent on Illicit Drugs	0.0010	0.0021	0.0021	0.0024	0.0016
Past Year, Dependent on Alcohol[1]	0.0009	0.0019	0.0022	0.0029	0.0023
Past Year Received Treatment for Illicit Drugs	0.0005	0.0015	0.0011	0.0016	0.0014
Past Year Received Treatment for Alcohol Use[2]	0.0003	0.0005	0.0009	0.0012	0.0010

Variable	50-54	55-59	60-64	50-64	65+
Past Month Cigarette Use	0.0140	0.0143	0.0162	0.0087	0.0070
Past Month Alcohol Use	0.0158	0.0180	0.0210	0.0111	0.0131
Past Month Use of Any Illicit Drug	0.0083	0.0070	0.0059	0.0044	0.0019
Past Month Use of Any Illicit Drug But Marijuana	0.0056	0.0056	0.0055	0.0033	0.0019
Past Month Cocaine Use	0.0025	0.0020	*	0.0012	0.0009
Past Year, Dependent on Illicit Drugs	0.0027	0.0034	*	0.0016	0.0010
Past Year, Dependent on Alcohol[1]	0.0055	0.0057	0.0075	0.0035	0.0011
Past Year Received Treatment for Illicit Drugs	0.0035	0.0019	0.0022	0.0017	*
Past Year Received Treatment for Alcohol Use[2]	0.0024	0.0027	0.0008	0.0014	0.0004

*Low precision; no estimate reported.

[1]Dependent on Alcohol but Not Dependent on Illicit Drugs.

[2]Received Treatment for Alcohol Use but Not for Illicit Drug Use.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007.

Table 3.2a Prevalence Rates, by Ethnicity and Race: Weighted Results, Based on 2007 NSDUH

Variable	Hispanic	White, Not Hispanic
Past Month Cigarette Use	0.2050	0.2562
Past Month Alcohol Use	0.4208	0.5613
Past Month Use of Any Illicit Drug	0.0660	0.0818
Past Month Use of Any Illicit Drug But Marijuana	0.0337	0.0392
Past Month Cocaine Use	0.0103	0.0085
Past Year, Dependent on Illicit Drugs	0.0171	0.0185
Past Year, Dependent on Alcohol[1]	0.0257	0.0300
Past Year Received Treatment for Illicit Drugs	0.0059	0.0087
Past Year Received Treatment for Alcohol Use[2]	0.0035	0.0054

Variable	Black, Not Hispanic	Other, Not Hispanic
Past Month Cigarette Use	0.2318	0.1936
Past Month Alcohol Use	0.3934	0.3895
Past Month Use of Any Illicit Drug	0.0950	0.0643
Past Month Use of Any Illicit Drug But Marijuana	0.0357	0.0296
Past Month Cocaine Use	0.0081	0.0029
Past Year, Dependent on Illicit Drugs	0.0269	0.0122
Past Year, Dependent on Alcohol[1]	0.0260	0.0237
Past Year Received Treatment for Illicit Drugs	0.0136	0.0061
Past Year Received Treatment for Alcohol Use[2]	0.0075	0.0046

[1]Dependent on Alcohol but Not Dependent on Illicit Drugs.

[2]Received Treatment for Alcohol Use but Not for Illicit Drug Use.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007.

Table 3.2b Standard Errors of Prevalence Rates, by Ethnicity and Race, Based on 2007 NSDUH

Variable	Hispanic	White, Not Hispanic
Past Month Cigarette Use	0.0082	0.0039
Past Month Alcohol Use	0.0103	0.0047
Past Month Use of Any Illicit Drug	0.0040	0.0022
Past Month Use of Any Illicit Drug But Marijuana	0.0030	0.0014
Past Month Cocaine Use	0.0017	0.0007
Past Year, Dependent on Illicit Drugs	0.0019	0.0009
Past Year, Dependent on Alcohol[1]	0.0027	0.0014
Past Year Received Treatment for Illicit Drugs	0.0012	0.0007
Past Year Received Treatment for Alcohol Use[2]	0.0009	0.0005

Variable	Black, Not Hispanic	Other, Not Hispanic
Past Month Cigarette Use	0.0088	0.0114
Past Month Alcohol Use	0.0108	0.0154
Past Month Use of Any Illicit Drug	0.0052	0.0066
Past Month Use of Any Illicit Drug But Marijuana	0.0037	0.0053
Past Month Cocaine Use	0.0017	0.0008
Past Year, Dependent on Illicit Drugs	0.0034	0.0017
Past Year, Dependent on Alcohol[1]	0.0032	0.0031
Past Year Received Treatment for Illicit Drugs	0.0028	0.0016
Past Year Received Treatment for Alcohol Use[2]	0.0020	0.0017

[1]Dependent on Alcohol but Not Dependent on Illicit Drugs.

[2]Received Treatment for Alcohol Use but Not for Illicit Drug Use.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007.

Table 3.3a Prevalence Rates, by Pregnancy among Women Aged 12 to 44: Weighted Results, Based on 2007 NSDUH

Variable	Pregnant	Not Pregnant
Past Month Cigarette Use	0.1567	0.2513
Past Month Alcohol Use	0.1197	0.4926
Past Month Use of Any Illicit Drug	0.0623	0.0880
Past Month Use of Any Illicit Drug But Marijuana	0.0247	0.0424
Past Month Cocaine Use	0.0062	0.0072
Past Year, Dependent on Illicit Drugs	0.0437	0.0203
Past Year, Dependent on Alcohol[1]	0.0114	0.0273
Past Year Received Treatment for Illicit Drugs	0.0058	0.0081
Past Year Received Treatment for Alcohol Use[2]	0.0031	0.0033

[1]Dependent on Alcohol but Not Dependent on Illicit Drugs.

[2]Received Treatment for Alcohol Use but Not for Illicit Drug Use.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007.

Table 3.3b Standard Errors of Prevalence Rates, by Pregnancy among Women Aged 12 to 44, Based on 2007 NSDUH

Variable	Pregnant	Not Pregnant
Past Month Cigarette Use	0.0163	0.0044
Past Month Alcohol Use	0.0147	0.0054
Past Month Use of Any Illicit Drug	0.0113	0.0026
Past Month Use of Any Illicit Drug But Marijuana	0.0076	0.0018
Past Month Cocaine Use	0.0045	0.0007
Past Year, Dependent on Illicit Drugs	0.0107	0.0011
Past Year, Dependent on Alcohol[1]	0.0032	0.0013
Past Year Received Treatment for Illicit Drugs	0.0027	0.0009
Past Year Received Treatment for Alcohol Use[2]	0.0018	0.0005

[1]Dependent on Alcohol but Not Dependent on Illicit Drugs.

[2]Received Treatment for Alcohol Use but Not for Illicit Drug Use.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007.

3.3.1 Variance Components

The notation for variance components associated with stratum, segment, dwelling unit, and person is shown in Table 3.4.

Table 3.4 Variance Component Notation

Level	Symbol
State Sampling Region	σ^2_{SSR}
Segment	$\sigma^2_{segment}$
Dwelling Unit	σ^2_{DU}
Person	σ^2_{person}
Total	$\sigma^2 = \sigma^2_{SSR} + \sigma^2_{segment} + \sigma^2_{DU} + \sigma^2_{person}$

Variance components were estimated using the method of moments in SAS PROC NESTED with equal weights. Variance components were computed for nine substance use and treatment variables of interest, controlling for age group. Variance components for the nine selected measures are shown in Table 3.5.

Table 3.5 Variance Components of Residuals as a Percentage of Total Variance for Selected Measures, Based on 2007 NSDUH

Age	Variable	Mean	Variance Component as a Percent of Total Variance			
			σ^2_{SSR}	$\sigma^2_{segment}$	σ^2_{DU}	σ^2_{person}
12 or Older	Past Month Cigarette Use	0.2424	1.4146	2.2500	25.6207	70.7147
	Past Month Alcohol Use	0.5114	2.5338	2.8712	20.9846	73.6104
	Past Month Use of Any Illicit Drug	0.0801	0.9874	1.3420	16.5964	81.0742
	Past Month Use of Any Illicit Drug But Marijuana	0.0374	0.3340	0.7356	7.8665	91.0639
	Past Month Cocaine Use	0.0084	0.1255	0.4452	10.4601	88.9692
	Past Year, Dependent on Illicit Drugs	0.0189	0.2981	0.2872	9.9387	89.4760
	Past Year, Dependent on Alcohol[1]	0.0286	0.1670	0.0000	6.4707	93.3623
	Past Year Received Treatment for Illicit Drugs	0.0087	0.0574	0.0000	9.8185	90.1242
	Past Year Received Treatment for Alcohol Use[2]	0.0054	0.0933	0.1523	11.4884	88.2660

[1]Dependent on Alcohol but Not Dependent on Illicit Drugs.

[2]Received Treatment for Alcohol Use but Not for Illicit Drug Use.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007.

3.3.2 Unequal Weighting Effects

The UWE for a domain, *d*, is

$$UWE_d = \frac{n_d \sum_{i \in d} w_i^2}{\left(\sum_{i \in d} w_i\right)^2}$$

and was estimated using the population coefficient of variation of domain weights as

$$UWE_d = 1 + CV_{w_d}^2 .$$

Table 3.6 shows UWEs for 2007 national estimates by eight age groups. Use of these UWEs presumes that the relative State sample sizes are maintained. For small sample States, the relative sample size is 1/75th of the national sample. For large States, the relative size is 4/75th of the total.

Table 3.6 Unequal Weighting Effects for the National Estimates, Based on 2007 NSDUH

Age Group	Unequal Weighting Effect
12 or Older	3.58627
12 to 14	1.72707
15 to 17	1.74739
18 to 25	1.75686
26 to 34	1.75555
35 to 49	1.62846
50 to 64	1.67092
65 or Older	1.65374

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007.

3.3.3 Cluster Sizes and Cluster Size Variation

The parametric models also require the average cluster size and the coefficient of variation of the cluster size at the segment level and at the dwelling unit level as shown in Table 3.7.

Table 3.7 Average Cluster Sizes and Coefficients of Variation: Based on 2007 NSDUH

Age Group Domain (*d*)	Weighted *n*	*n*	Respondents/DU		Respondents/Segment	
			$\overline{m}_{d,DU}$	$CV_{m_{d,DU}}$	$\overline{m}_{d,seg}$	$CV_{m_{d,seg}}$
12 to 14	12,267,761	11,022	0.0779	3.6733	1.5308	0.9994
15 to 17	12,973,328	11,411	0.0807	3.5990	1.5849	0.9770
18 to 25	32,730,854	22,187	0.1568	2.7844	3.0815	1.0617
26 to 34	35,300,404	6,908	0.0488	4.8620	0.9594	1.3283
35 to 49	64,843,855	9,976	0.0705	3.8593	1.3856	1.0113
50 to 64	53,427,744	3,940	0.0278	6.3012	0.5472	1.5284
65 or Older	36,301,265	2,426	0.0171	8.2079	0.3369	2.0654
12 to 17	25,241,088	22,433	0.1586	2.7205	3.1157	0.8168
26 or Older	189,873,266	23,250	0.1643	2.5399	3.2292	0.7627
12 or Older	247,845,208	67,870	0.4797	1.5180	9.4264	0.5804

CV = coefficients of variation; DU = dwelling unit.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007.

3.3.4 Variance Models

The goal was to determine the optimal sample distribution subject to multiple precision requirements established by SAMHSA. In summary, the precision requirements are that the expected relative standard error (RSE) on a prevalence of 10 percent not exceed the following:

- 3.00 percent for total population statistics, and
- 5.00 percent for statistics in three age group domains: 12 to 17, 18 to 25, and 26 or older.

For this reason, the sample optimization focused on the following age domains:

- 12 to 17,
- 18 to 25,

- 26 or older, and
- 12 or older.

Model 1: For persons aged 12 to 17 (two age groups combined, where age group is designated by the subscript *a*),

$$Var(\hat{p}_{12-17}) = \sum_{a=1}^{2} W_a^2 \frac{p_a(1-p_a)}{n_a} UWE_a \{\sigma_{seg}^2 \bar{m}_{12-17,seg}(1 + CV_{m_{12-17,seg}}^2) + \sigma_{DU}^2 \bar{m}_{12-17,DU}(1+CV_{m_{12-17,DU}}^2) + \sigma_{per}^2\},$$

where $\sigma_{segment}^2$, σ_{DU}^2, and σ_{person}^2 are variance components as a percent of total variance, $\bar{m}_{12-17,seg}$ is the average number of respondents per segment in the 12 to 17 age group, and $\bar{m}_{12-17,DU}$ is the average number of respondents per selected dwelling unit in the 12 to 17 age group.

The term in the right brackets is treated as a fixed design effect for each domain estimate to be optimized. This general primary stratification and clustering effect is designated for the combined age domain, *d*, by $C_{eff,d}$. This term reflects the combined effects of primary sample stratification and clustering. Table 3.8 presents these combined effects for the four domains of interest and the nine NSDUH measures. The optimization process assumes the same number and size of segments as has been used in the current NSDUH design. Increasing the cluster size in the formal optimization process is an option discussed in Section 3.5.

Model 2: For all persons aged 18 to 25,

$$Var(\hat{p}_a) = \frac{p_a(1-p_a)}{n_a} UWE_a \{\sigma_{seg}^2 \bar{m}_{a,seg}(1 + CV_{m_{a,seg}}^2) + \sigma_{DU}^2 \bar{m}_{a,DU}(1+CV_{m_{a,DU}}^2) + \sigma_{person}^2\},$$

where a = 18 to 25.

Table 3.8 Combined Effects of Primary Sample Stratification and Clustering

	Age Domain			
Variable	**12-17**	**18-25**	**26+**	**12+**
Past Month Cigarette Use	1.1654	1.2063	1.1357	1.3968
Past Month Alcohol Use	1.1648	1.2123	1.1397	1.4306
Past Month Use of Any Illicit Drug	1.1016	1.1265	1.0825	1.2429
Past Month Use of Any Illicit Drug but Marijuana	1.0537	1.0668	1.0445	1.1280
Past Month Cocaine Use	1.0522	1.0624	1.0405	1.1116
Past Year, Dependent on Illicit Drugs	1.0421	1.0500	1.0311	1.0885
Past Year, Dependent on Alcohol[1]	1.0198	1.0224	1.0128	1.0362
Past Year Received Treatment for Illicit Drugs	1.0321	1.0360	1.0214	1.0569
Past Year Received Treatment for Alcohol Use[2]	1.0436	1.0503	1.0311	1.0840

[1]Dependent on Alcohol but Not Dependent on Illicit Drugs.

[2]Received Treatment for Alcohol Use but Not for Illicit Drug Use.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007.

Model 3: For all persons aged 26 or older (four age groups combined),

$$Var(\hat{p}_{26+}) = \sum_{a=4}^{7} W_a^2 \frac{p_a(1-p_a)}{n_a} UWE_a \{\sigma_{seg}^2 \bar{m}_{26+,seg} (1 + CV^2_{m_{26+,seg}}) + \sigma_{DU}^2 \bar{m}_{26+,DU} (1 + CV^2_{m_{26+,DU}}) + \sigma^2_{person}\}.$$

Model 4: For all persons aged 12 or older (seven age groups combined),

$$Var(\hat{p}_{12+}) = \sum_{a=1}^{7} W_a^2 \frac{p_a(1-p_a)}{n_a} UWE_a \{\sigma_{seg}^2 \bar{m}_{12+,seg} (1 + CV^2_{m_{12+,seg}}) + \sigma_{DU}^2 \bar{m}_{12+,DU} (1 + CV^2_{m_{12+,DU}}) + \sigma^2_{person}\}.$$

3.4 Solutions to the Age Group Optimization

Solutions to the sample optimization were achieved using Chromy's algorithm (Chromy, 1987). In summary, given specified variance bounds (V*), the algorithm simultaneously solves for $n_a, a = 1, 2, ..., 7,$ in variance models 1 through 4 above, while minimizing cost. In other words, the algorithm minimizes

$$C = \sum_{a=1}^{7} C_a n_a$$

subject to

$$(1)\ \ Var(\hat{p}) = \sum_{a=1}^{7} \frac{Var(\hat{p}_a)}{n_a} \leq V^*, \text{ and}$$

$$(2)\ \ n_a \geq 0 \text{ for } a = 1, 2, ..., 7.$$

The process is iterative until the optimum solution is reached. In general, the solution will involve one or more age-specific measures for which $Var(\hat{p}) = V^*$. This age-specific measure is called the "driving" measure.

Note that for the optimization exercise, it is assumed that the cost of interviewing a 12 year old is the same as the cost of interviewing a 65 year old. Therefore, the cost component is set to 1 for all age groups. Section 3.5 describes how the cost of each oversampling design is to be estimated.

Table 3.9 presents the results of three solutions to the age group optimization. Each of these solutions is discussed in detail in the following sections.

Table 3.9 Results of Formal Optimization

Age Domain	Current[1]	Solution 1	Solution 2	Solution 3
12 to 14	11,054	2,507	5,911	8,825
15 to 17	11,446	3,877	9,140	13,644
18 to 25	22,500	7,669	18,077	22,476
26 to 34	6,300	2,741	6,460	4,231
35 to 49	9,700	4,919	11,598	7,594
50 to 64	4,023	4,173	9,839	6,443
65 or Older	2,477	2,744	6,469	4,235
12 or Older	67,500	28,630	67,494	67,448
Large State	3,600	1,527	3,600	3,597
Small State	900	382	900	899

[1] Current sample sizes for 12 to 14 and 15 to 17 are estimated using the 2007 NSDUH ratio to persons 12 to 17. Similarly, 50 to 64 and 65 or older are estimated using the 2007 ratio to persons 50 or older.

3.4.1 Solution 1: Relaxed Sample Size Requirements

For the first solution, all sample size requirements were relaxed, including the requirement that the national sample size be 67,500 interviews. Table 3.9 shows that the SAMHSA precision requirements can be met with a national sample size of 28,630 persons (Solution 1). As shown in Table 3.10, the solution is driven by past month cigarette use and past month illicit drug except marijuana use among 12 to 17 year olds, past month alcohol use among 18 to 25 year olds, and past month alcohol use among persons 12 or older.

3.4.2 Solution 2: Fixed Overall Sample Size

For the second solution (Solution 2), the specified bounds were tightened uniformly over major age domains so that the solution approximates the current overall sample size of 67,500. Like Solution 1, Solution 2 relaxes the requirement that the sample be equally distributed among the 12 to 17, 18 to 25, and 26 or older age groups. Solution 2 distributes the sample to the 12 to 17, 18 to 25, and 26 or older age groups in proportions of 0.22, 0.27, and 0.51, respectively. Like Solution 1, Solution 2 is driven by past month cigarette use and past month illicit drug except marijuana use among 12 to 17 year olds, past month alcohol use among 18 to 25 year olds, and past month alcohol use among persons 12 or older (Table 3.11).

3.4.3 Solution 3: Equal Distribution among Three Major Age Groups

For the third and final solution (Solution 3), the specified bounds were adjusted within major age domains so that the solution was approximately equally distributed to the three major age groups (12 to 17, 18 to 25, and 26+) and the overall sample size was approximately 67,500. Like Solutions 1 and 2, Solution 3 is driven by past month cigarette use and past month illicit drug except marijuana use among 12 to 17 year olds and past month alcohol use among 18 to 25 year olds. Unlike the other solutions, Solution 3 is driven by past month alcohol use among persons 26 or older (Table 3.12), but not among persons 12 or older.

Table 3.10 Expected Relative Standard Errors, by Age Group: Solution 1 (Oversampling Optimization)

Domain	Variable	Specified		Achieved	
		SRSE	RSE	SRSE	RSE
12 to 17	**Past Month Cigarette Use**	**0.05000**	**0.0505**	**0.0500**	**0.0505**
	Past Month Alcohol Use	0.05000	0.0383	0.0496	0.0380
	Past Month Use of Any Illicit Drug	0.05000	0.0513	0.0498	0.0511
	Past Month Use of Any Illicit Drug but Marijuana	**0.05000**	**0.0752**	**0.0500**	**0.0752**
	Past Month Cocaine Use	0.05000	0.2597	0.0480	0.2494
	Past Year, Dependent on Illicit Drugs	0.05000	0.1090	0.0479	0.1044
	Past Year, Dependent on Alcohol but Not Illicit Drugs	0.05000	0.1333	0.0477	0.1270
	Past Year Received Treatment for Illicit Drugs	0.05000	0.1739	0.0477	0.1661
	Past Year Received Treatment for Alcohol Use but Not Illicit Drugs	0.05000	0.4234	0.0488	0.4135
18 to 25	Past Month Cigarette Use	0.05000	0.0221	0.0499	0.0221
	Past Month Alcohol Use	**0.05000**	**0.0133**	**0.0500**	**0.0133**
	Past Month Use of Any Illicit Drug	0.05000	0.0336	0.0482	0.0324
	Past Month Use of Any Illicit Drug but Marijuana	0.05000	0.0562	0.0469	0.0527
	Past Month Cocaine Use	0.05000	0.1249	0.0468	0.1169
	Past Year, Dependent on Illicit Drugs	0.05000	0.0695	0.0465	0.0646
	Past Year, Dependent on Alcohol but Not Illicit Drugs	0.05000	0.0677	0.0459	0.0622
	Past Year Received Treatment for Illicit Drugs	0.05000	0.1375	0.0462	0.1271
	Past Year Received Treatment for Alcohol Use but Not Illicit Drugs	0.05000	0.1671	0.0465	0.1556
26 or Older	Past Month Cigarette Use	0.05000	0.0296	0.0336	0.0199
	Past Month Alcohol Use	0.05000	0.0154	0.0338	0.0104
	Past Month Use of Any Illicit Drug	0.05000	0.0673	0.0331	0.0446
	Past Month Use of Any Illicit Drug but Marijuana	0.05000	0.0970	0.0328	0.0635
	Past Month Cocaine Use	0.05000	0.1934	0.0328	0.1269
	Past Year, Dependent on Illicit Drugs	0.05000	0.1497	0.0326	0.0976
	Past Year, Dependent on Alcohol but Not Illicit Drugs	0.05000	0.1032	0.0322	0.0665
	Past Year Received Treatment for Illicit Drugs	0.05000	0.1894	0.0324	0.1227
	Past Year Received Treatment for Alcohol Use but Not Illicit Drugs	0.05000	0.2331	0.0326	0.1518
12 or Older	Past Month Cigarette Use	0.03000	0.0177	0.0298	0.0176
	Past Month Alcohol Use	**0.03000**	**0.0098**	**0.0300**	**0.0098**
	Past Month Use of Any Illicit Drug	0.03000	0.0339	0.0260	0.0294
	Past Month Use of Any Illicit Drug but Marijuana	0.03000	0.0507	0.0254	0.0429
	Past Month Cocaine Use	0.03000	0.1088	0.0263	0.0952
	Past Year, Dependent on Illicit Drugs	0.03000	0.0720	0.0238	0.0573
	Past Year, Dependent on Alcohol but Not Illicit Drugs	0.03000	0.0583	0.0253	0.0492
	Past Year Received Treatment for Illicit Drugs	0.03000	0.1066	0.0255	0.0906
	Past Year Received Treatment for Alcohol Use but Not Illicit Drugs	0.03000	0.1363	0.0265	0.1202

RSE = relative standard error; SRSE = standardized relative standard error.

Note: Bolded lines indicate drivers.

Table 3.11 Expected Relative Standard Errors, by Age Group: Solution 2 (Oversampling Optimization)

		Specified		Achieved	
Domain	**Variable**	**SRSE**	**RSE**	**SRSE**	**RSE**
12 to 17	**Past Month Cigarette Use**	**0.0326**	**0.0329**	**0.0326**	**0.0329**
	Past Month Alcohol Use	0.0326	0.0249	0.0323	0.0247
	Past Month Use of Any Illicit Drug	0.0326	0.0334	0.0324	0.0333
	Past Month Use of Any Illicit Drug but Marijuana	**0.0326**	**0.0490**	**0.0326**	**0.0490**
	Past Month Cocaine Use	0.0326	0.1691	0.0313	0.1624
	Past Year, Dependent on Illicit Drugs	0.0326	0.0710	0.0312	0.0680
	Past Year, Dependent on Alcohol but Not Illicit Drugs	0.0326	0.0868	0.0310	0.0827
	Past Year Received Treatment for Illicit Drugs	0.0326	0.1133	0.0311	0.1082
	Past Year Received Treatment for Alcohol Use but Not Illicit Drugs	0.0326	0.2758	0.0318	0.2693
18 to 25	Past Month Cigarette Use	0.0326	0.0144	0.0325	0.0144
	Past Month Alcohol Use	**0.0326**	**0.0086**	**0.0326**	**0.0086**
	Past Month Use of Any Illicit Drug	0.0326	0.0219	0.0314	0.0211
	Past Month Use of Any Illicit Drug but Marijuana	0.0326	0.0366	0.0305	0.0344
	Past Month Cocaine Use	0.0326	0.0814	0.0305	0.0762
	Past Year, Dependent on Illicit Drugs	0.0326	0.0452	0.0303	0.0421
	Past Year, Dependent on Alcohol but Not Illicit Drugs	0.0326	0.0441	0.0299	0.0405
	Past Year Received Treatment for Illicit Drugs	0.0326	0.0895	0.0301	0.0828
	Past Year Received Treatment for Alcohol Use but Not Illicit Drugs	0.0326	0.1089	0.0303	0.1013
26 or Older	Past Month Cigarette Use	0.0326	0.0193	0.0219	0.0129
	Past Month Alcohol Use	0.0326	0.0100	0.0220	0.0068
	Past Month Use of Any Illicit Drug	0.0326	0.0438	0.0216	0.0290
	Past Month Use of Any Illicit Drug but Marijuana	0.0326	0.0632	0.0213	0.0414
	Past Month Cocaine Use	0.0326	0.1260	0.0214	0.0826
	Past Year, Dependent on Illicit Drugs	0.0326	0.0975	0.0212	0.0636
	Past Year, Dependent on Alcohol but Not Illicit Drugs	0.0326	0.0672	0.0210	0.0433
	Past Year Received Treatment for Illicit Drugs	0.0326	0.1233	0.0211	0.0799
	Past Year Received Treatment for Alcohol Use but Not Illicit Drugs	0.0326	0.1518	0.0212	0.0989
12 or Older	Past Month Cigarette Use	0.0195	0.0115	0.0194	0.0114
	Past Month Alcohol Use	**0.0195**	**0.0064**	**0.0195**	**0.0064**
	Past Month Use of Any Illicit Drug	0.0195	0.0221	0.0169	0.0191
	Past Month Use of Any Illicit Drug but Marijuana	0.0195	0.0330	0.0165	0.0280
	Past Month Cocaine Use	0.0195	0.0709	0.0171	0.0620
	Past Year, Dependent on Illicit Drugs	0.0195	0.0469	0.0155	0.0373
	Past Year, Dependent on Alcohol but Not Illicit Drugs	0.0195	0.0380	0.0165	0.0321
	Past Year Received Treatment for Illicit Drugs	0.0195	0.0694	0.0166	0.0590
	Past Year Received Treatment for Alcohol Use but Not Illicit Drugs	0.0195	0.0888	0.0172	0.0783

RSE = relative standard error; SRSE = standardized relative standard error.

Note: Bolded lines indicate drivers.

Table 3.12 Expected Relative Standard Errors, by Age Group: Solution 3 (Oversampling Optimization)

		Specified		Achieved	
Domain	**Variable**	**SRSE**	**RSE**	**SRSE**	**RSE**
12 to 17	**Past Month Cigarette Use**	**0.0267**	**0.0269**	**0.0267**	**0.0269**
	Past Month Alcohol Use	0.0267	0.0204	0.0264	0.0202
	Past Month Use of Any Illicit Drug	0.0267	0.0274	0.0265	0.0272
	Past Month Use of Any Illicit Drug but Marijuana	**0.0267**	**0.0401**	**0.0267**	**0.0401**
	Past Month Cocaine Use	0.0267	0.1384	0.0256	0.1329
	Past Year, Dependent on Illicit Drugs	0.0267	0.0581	0.0255	0.0557
	Past Year, Dependent on Alcohol but Not Illicit Drugs	0.0267	0.0710	0.0254	0.0677
	Past Year Received Treatment for Illicit Drugs	0.0267	0.0927	0.0254	0.0885
	Past Year Received Treatment for Alcohol Use but Not Illicit Drugs	0.0267	0.2257	0.0260	0.2204
18 to 25	Past Month Cigarette Use	0.0292	0.0129	0.0291	0.0129
	Past Month Alcohol Use	**0.0292**	**0.0077**	**0.0292**	**0.0077**
	Past Month Use of Any Illicit Drug	0.0292	0.0196	0.0282	0.0189
	Past Month Use of Any Illicit Drug but Marijuana	0.0292	0.0328	0.0274	0.0308
	Past Month Cocaine Use	0.0292	0.0730	0.0273	0.0683
	Past Year, Dependent on Illicit Drugs	0.0292	0.0406	0.0272	0.0378
	Past Year, Dependent on Alcohol but Not Illicit Drugs	0.0292	0.0396	0.0268	0.0363
	Past Year Received Treatment for Illicit Drugs	0.0292	0.0803	0.0270	0.0742
	Past Year Received Treatment for Alcohol Use but Not Illicit Drugs	0.0292	0.0976	0.0272	0.0909
26 or Older	Past Month Cigarette Use	0.0272	0.0161	0.0270	0.0160
	Past Month Alcohol Use	**0.0272**	**0.0083**	**0.0272**	**0.0083**
	Past Month Use of Any Illicit Drug	0.0272	0.0365	0.0267	0.0359
	Past Month Use of Any Illicit Drug but Marijuana	0.0272	0.0527	0.0264	0.0511
	Past Month Cocaine Use	0.0272	0.1051	0.0264	0.1021
	Past Year, Dependent on Illicit Drugs	0.0272	0.0813	0.0263	0.0786
	Past Year, Dependent on Alcohol but Not Illicit Drugs	0.0272	0.0561	0.0259	0.0536
	Past Year Received Treatment for Illicit Drugs	0.0272	0.1029	0.0261	0.0988
	Past Year Received Treatment for Alcohol Use but Not Illicit Drugs	0.0272	0.1267	0.0262	0.1222
12 or Older	Past Month Cigarette Use	0.0293	0.0173	0.0235	0.0138
	Past Month Alcohol Use	0.0293	0.0095	0.0237	0.0077
	Past Month Use of Any Illicit Drug	0.0293	0.0331	0.0199	0.0225
	Past Month Use of Any Illicit Drug but Marijuana	0.0293	0.0496	0.0195	0.0329
	Past Month Cocaine Use	0.0293	0.1063	0.0204	0.0740
	Past Year, Dependent on Illicit Drugs	0.0293	0.0704	0.0180	0.0432
	Past Year, Dependent on Alcohol but Not Illicit Drugs	0.0293	0.0570	0.0197	0.0383
	Past Year Received Treatment for Illicit Drugs	0.0293	0.1041	0.0198	0.0704
	Past Year Received Treatment for Alcohol Use but Not Illicit Drugs	0.0293	0.1332	0.0207	0.0941

RSE = relative standard error; SRSE = standardized relative standard error.

Note: Bolded lines indicate drivers.

3.5 Specifications for Several Alternative Sample Designs

In Sections 3.2 to 3.4, several age groups and other demographic groups for possible oversampling are investigated. In addition, three solutions were prepared for the optimal allocation of the NSDUH sample to seven age groups: 12 to 14, 15 to 17, 18 to 25, 26 to 34, 35 to 49, 50 to 64, and 65 or older. The preliminary sample optimization work described thus far only scratched the surface of possibilities. Thus, the following options were discussed with SAMHSA to determine which directions would be of greatest interest:

1. investigating other age and demographic groups of interest for oversampling, such as race and pregnancy, as mentioned earlier;

2. doubling or tripling the cluster size and investigating the effect on the sample optimization;
3. altering the specified variance bounds within major age domains so that less prevalent drug use measures drive the solution; and
4. defining sample size requirements (overall and within major age groups).

Based on these discussions and drawing from earlier results, SAMHSA provided specifications for several alternative designs to investigate. The specifications for these designs are provided in the rest of this section, and the design evaluations are provided in Sections 3.6 to 3.12.

3.5.1 Sample Size Specifications

The overall sample sizes that were studied are shown in Table 3.13. The national design reduces to essentially a proportional allocation based on the U.S. civilian, noninstitutionalized population aged 12 or older. Table 3.13a shows the proportion of the U.S. population in each State and the percentage of 12 to 17 year olds, 18 to 25 year olds, and persons aged 26 or older within each State. The hybrid design supplements the national sample in many States to achieve an effective sample size[10] of at least 400 for past month marijuana use estimates for persons aged 12 or older. Section 3.7 addresses the nominal and effective sample sizes based on some assumptions about design effects for each age allocation by cluster size combination under study.

Table 3.13 Sample Size Specifications

	Current Design	National Design	Hybrid Design
Total Sample Size	67,500	40,000	40,000 plus supplements
State Sample Size		No specific State requirement—basic proportional allocation to 12 or older population	Supplement when needed for effective sample size of 400 for past month marijuana use
Eight Largest States	3,600		
Remaining States	900		

[10] The effective sample size (n_{eff}) is the nominal sample size (n) divided by the design effect ($deff$), or $n_{eff} = n / deff$. The design effect is the ratio of the variance of an estimate obtained under the sample design to the variance of the estimate under simple random sample (SRS) of the same size. The variances of a proportion p under the competing designs are treated in Appendix A. The variance of estimate of a proportion, p, under SRS is $p(1-p)/n$.

Table 3.13a Percentage of Persons Aged 12 or Older, by State and by Age Group within State, 2007 NSDUH

State	Percentage of U.S. Population Aged 12 or Older	Percentage Aged 12 to 17 within State	Percentage Aged 18 to 25 within State	Percentage Aged 26 or Older within State
Total U.S.	100.00	10.18	13.21	76.61
AK	0.22	11.86	13.95	74.19
AL	1.54	10.10	13.06	76.84
AR	0.93	10.08	12.69	77.22
AZ	2.07	10.45	13.06	76.49
CA	12.04	10.85	14.20	74.94
CO	1.60	9.77	13.03	77.20
CT	1.18	10.10	12.15	77.74
DC	0.20	7.51	16.80	75.69
DE	0.29	9.84	12.85	77.30
FL	6.16	9.06	11.63	79.31
GA	3.08	10.80	13.03	76.16
HI	0.42	9.26	12.11	78.63
IA	1.00	9.96	13.83	76.20
ID	0.48	11.08	13.60	75.32
IL	4.25	10.34	13.56	76.10
IN	2.10	10.34	13.08	76.58
KS	0.91	10.37	14.31	75.32
KY	1.41	9.76	12.16	78.08
LA	1.41	10.59	14.60	74.81
MA	2.20	9.40	13.25	77.35
MD	1.87	10.24	12.78	76.98
ME	0.45	9.28	11.17	79.55
MI	3.38	10.53	13.00	76.47
MN	1.74	10.08	13.46	76.45
MO	1.95	10.18	12.93	76.89
MS	0.95	11.04	14.06	74.90
MT	0.32	9.84	13.19	76.97
NC	2.98	9.91	12.42	77.67
ND	0.21	9.52	17.02	73.47
NE	0.58	10.28	14.50	75.23
NH	0.45	9.94	11.91	78.15
NJ	2.92	9.99	11.84	78.17

(continued)

Table 3.13a Percentage of Persons Aged 12 or Older, by State and by Age Group within State, 2007 NSDUH (continued)

State	Percentage of U.S. Population Aged 12 or Older	Percentage Aged 12 to 17 within State	Percentage Aged 18 to 25 within State	Percentage Aged 26 or Older within State
NM	0.65	10.52	14.11	75.36
NV	0.84	10.23	11.53	78.23
NY	6.53	9.70	13.57	76.74
OH	3.84	10.16	12.75	77.10
OK	1.18	10.18	14.04	75.78
OR	1.27	9.47	12.21	78.32
PA	4.21	9.68	12.68	77.64
RI	0.36	9.49	14.12	76.39
SC	1.46	10.03	12.64	77.33
SD	0.26	10.27	14.08	75.64
TN	2.05	9.80	12.13	78.07
TX	7.63	11.14	14.27	74.59
UT	0.83	11.99	18.24	69.76
VA	2.53	9.82	12.97	77.20
VT	0.22	9.45	12.68	77.86
WA	2.17	9.92	12.49	77.60
WI	1.88	9.98	13.17	76.85
WV	0.62	8.80	11.60	79.59
WY	0.17	9.92	13.70	76.39

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007.

3.5.2 Cluster Size Options

As requested by SAMHSA, each of the three designs was evaluated with a cluster (segment) size of about 9 and about 27. This has been interpreted as the current expected cluster size of 9.375 and an expected cluster size 3 times as large (or 28.125); both of these numbers represent expected cluster sizes on average over the whole sample. This interpretation allows for reducing the number of strata by defining larger geographic strata and maintaining the number of segments per stratum in the current design.[11] A final solution to the hybrid design may require an allowance for some variation in cluster size among States to permit integer allocation of strata to States.[12]

[11] The current design assigns eight segments to each stratum. This design permits half of the segments to be replaced each year and facilitates fielding the sample as four quarterly surveys with two segments worked during each quarter. Alternative designs will maintain this feature if at all possible. Adjustments will be achieved by changing the number and size of the strata and then adjusting the cluster size to achieve the overall sample size requirement. For the hybrid option, this will need to be done within the State and may lead to some planned cluster size variation.

[12] Dr. Robert Groves suggested the possible need for variable cluster sizes during an expert consultant NSDUH meeting held in November 2008.

3.5.3 Allocation to Age Groups

In addition, the specifications included three options for sample allocation to age groups, as shown in Table 3.14. Allocation 1 represents the current design with the allocation to three groups comprising persons aged 26 or older based on the 2008, 2009, and 2010 planned sample allocations. The allocation within the 26 or older groups is based on simultaneously achieving acceptable precision levels on specified survey measures for the combined age group (e.g., see pp. A11-A15 of Morton, Chromy, & Hirsch, 2008a). The allocation to the 12 to 17 and 18 to 25 age groups is set to provide a high level of precision for reporting substance use measures for these special age groups. Allocation 2 reduces the sample for the first two age groups and allows for some reduction in precision for these two age groups; it allocates the remaining 50 percent of the sample dedicated to persons aged 26 or older among the three component age groups (26 to 34, 35 to 49, and 50 or older) in proportion to their 2007 population estimates. Allocation 3 also allows for lower precision for the first two age groups, but it allocates those aged 26 or older in the same proportions as allocation 1. The component age groups are proportionately increased to compensate in the reduced allocation to persons aged 12 to 17 and persons aged 18 to 25. These specifications led to 18 potential designs (3 sample size specifications by 2 cluster sizes by 3 age group allocations).

Table 3.14 Sample Allocation, by Age Group Options, Percentages

Age Group	Allocation 1: Current Design	Allocation 2	Allocation 3
12 to 17	33.33	25.00	25.00
18 to 25	33.33	25.00	25.00
26 to 34	9.33	9.30	14.00
35 to 49	14.37	17.08	21.56
50 or Older	9.63	23.63	14.44

3.5.4 Key Outcome Variables

For this component of the Sampling Issues Methods Study, the list of outcome variables was expanded by SAMHSA to the following:

1. past month alcohol use, 12 to 20, 50 or older, 12 or older, 12 or older for Asian or Pacific Islander (API), American Indian or Alaska Native (AIAN), pregnant women;
2. past month binge alcohol use, 18 to 25, 12 or older;
3. past month marijuana use, 12 or older, 12 to 17, 18 to 25, 50 or older, 12 or older for API, AIAN, pregnant women;
4. past month cigarette use, 12 to 17, 12 or older;
5. past month pain reliever use, 18 to 25, 12 or older;
6. past year alcohol disorder, 12 or older;
7. past year illicit drug disorder, 12 or older;
8. substance use disorder, 50 or older;

9. past year specialty substance use treatment, 12 or older;
10. serious mental illness (SMI) proxy, 18 or older; and
11. past year major depressive episode (MDE), 18 or older.

3.5.5 Evaluation Criteria

Alternative designs were evaluated in six dimensions:

1. ***Sampling frame structure***: This topic covers the number and size of the primary strata, the number of selected primary sampling units (PSUs or segments) per strata and survey year, and the target sample size for each PSU. Under each design, stratification and sampling unit definitions were specified as a preliminary step to developing appropriate cost and variance models. The evaluation of this design dimension will address possible additional flexibility in specifying the design structure and is discussed in Section 3.7.
2. ***Cost***: The variable cost models developed for studying cluster size in terms of the numbers of segments and sample persons were refined to include a sample dwelling unit (SDU) component and a PSU component and are discussed in Section 3.8.
3. ***Precision of estimates for annual estimates and trend measures across nonadjacent[13] years***: Comparative results are presented in Section 3.9 in terms of prevalence estimates and their RSEs.
4. ***Field management/structural issues***: Each design is examined in Section 3.10 to determine its impact on the size, organization, and structure of the data collection staff.
5. ***State and substate estimates***: Each design is examined in Section 3.11 to determine its impact on State and substate estimates.
6. ***Response rates***: The impact on overall expected response rates (for those aged 12 or older) is examined in Section 3.12 for each design.

3.6 Effective Sample Sizes for Hybrid Designs

The hybrid sample design specifications called for achieving a minimum effective sample size of 400 for State estimates of past month marijuana use by persons aged 12 or older. Table 3.15 shows 2007 NSDUH estimates of the design effects and translates them into required State sample sizes. The design effect for the estimate of a proportion is the variance of the estimate divided by what the variance would have been had the proportion been estimated using SRS. The effective sample size of an estimate is its nominal sample size divided by its design effect.

[13] Estimates from adjacent years are subject to a small positive correlation under the current design because half of the segments are reused each year. New dwelling unit selections are employed each year, even in the reused segments. Confining comparisons to nonadjacent years makes it reasonable to assume that estimates for the 2 years are uncorrelated, and it simplifies the modeling of the SEs of trend measures. For adjacent years, the SEs of differences actually would be reduced somewhat because a positive correlation reduces the variance of difference estimates.

The design effects calculations in Table 3.15 assume that whenever a State estimate is required, State sampling (SS) regions (similar to the current ones) will be treated as strata rather than as PSUs, per the national design strategy.

Table 3.15 Estimated Design Effects and Minimum Nominal Sample Sizes to Achieve an Effective Sample Size of 400 for Past Month Marijuana Use

	Design Effect, by Expected Cluster Size[1]		Required Nominal Sample Size, by Expected Cluster Size (= 400 × design effect)	
Age Group Allocation	**Cluster Size = 9.375**	**Cluster Size = 28.125**	**Cluster Size = 9.375**	**Cluster Size = 28.125**
1	1.861	2.280	745	913
2	1.535	1.880	614	753
3	1.415	1.733	566	694

[1] See Equation A4 in Appendix A.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007.

Based on these requirements, State sample sizes and the total national sample are shown in Table 3.16. The national design increases the sample allocation in one of the largest eight States (California) and in four of the remaining States (Georgia, New Jersey, North Carolina, and Virginia); the sample size in all other States is reduced compared with the current design. The hybrid design increases the sample size in the smaller States relative to the national design.

3.7 Sampling Plans

3.7.1 Descriptions

Eighteen designs were evaluated in this NSDUH study. For the national design, an extra stage of sampling was introduced. PSUs are normally larger areas composed of metropolitan areas or sets of counties in such a design. They also could be arbitrarily defined geographic areas, such as the currently defined SS regions (e.g., see p. 8 of Morton et al., 2008b). The variance component analysis conducted on the 2007 data identified the SS region as a stage of sampling and provided estimates of variance components associated with the SS regions. Consequently, the current SS regions are used as the PSUs in modeling the variances and costs for the national designs evaluated.

For all of the designs considered, SS regions continued to be defined in terms of contiguous areas formed by combining whole census tracts (see Morton et al., 2008b) to form approximately equal-sized geographic regions within each State. In addition to contiguity, the current methodology requires that regions be defined to facilitate travel relative to major rivers and mountain ranges.

Table 3.16 Expected State Sample Sizes, by Design Option and Age Allocation

State	Rank	Current Design	National	Hybrid Design Options to Achieve Effective Sample Size = 400					
				Cluster Size = 9.375			Cluster Size = 28.125		
				Allocation 1	Allocation 2	Allocation 3	Allocation 1	Allocation 2	Allocation 3
CA	1	3,600	4,817	4,817	4,817	4,817	4,817	4,817	4,817
TX	2	3,600	3,051	3,051	3,051	3,051	3,051	3,051	3,051
NY	3	3,600	2,613	2,613	2,613	2,613	2,613	2,613	2,613
FL	4	3,600	2,464	2,464	2,464	2,464	2,464	2,464	2,464
IL	5	3,600	1,702	1,702	1,702	1,702	1,702	1,702	1,702
PA	6	3,600	1,684	1,684	1,684	1,684	1,684	1,684	1,684
OH	7	3,600	1,535	1,535	1,535	1,535	1,535	1,535	1,535
MI	8	3,600	1,352	1,352	1,352	1,352	1,352	1,352	1,352
GA	9	900	1,233	1,233	1,233	1,233	1,233	1,233	1,233
NC	10	900	1,191	1,191	1,191	1,191	1,191	1,191	1,191
NJ	11	900	1,167	1,167	1,167	1,167	1,167	1,167	1,167
VA	12	900	1,014	1,014	1,014	1,014	1,014	1,014	1,014
MA	13	900	878	878	878	878	913	878	878
WA	14	900	867	867	867	867	913	867	867
IN	15	900	839	839	839	839	913	839	839
AZ	16	900	826	826	826	826	913	826	826
TN	17	900	820	820	820	820	913	820	820
MO	18	900	781	781	781	781	913	781	781
WI	19	900	753	753	753	753	913	753	753
MD	20	900	749	749	749	749	913	753	749
MN	21	900	695	745	695	695	913	753	695
CO	22	900	642	745	642	642	913	753	694
AL	23	900	615	745	615	615	913	753	694
SC	24	900	582	745	614	582	913	753	694
KY	25	900	564	745	614	566	913	753	694
LA	26	900	562	745	614	566	913	753	694
OR	27	900	507	745	614	566	913	753	694
OK	28	900	472	745	614	566	913	753	694
CT	29	900	471	745	614	566	913	753	694
IA	30	900	399	745	614	566	913	753	694
MS	31	900	378	745	614	566	913	753	694
AR	32	900	374	745	614	566	913	753	694
KS	33	900	364	745	614	566	913	753	694

(continued)

Table 3.16 Expected State Sample Sizes, by Design Option and Age Allocation (continued)

State	Rank	Current Design	National	Hybrid Design Options to Achieve Effective Sample Size = 400					
				Cluster Size = 9.375			Cluster Size = 28.125		
				Allocation 1	Allocation 2	Allocation 3	Allocation 1	Allocation 2	Allocation 3
NV	34	900	337	745	614	566	913	753	694
UT	35	900	331	745	614	566	913	753	694
NM	36	900	259	745	614	566	913	753	694
WV	37	900	248	745	614	566	913	753	694
NE	38	900	233	745	614	566	913	753	694
ID	39	900	194	745	614	566	913	753	694
ME	40	900	182	745	614	566	913	753	694
NH	41	900	180	745	614	566	913	753	694
HI	42	900	170	745	614	566	913	753	694
RI	43	900	144	745	614	566	913	753	694
MT	44	900	129	745	614	566	913	753	694
DE	45	900	115	745	614	566	913	753	694
SD	46	900	105	745	614	566	913	753	694
AK	47	900	87	745	614	566	913	753	694
VT	48	900	86	745	614	566	913	753	694
ND	49	900	86	745	614	566	913	753	694
DC	50	900	81	745	614	566	913	753	694
WY	51	900	70	745	614	566	913	753	694
Total		67,500	40,000	53,432	49,481	48,153	59,431	53,685	51,852

Within SS regions, census tracts are the PSUs[14] in the current 50-State design. Census tracts are selected with probability proportional to size (PPS) according to a minimum replacement scheme. For each tract selection, area segments consisting of one or more census blocks are defined, and one segment is selected for the sample. This method is employed to ensure that segments are wholly contained within census tracts. For discussion purposes, the combined selection of a census tract and a segment within the census tract is referred to as segment selection, and the ultimate segments are referred to as PSUs for the variance component analysis and for sample allocation purposes.

As noted above, for the national design options, the SS regions are treated as PSUs. The segments in the 50-State design become secondary sampling units (SSUs).

The quarterly design (with eight segments per SS region) also was maintained for all of the designs, with the additional assumptions of the current dwelling unit listing methods and

[14] Adjacent census tracts were combined if a single tract did not contain a specified minimum number of dwelling units. In all cases, the PSU consisted of one or more census tracts.

associated costs. If some form of address-based sampling (ABS) is adopted in the future, cost and variance models will need to be adjusted to reflect this change, and optimum cluster sizes may change. Although segments probably will be larger under ABS sampling, which may result in a slight reduction in variance because of a marginally smaller intracluster correlation, the reduction in variance will be roughly offset by the use of the check for housing units missed (CHUM) procedure. The larger impact of ABS will be on reducing the cost per segment and thereby changing the cost/variance tradeoff currently favoring an increase in the number of respondents per segment from 9 to 29. Such a large increase will no longer be optimal.

Table 3.17 shows the design parameters for six sample designs using the current national sample size of 67,500 and assumes State samples as currently allocated. For comparing the impact of increasing cluster size, an average cluster size of 28.125 was used to reflect increasing clusters by a factor of three. This has the corresponding effect of decreasing the number of SS regions and segments to one third of their 2009 survey values. Under the current age allocation scheme, the sample requirement for 12 to 17 year olds drives the screening requirement in most States with 18 to 25 year olds sampled at a slightly lower rate on average. The next highest sampling rate is for 26 to 34 year olds and is about one fifth of the sampling rate of the two younger age groups. For modeling purposes, the number of completed screeners was derived as a constant ratio to the 12- to 17-year-old sampling size requirement. Implicit in this ratio is the expected interview response rates. Table 3.17a provides weighted interview response rates by five age groups and all persons 12 or older from the 2007 NSDUH. The ratio to 12- to 17-year-old sample size approach also was used to estimate the required number of completed screening interviews for the national and hybrid designs.

Table 3.18 shows the design parameters for six sample designs based on a national sample of 40,000 persons.[15] For modeling purposes, the current SS regions were treated as PSUs. The number of segments required was rounded to the nearest multiple of 8 to preserve the quarterly design with two segments observed each quarter; a replacement of half of the segments each year also continues to be assumed for the national design. When considering large cluster sizes, larger geographic strata could have been used, but maintaining the 900 strata and using larger minimum segment sizes also appears feasible and was assumed for three designs with larger cluster sizes.[16] It was further assumed that the sample PSUs will be selected in advance for a typical 5-year design plan. A sufficient number of segments will be selected within each PSU to replace half of them with each new sample year.

[15] A formal multiple constraint optimization was conducted that set the variance bounds at SAMHSA's current precision requirements. That is, the expected RSE on a prevalence of 10 percent could not exceed 3 percent for the total population estimates and 5 percent for statistics in three age group domains (i.e., 12 to 17, 18 to 25, and 26 or older). The solution suggested that the optimal number of persons was approximately 40,000 and the optimal cluster size was 29 (see Chapter 2). Thus, a national sample of 40,000 persons was evaluated in this chapter.

[16] The major disadvantage to decreasing the number of strata is that it limits the ability to produce small area estimates as discussed in Section 3.11.3. Otherwise, decreasing the number of strata is favored or neutral.

Table 3.17 Summary of Design Parameters under Various 50-State Designs with the Current Sample Size

Design Parameter	Cluster Size = 9.375			Cluster Size = 28.125		
	Allocation 1	Allocation 2	Allocation 3	Allocation 1	Allocation 2	Allocation 3
Age Group Sample Size						
12 to 17	22,500	16,875	16,875	22,500	16,875	16,875
18 to 25	22,500	16,875	16,875	22,500	16,875	16,875
26 to 34	6,300	6,275	9,450	6,300	6,275	9,450
35 to 49	9,700	11,526	14,550	9,700	11,526	14,550
50 or Older	6,500	15,949	9,750	6,500	15,949	9,750
Total	67,500	67,500	67,500	67,500	67,500	67,500
Total Sample						
SS Regions (Frame)	900	900	900	300	300	300
SS Regions (Sample)	900	900	900	300	300	300
Segments	7,200	7,200	7,200	2,400	2,400	2,400
Expected Completed Screening Interviews	140,041	105,031	105,031	140,041	105,031	105,031
Expected Completed Interviews	67,500	67,500	67,500	67,500	67,500	67,500

SS = State sampling.

Table 3.17a Expected Interview Response Rates, by Age Group

Age Group	Expected Interview Response Rate (Percent)
12 or Older	73.94
12 to 17	85.35
18 to 25	79.76
26 to 34	75.36
35 to 49	73.37
50 or Older	68.49

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007.

Table 3.18 Summary of Design Parameters for National Sample of 40,000

Design Parameter	Cluster Size = 9.375			Cluster Size = 28.125		
	Allocation 1	Allocation 2	Allocation 3	Allocation 1	Allocation 2	Allocation 3
Age Group Sample Size						
12 to 17	13,333	10,000	10,000	13,333	10,000	10,000
18 to 25	13,333	10,000	10,000	13,333	10,000	10,000
26 to 34	3,733	3,718	5,600	3,733	3,718	5,600
35 to 49	5,748	6,830	8,622	5,748	6,830	8,622
50 or Older	3,852	9,451	5,778	3,852	9,451	5,778
Total	40,000	40,000	40,000	40,000	40,000	40,000
Total Sample						
SS Regions (Frame)	900	900	900	900	900	900
SS Regions (Sample)	533	533	533	178	178	178
Segments (Rounded to Multiple of 8)	4,264	4,264	4,264	1,424	1,424	1,424
Expected Completed Screening Interviews	82,987	62,240	62,240	82,987	62,240	62,240
Expected Completed Interviews	40,000	40,000	40,000	40,000	40,000	40,000

SS = State sampling.

Table 3.19 shows design parameters for six hybrid designs based on supplementing the national sample to achieve an effective sample size of 400 in each State when estimating past month marijuana use for persons 12 or older. Because State samples are required under all these options, it was assumed that SS regions would be defined within each State, and segments would serve as first-stage sampling units; this approach is similar to current design with fewer SS regions formed in most States. Although cost and variance were modeled based on national average cluster sizes, the preservation of SS regions with eight segments in each one would require that target cluster sizes vary considerably among States. A fully developed sampling plan for each State would need to be developed before implementing any of the hybrid designs.[17]

3.7.2 Impact on Sampling Frame Construction and Sample Frame Structure

Changing the allocation to age groups affects the sampling scheme within the dwelling units through an adjustment of the target selection probabilities. It also may change the number of SDUs required to yield the respondent sample as shown in Tables 3.17, 3.18, and 3.19. The following discussion considers the six designs defined by three sample size specifications and the two cluster sizes.

[17] Developing a sampling plan for each State is a one-time cost and has not been taken into consideration when costing the hybrid models.

Table 3.19 **Summary of Design Parameters for Hybrid Designs**

Design Parameter	Cluster Size = 9.375			Cluster Size = 28.125		
	Allocation 1	Allocation 2	Allocation 3	Allocation 1	Allocation 2	Allocation 3
Age Group Sample Size						
12 to 17	17,811	12,370	12,038	19,810	13,421	12,963
18 to 25	17,811	12,370	12,038	19,810	13,421	12,963
26 to 34	4,987	4,600	6,741	5,547	4,990	7,259
35 to 49	7,678	8,449	10,380	8,540	9,167	11,177
50 or Older	5,145	11,692	6,955	5,723	12,685	7,490
Total	53,432	49,481	48,153	59,431	53,685	51,852
Total Sample						
SS Regions (Frame)	712	660	642	264	239	230
SS Regions (Sample)	712	660	642	264	239	230
Segments (Rounded to Multiple of 8)	5,696	5,280	5,136	2,112	1,912	1,840
Expected Completed Screening Interviews	110,854	76,993	74,927	123,300	83,534	80,682
Expected Completed Interviews	53,432	49,481	48,153	59,431	53,685	51,852

SS = State sampling.

Current 50-State design with a sample size of 67,500 total respondents and an average cluster size of 9.375 respondents. This design has a number of national-level features:

- This design has 900 primary strata, with an annual sample of 7,200 segments (as shown in Table 3.17).
- The primary strata, called SS regions, may vary in population size across States, but they are designed to be of approximately equal population within a State.
- Large sample States have 48 SS regions, while small sample States have 12 SS regions.
- Within each SS region, two sample segments are selected for each quarterly survey for a total of eight segments per year. Half of the segments (or four per SS region) are retired each year and replaced with a new sample the following year.
- A coordinated multiyear sample can be selected to provide a PPS sample. A 10-year super sample draw allows spanning the period between publications of decennial

censuses; each annual sample then is an equal probability sample from the super sample.

- The minimum number of dwelling units is set at 100 in rural areas and 150 in urban areas to yield adequate numbers of selected respondents for at least two annual surveys. The sample of segments is selected by first drawing a census tract or set of census tracts that meet the minimum size requirement. Sampling of the compact area segment within each census tract is postponed until the year the census tract first enters the sample.

A 50-State design with a sample of 67,500 total respondents and an average cluster of 28.125 respondents. The following features change:

- This design has 300 primary strata with an annual sample of 2,400 segments (Table 3.17).

- SS regions will be 3 times as large as in the current design. A simple way of forming them will be to combine three adjacent current SS regions. Alternatively, if the larger SS regions were formed first, each could be subsequently partitioned into three SS regions to provide flexibility in returning to the current design. Another approach would be to start with the current 900 SS regions and draw the sample of 7,200 segments. Then, if desired, the larger SS regions could be formed by combining adjacent SS regions, and the sample of segments could be selected as a subsample of the original sample in each new SS region.

- Within each SS region, the structure will be same except for larger units.

- Multiyear sample draws will continue to be feasible.

- Large sample States will have 16 SS regions, while small sample States will have 4.

- The minimum number of dwelling units per segment will have to be increased by about a factor of 3 to provide for an adequate yield of respondents. It is currently rare for census tracts to fail to meet the minimum size requirements for segments, but it does occur occasionally for industrial and special tracts. Increasing the minimum size by a factor of 3 will not be expected to materially increase the number of census tracts that need to be combined to form PSUs.

National sample of 40,000 with an average cluster size close to 9.375. This option calls for a number of major changes in the structure of the sampling frame and the sample, as shown in Table 3.18:

- The numbers in Table 3.18 provide a solution with an average cluster size near 9.375. A total of 533 SS regions with an average cluster size of 9.381 is shown in Table 3.18. To preserve the within SS region sample structure with 8 segments selected in each SS region, the total number of segments is set as 8 times the number of SS regions, or 4,264.

- The SS regions in Table 3.18 are no longer sample strata, but are instead treated as PSUs, which introduces an additional stage in the sampling frame and sample structure. This approach facilitated computation of variance components and the relation of cost to the current 50-State design. It also provides the flexibility to return to a 50-State design in the future by including all 900 SS regions in the sample instead of just the 533 shown in Table 3.18.

- A combination of explicit and implicit stratification will be applied to the SS regions, and the PPS sequential sampling scheme will be applied to select the sample in each explicit stratum (the sampling scheme now used to select sample census tracts). Size measures will be based on an estimated population, so the sample distribution will shift toward highly populated areas. Optionally, a super sample of SS regions could be selected to allow some replacement of the SS region sample for each annual survey, but the cost models in this report assume sample replacement occurring only with new segments within selected SS regions.

- Within selected SS regions, the sample allocation to quarters and the replacement schemes will remain as above for the current 50-State design.

- A multiyear sample could be selected with some combination of replacement schemes at the SS region and segment within SS region possible.

- There would be no change from the current design in segment formation except that it would be restricted to SS regions selected into the sample each year.

National sample of 40,000 with an average cluster size near 28.125. This sample can best be described as modification to the national sample of 40,000 with an average cluster size near 9.375:

- Table 3.18 shows that this design can be achieved by selecting 178 SS regions and 1,424 segments (8 per SS region annually) to achieve an average cluster size of 28.090.

- As noted above, the flexibility exists to switch back to the 50-State design by treating all SS regions as strata (selected with probability 1) and increasing the sample size to 67,500. If the approximate solution were adjusted to 534 SS regions for the option above, then flexibility to move between the 178 SS regions and 534 SS regions with a sample size of 40,000 will be achieved in a manner similar to that applied to the 50-State designs.

Hybrid designs with a cluster size of about 9.375 or about 28.125. These designs must be tailored to the requirements of each State:

- Table 3.19 shows an approximate solution to the specified sampling requirements developed based on the expected results but without working out the details in each State for each of the six sample size by cluster size combinations.

- Even within State, the nominal sample size requirements vary substantially across the six options. The majority of States will receive the smallest sample allocations shown in Table 3.18, with the range of nominal sample sizes between 566 for allocation 3 with a cluster size near 9.375 and 913 for allocation 1 with a cluster size near 28.125.

- To preserve the sample structure within SS regions for sample replacement and quarterly coverage, it is easiest to think in terms of blocks of a targeted size of 75 (8 × 9.375) for the smaller cluster size and a targeted size of 225 (8 × 28.125) for the larger cluster size specified. Forming four SS regions for the targeted cluster size of 28.125 will result in 32 segments per annual sample with an average size of 28.531 to achieve a State sample size of 913 persons under age allocation scheme 1.

- Increasing the number of SS regions by a factor of 3 to 12 for designs with a target cluster of size of 9.375 will results in 96 segments. With a reduced target of 633 under allocation 3, the average cluster size will need to be reduced to 6.594.

- An attractive alternative in this case will be to form 8 SS regions with 8 segments per SS region for a total of 64 segments with an average cluster size of 8.844 (566/64).

The enumeration of more options is feasible. For planning purposes, it may be sufficient to note that solutions exist that preserve the design structure within SS regions for each of the six cluster size by age allocation plans. The within SS region structure also could be relaxed in a number of ways. For example, the segment sample allocated to quarters could be allowed to vary by at most one segment within States while maintaining equal quarterly allocations nationally or regionally. The allocation of segments to quarters also could be applied after drawing a sample of segments so that a smaller number of segments will be selected per stratum (SS region). The number of SS regions then could be increased to ensure geographic spread and representation within State at the expense of equal quarterly representation within each stratum.

3.8 Cost Comparisons

The cost figures[18] included in this report are based on 2006 calendar year cost data. Using best knowledge and logical assumptions, task-level costs were divided into (a) variable costs per geographic cluster, (b) variable costs per interview, and (c) fixed costs. Only variable costs were included in the analysis. This included a portion of the costs from the instrumentation development and training and field materials (Task 2), sample design (Task 3), sample selection (Task 4), and data management and processing (Task 7), as well as the majority of costs from the field preparations (Task 5) and screening and interviewing (Task 6).[19]

---

[18] The characteristics of census tracts played no direct role in determining these cost estimates. However, because the first stage of selection is census tracts and sample segments are contained within a single census tract to the extent possible, sample segments associated with densely populated census tracts will in theory have lower per screening and interview unit costs than more sparsely populated census tracts because of less travel time between travel points.

[19] It was assumed that the fixed costs and the per unit variable field costs will be essentially the same as under the current design with some allowance for salary increases and inflation. No adjustment was included for changes in the questionnaire structure, weighting, imputation, and other methods.

Key design parameters and total costs for 18 designs are compared in Tables 3.20, 3.21, and 3.22. Cost savings shown in the tables are relative to the current design with an average cluster size of 9.375 and allocation 1.

Table 3.20 Cost Comparisons for Various 50-State Designs with the Current Sample Size

	Cluster Size = 9.375			Cluster Size = 28.125		
	Allocation 1	**Allocation 2**	**Allocation 3**	**Allocation 1**	**Allocation 2**	**Allocation 3**
Design Parameter						
PSUs (SS Regions)	900	900	900	300	300	300
Segments	7,200	7,200	7,200	2,400	2,400	2,400
Completed Screenings	140,041	105,031	105,031	140,041	105,031	105,031
Completed Interviews	67,500	67,500	67,500	67,500	67,500	67,500
Variable Costs ($ in Thousands)	30,840	29,265	28,972	24,786	23,211	22,979
Cost Savings ($ in Thousands)	0	1,575	1,868	6,054	7,629	7,862

PSU = primary sampling unit; SS = State sampling.

Table 3.20 shows that under the current sample size requirement, the greatest savings will occur by increasing the cluster size and lowering the sampling rates for the younger age groups. This within-table conclusion carries over to the national and hybrid designs.

Table 3.21 Cost Comparisons for National Design of 40,000

	Cluster Size = 9.375			Cluster Size = 28.125		
	Allocation 1	**Allocation 2**	**Allocation 3**	**Allocation 1**	**Allocation 2**	**Allocation 3**
Design Parameter						
PSUs (SS Regions)	533	533	533	178	178	178
Segments	4,264	4,264	4,264	1,424	1,424	1,424
Completed Screenings	82,987	62,240	62,240	82,987	62,240	62,240
Completed Interviews	40,000	40,000	40,000	40,000	40,000	40,000
Variable Costs ($ in Thousands)	18,272	17,339	17,165	14,690	13,757	13,619
Cost Savings ($ in Thousands)	12,568	13,502	13,675	16,150	17,084	17,221

PSU = primary sampling unit; SS = State sampling.

The national design (Table 3.21) with a sample of 40,000 offers the greatest cost reductions. Hybrid designs still produce major cost reductions while preserving minimal sample sizes for some State estimates (Table 3.22). Cost savings should be weighed against loss of precision, as discussed in the next section.

Table 3.22 Cost Comparisons for Hybrid Designs

	Cluster Size = 9.375			Cluster Size = 28.125		
	Allocation 1	Allocation 2	Allocation 3	Allocation 1	Allocation 2	Allocation 3
Design Parameter						
PSUs (SS Regions)	712	660	642	264	239	230
Segments	5,696	5,280	5,136	2,112	1,912	1,840
Completed Screenings	110,854	76,993	74,927	123,300	83,534	80,682
Completed Interviews	53,432	49,481	48,153	59,431	53,685	51,852
Variable Costs ($ in Thousands)	24,408	21,455	20,668	21,822	18,464	17,647
Cost Savings ($ in Thousands)	6,432	9,385	10,173	9,018	12,376	13,193

PSU = primary sampling unit; SS = State sampling.

3.9 Precision of Estimates

Tables 3.23, 3.24, and 3.25 compare the RSEs for the 11 specified variables. Appendix A describes the model used to project the SEs.

To help understand and summarize the data in Tables 3.23 to 3.25 in a comparative way, a measure of relative precision was calculated for each variable under each design plan; the current design was used as the base plan. Cochran (1963, p. 102) defined the relative precision of design 2 to design 1 and the ratio of the variance under design 1 to the variance under design 2 expressed as a percentage. As an example, a relative precision of 133 percent can also be interpreted as a reduction in variance of about 25 percent. The relative precision of the plans is most highly related to sample size, but the impact also varies by age domain. To allow some summarization among similar domains, four domain categories were used to classify the specified outcome variables:

- those pertaining to all persons 12 or older (12 outcome variables);
- those pertaining to young populations with ranges under 25 (six outcome variables);
- those pertaining to persons aged 50 or older (three outcome variables); and
- those pertaining to intermediate or mixed groups, persons aged 18 or older, or pregnant women (four outcome variables).

Table 3.26 summarizes the relative precision of several plans using the geometric means of the relative precision estimates over each domain category.

Table 3.23 Estimated Prevalence Estimates and Relative Standard Errors for the 50-State Design

Measure	Domain	Prevalence	Relative Standard Error (RSE) per Allocation					
			Cluster Size = 9.375			Cluster Size = 28.125		
			1	2	3	1	2	3
Past Year Alcohol Use Disorder	12+	0.0752	0.0231	0.0199	0.0197	0.0246	0.0212	0.0209
Substance Use Disorder	50+	0.0371	0.0862	0.0550	0.0704	0.102	0.0651	0.0833
Past Year Illicit Drug Use Disorder	12+	0.0276	0.0330	0.0305	0.0292	0.0343	0.0317	0.0303
Past Month Alcohol Use	12-20	0.2789	0.0130	0.0150	0.0150	0.0149	0.0172	0.0172
Past Month Alcohol Use	50+	0.4692	0.0175	0.0112	0.0143	0.0186	0.0118	0.0152
Past Month Alcohol Use	AIAN, 12+	0.4181	0.0471	0.0362	0.0392	0.0514	0.0395	0.0428
Past Month Alcohol Use	12+	0.5114	0.0084	0.0065	0.0070	0.0103	0.0079	0.0086
Past Month Alcohol Use	API,12+	0.3670	0.0484	0.0396	0.0402	0.0532	0.0436	0.0442
Past Month Alcohol Use	Pregnant, 12+	0.1197	0.1321	0.1327	0.1159	0.1369	0.1376	0.1201
Past Month Pain Reliever Use	18-25	0.0464	0.0410	0.0474	0.0474	0.0426	0.0492	0.0492
Past Month Pain Reliever Use	12+	0.0209	0.0429	0.0374	0.0369	0.0459	0.0400	0.0395
Past Month Binge Alcohol Use	18-25	0.4178	0.0112	0.0129	0.0129	0.0122	0.0141	0.0141
Past Month Binge Alcohol Use	12+	0.2331	0.0136	0.0112	0.0114	0.0157	0.0129	0.0131
Past Month Cigarette Use	12-17	0.0983	0.0287	0.0332	0.0332	0.0315	0.0363	0.0363
Past Month Cigarette Use	12+	0.2424	0.0145	0.0116	0.0121	0.0172	0.0138	0.0143
Past Year MDE (18+)	18+	0.0746	0.0283	0.0226	0.0235	0.0301	0.0240	0.0250
Past Month Marijuana Use	12-17	0.0668	0.0343	0.0397	0.0397	0.0361	0.0417	0.0417
Past Month Marijuana Use	18-25	0.1643	0.0210	0.0243	0.0243	0.0224	0.0258	0.0258
Past Month Marijuana Use	50+	0.0154	0.1304	0.0832	0.1065	0.1336	0.0853	0.1091
Past Month Marijuana Use	AIAN, 12+	0.0729	0.0956	0.0914	0.0838	0.0992	0.0949	0.0869
Past Month Marijuana Use	12+	0.0583	0.0247	0.0223	0.0215	0.0273	0.0247	0.0238
Past Month Marijuana Use	API,12+	0.0279	0.1516	0.1405	0.1369	0.1579	0.1463	0.1425
Past Month Marijuana Use	Pregnant, 12+	0.0404	0.2004	0.2113	0.1886	0.2032	0.2143	0.1912
SMI Proxy (18+)	18+	0.1092	0.0218	0.0178	0.0182	0.0233	0.0190	0.0194
Past Year Specialty Substance Use Treatment	12+	0.0097	0.0710	0.0599	0.0596	0.0737	0.0622	0.0619

AIAN = American Indian or Alaska Native; API = Asian or Other Pacific Islander; MDE = major depressive episode; SMI = serious mental illness.

Note: Prevalence is expressed in absolute terms (e.g., 0.0752 = 7.52 percent).

Table 3.24 Estimated Prevalence Estimates and Relative Standard Errors for National Design

Measure	Domain	Preva-lence	Relative Standard Error (RSE) per Allocation					
			Cluster Size = 9.375			Cluster Size = 28.125		
			1	2	3	1	2	3
Past Year Alcohol Use Disorder	12+	0.0752	0.0302	0.0261	0.0257	0.0387	0.0335	0.0330
Substance Use Disorder	50+	0.0371	0.1044	0.0666	0.0852	0.1355	0.0865	0.1106
Past Year Illicit Drug Use Disorder	12+	0.0276	0.0402	0.0373	0.0356	0.0475	0.0440	0.0420
Past Month Alcohol Use	12-20	0.2789	0.0183	0.0211	0.0211	0.0255	0.0294	0.0294
Past Month Alcohol Use	50+	0.4692	0.0214	0.0137	0.0175	0.0261	0.0167	0.0213
Past Month Alcohol Use	AIAN, 12+	0.4181	0.1001	0.0811	0.0838	0.1297	0.1051	0.1086
Past Month Alcohol Use	12+	0.5114	0.0130	0.0100	0.0108	0.0194	0.0150	0.0161
Past Month Alcohol Use	API,12+	0.3670	0.0571	0.0470	0.0476	0.0749	0.0616	0.0624
Past Month Alcohol Use	Pregnant, 12+	0.1197	0.1624	0.1648	0.1439	0.1880	0.1887	0.1648
Past Month Pain Reliever Use	18-25	0.0464	0.0472	0.0545	0.0545	0.0517	0.0597	0.0597
Past Month Pain Reliever Use	12+	0.0209	0.0506	0.0442	0.0435	0.0587	0.0513	0.0505
Past Month Binge Alcohol Use	18-25	0.4178	0.0145	0.0168	0.0168	0.0187	0.0216	0.0216
Past Month Binge Alcohol Use	12+	0.2331	0.0192	0.0159	0.0161	0.0269	0.0222	0.0225
Past Month Cigarette Use	12-17	0.0983	0.0363	0.0419	0.0419	0.0461	0.0532	0.0532
Past Month Cigarette Use	12+	0.2424	0.0203	0.0164	0.0169	0.0287	0.0232	0.0239
Past Year MDE	18+	0.0746	0.0314	0.0251	0.0261	0.0333	0.0267	0.0277
Past Month Marijuana Use	12-17	0.0668	0.0424	0.0489	0.0489	0.0513	0.0592	0.0592
Past Month Marijuana Use	18-25	0.1643	0.0266	0.0308	0.0308	0.0332	0.0383	0.0383
Past Month Marijuana Use	50+	0.0154	0.1510	0.0964	0.1234	0.1686	0.1076	0.1377
Past Month Marijuana Use	AIAN, 12+	0.0729	0.2143	0.2078	0.1882	0.2502	0.2426	0.2197
Past Month Marijuana Use	12+	0.0583	0.0338	0.0306	0.0294	0.0456	0.0414	0.0398
Past Month Marijuana Use	API,12+	0.0279	0.1680	0.1573	0.1532	0.1981	0.1855	0.1807
Past Month Marijuana Use	Pregnant, 12+	0.0404	0.2388	0.2518	0.2247	0.2557	0.2695	0.2405
SMI Proxy (18+)	18+	0.1092	0.0242	0.0198	0.0202	0.0258	0.0211	0.0216
Past Year Specialty Substance Use Treatment	12+	0.0097	0.0789	0.0668	0.0663	0.0819	0.0693	0.0688

AIAN = American Indian or Alaska Native; API = Asian or Other Pacific Islander; MDE = major depressive episode; SMI = serious mental illness.

Note: Prevalence is expressed in absolute terms (e.g., 0.0752 = 7.52 percent).

Table 3.25 Estimated Prevalence Estimates and Relative Standard Errors for Hybrid Designs

Measure	Domain	Preva-lence	Relative Standard Error (RSE) per Allocation					
			Cluster Size = 9.375			Cluster Size = 28.125		
			1	2	3	1	2	3
Past Year Alcohol Use Disorder	12+	0.0752	0.0280	0.0247	0.0244	0.0330	0.0301	0.0304
Substance Use Disorder	50+	0.0371	0.0977	0.0635	0.0816	0.1194	0.0793	0.1034
Past Year Illicit Drug Use Disorder	12+	0.0276	0.0376	0.0355	0.0340	0.0415	0.0401	0.0391
Past Month Alcohol Use	12-20	0.2789	0.0168	0.0198	0.0199	0.0215	0.0262	0.0270
Past Month Alcohol Use	50+	0.4692	0.0200	0.0130	0.0167	0.0227	0.0152	0.0198
Past Month Alcohol Use	AIAN, 12+	0.4181	0.0895	0.0744	0.0772	0.1047	0.0909	0.0976
Past Month Alcohol Use	12+	0.5114	0.0118	0.0094	0.0101	0.0161	0.0132	0.0147
Past Month Alcohol Use	API,12+	0.3670	0.0536	0.0449	0.0456	0.0652	0.0559	0.0580
Past Month Alcohol Use	Pregnant, 12+	0.1197	0.1532	0.1567	0.1373	0.1645	0.1721	0.1535
Past Month Pain Reliever Use	18-25	0.0464	0.0445	0.0523	0.0525	0.0465	0.0555	0.0563
Past Month Pain Reliever Use	12+	0.0209	0.0475	0.0423	0.0418	0.0521	0.0473	0.0474
Past Month Binge Alcohol Use	18-25	0.4178	0.0135	0.0159	0.0159	0.0160	0.0195	0.0200
Past Month Binge Alcohol Use	12+	0.2331	0.0177	0.0149	0.0152	0.0226	0.0198	0.0206
Past Month Cigarette Use	12-17	0.0983	0.0338	0.0398	0.0399	0.0398	0.0481	0.0493
Past Month Cigarette Use	12+	0.2424	0.0187	0.0154	0.0160	0.0242	0.0206	0.0219
Past Year MDE	18+	0.0746	0.0298	0.0242	0.0252	0.0308	0.0252	0.0264
Past Month Marijuana Use	12-17	0.0668	0.0396	0.0466	0.0467	0.0446	0.0538	0.0550
Past Month Marijuana Use	18-25	0.1643	0.0248	0.0292	0.0293	0.0286	0.0346	0.0354
Past Month Marijuana Use	50+	0.0154	0.1423	0.0924	0.1186	0.1503	0.0994	0.1294
Past Month Marijuana Use	AIAN, 12+	0.0729	0.1911	0.1905	0.1731	0.2020	0.2103	0.1976
Past Month Marijuana Use	12+	0.0583	0.0312	0.0289	0.0278	0.0385	0.0369	0.0365
Past Month Marijuana Use	API,12+	0.0279	0.1592	0.1513	0.1480	0.1769	0.1712	0.1697
Past Month Marijuana Use	Pregnant, 12+	0.0404	0.2241	0.2405	0.2154	0.2281	0.2490	0.2261
SMI Proxy (18+)	18+	0.1092	0.0229	0.0191	0.0195	0.0238	0.0200	0.0206
Past Year Specialty Substance Use Treatment	12+	0.0097	0.0748	0.0643	0.0640	0.0755	0.0654	0.0657

AIAN = American Indian or Alaska Native; API = Asian or Other Pacific Islander; MDE = major depressive episode; SMI = serious mental illness.

Note: Prevalence is expressed in absolute terms (e.g., 0.0752 = 7.52 percent).

Table 3.26 Relative Precision, by Design and Domain Group, Percentages

Design and Estimate Domain Group	Cluster Size = 9.375			Cluster Size = 28.125		
	Allocation 1	Allocation 2	Allocation 3	Allocation 1	Allocation 2	Allocation 3
50-State Design						
All Ages (12+)	100	137	137	83	114	114
Mixed (18+ or Pregnant)	100	112	129	93	104	119
Old	100	208	148	82	172	121
Young	100	89	83	87	77	72
National Design						
All Ages (12+)	51	69	70	32	43	44
Mixed (18+ or Pregnant)	72	89	98	63	76	83
Old	72	171	105	48	117	72
Young	64	47	47	40	30	30
Hybrid Designs						
All Ages (12+)	60	78	78	44	53	51
Mixed (18+ or Pregnant)	83	97	106	77	87	93
Old	79	188	114	62	139	82
Young	72	52	51	53	36	35

Examination of the columns supports decreases in precision with decreases in sample size. Precision also decreases with increasing cluster size, but the decreases in precision are accompanied by decreases in cost (see Tables 3.20, 3.21, and 3.22). Allocations 2 and 3, which sample older persons at higher rates, have better precision than the currently used allocation 1 for three of the four domain groups. The young domains have much lower precision under these allocation plans. Although the effects of design changes on precision are informative, any decisions about allocations among age groups should be based on data needs and public policy requirements rather than comparison with the current design.

3.10 Field Management Issues

3.10.1 Increasing Cluster Size

Increasing the cluster size by a factor of 3 is proposed in 9 of the 18 sampling plans. Currently, field interviewers (FIs) are assigned two or more segments to work per quarter. Many of them can handle a larger workload—as many as four or five segments during the quarter. Tripling the cluster size and maintaining an assignment of two clusters (segments) will provide an adequate assignment to keep the most productive interviewers occupied during the quarter. Some attrition of less productive staff is expected, such as those who could not work more than 5 to 10 hours per week or those who could not travel more than 50 miles from home. Once these FIs resigned or were terminated for lack of work in their area, the FI team will stabilize and the turnover rate will remain about what it is under the current design. Because the overall size of the FI team will diminish, the size of the field management team is expected to shrink as well.

Section 3.10.5 presents greater detail on the expected number of FIs and field management staff based on the various cluster sizes.

Under most of the alternative designs, a larger proportion of FIs will most likely be working full time (i.e., 30 to 40 hours per week) than under the current design. A larger proportion of highly skilled FIs working more hours per week might decrease the FI turnover rate overall, but this might also increase field costs because FIs on average will be farther from the sampled segments. It is not clear that increased travel costs from using fewer FIs will offset any savings brought by a lower turnover rate. Even so, some gain might be seen in response rates because of more highly qualified FIs completing a larger proportion of cases. Another consideration is that some State laws may require that FIs working over a certain number of hours per week be offered some form of fringe benefits by the employer, which will increase wage costs. However, as noted above, if the overall staff size shrank sufficiently, the size of the field management staff could shrink also, which could offset increased FI wage costs.

Overall, having more full-time FIs might increase costs slightly initially, might lead to higher response rates, and over time might pay for itself based on lower FI turnover rates and a shrunken field management staff size. This approach needs to be explored further before drawing more solid conclusions. For example, one possibility that could be explored is guaranteeing the equivalent of full-time hours to the most qualified FIs who are available for travel.

Some segments will naturally vary from the expected average because of unexpected population changes or poor advance information. These circumstances currently occur, also. The challenge to manage each interviewer's workload and to provide assistance within some segments will remain. The major effect will be fewer travel points as a result of simultaneously reducing the number of segments.

3.10.2 Decreasing the Sampling Rate for Young Age Groups

As noted in Section 3.8 on cost comparisons, reducing the youth sample will decrease the number of screening interviews required to yield the required sample. The older age groups typically have lower response rates. The unweighted response rate should be expected to decrease as the sample distribution shifts toward older age groups; the weighted response rate may also decrease slightly because response rates for older persons appear to be better when they are selected along with a younger person, which will happen less often.

Some of the savings from reduced screenings may be offset by more callbacks required to obtain responses from older sample members. Much of this is speculation and will not be known until it is tried.

3.10.3 Fewer Travel Points

Larger clusters will mean fewer travel points. For the current and hybrid designs, the field organization will still be organized around States. More field supervisors (FSs) will handle multiple State assignments as the sample size is reduced under the hybrid designs. An attempt has been made to incorporate the shift to more travel within and less between segments in developing the cost model parameters.

Considerable challenges are experienced in maintaining required sample sizes by State. The variation in yield may increase as the number of segments per State decreases. Maintaining adequate State sample sizes without large excesses may become more challenging.

3.10.4 National Design

Moving to a national design will shift the sample distribution from sparsely populated and mostly rural areas to higher-density urban areas. Theoretically, weighted response rates should remain about the same, but unweighted response rates should be expected to decrease.

Some realignment of staff will be required. With a reduced sample size of 40,000, most of the reductions in interviewing staff will occur in the low-population States. Only one State (California) will have some increase in sample size.

The structure and organization of the field staff will no longer focus on State requirements. The employment of each FI under the current design is contingent on adequate performance and the availability of work in the FI's local area, unless the FI is designated as a traveler, in which case employment is contingent on adequate performance and the availability of work anywhere to which the FI is able to travel. With the stability of the location of SS regions and segments under the current design, FIs are never terminated as a result of a lack of work in their local area. As the location of sample segments shifts under a national design, it is expected that some portion of FIs will be terminated at the end of each quarter or year based on a lack of work in their local area. New FIs will need to be recruited and trained for any newly sampled areas. Drawing a sample of SS regions as PSUs will help concentrate the work into more reasonable coverage areas for a single interviewer assigned to one PSU.

The control of the quarterly sample allocation will focus on national yield only and simplify the process of quarterly projections and sample size adjustments.

3.10.5 Field Staff Sizes

Based on all of the above, Tables 3.27 to 3.29 provide approximate numbers of FIs, FSs, regional supervisors (RSs), and regional directors (RDs) under each of the designs shown in Tables 3.20 to 3.22 in Section 3.8.

3.11 State and Substate Estimates

3.11.1 National Design

The national design is not suitable for State or substate estimation. Only the 12 largest population States have annual nominal sample sizes in excess of the current target of 900 for small sample States. Even in States with relatively large samples, there is no guarantee that substate sample sizes will be adequate for reasonably accurate small area estimation (SAE). Because the SS regions are not strata in the national design, their sample sizes are not targeted and will be subject to random variation.

Table 3.27 Field Staff Size for Various 50-State Designs with the Current Sample Size

	Cluster Size = 9.375			Cluster Size = 28.125		
	Allocation 1	Allocation 2	Allocation 3	Allocation 1	Allocation 2	Allocation 3
Design Parameter						
PSUs (SS Regions)	900	900	900	300	300	300
Segments	7,200	7,200	7,200	2,400	2,400	2,400
Completed Screenings	140,041	105,031	105,031	140,041	105,031	105,031
Completed Interviews	67,500	67,500	67,500	67,500	67,500	67,500
Field Interviewers (FIs)	680	636	636	537	493	493
Field Supervisors (FSs)	38	36	36	30	28	28
Regional Supervisors (RSs)	7	7	7	6	5	5
Regional Directors (RDs)	3	3	3	2	2	2

PSU = primary sampling unit; SS = State sampling.

Table 3.28 Field Staff Size for National Design of 40,000

	Cluster Size = 9.375			Cluster Size = 28.125		
	Allocation 1	Allocation 2	Allocation 3	Allocation 1	Allocation 2	Allocation 3
Design Parameter						
PSUs (SS Regions)	533	533	533	178	178	178
Segments	4,264	4,264	4,264	1,424	1,424	1,424
Completed screenings	82,987	62,240	62,240	82,987	62,240	62,240
Completed interviews	40,000	40,000	40,000	40,000	40,000	40,000
Field Interviewers (FIs)	443	416	416	358	332	332
Field Supervisors (FSs)	25	23	23	20	19	19
Regional Supervisors (RSs)	5	4	4	4	3	3
Regional Directors (RDs)	2	2	2	2	0	0

PSU = primary sampling unit; SS = State sampling.

Table 3.29 Field Staff Size for Hybrid Designs

	Cluster Size = 9.375			Cluster Size = 28.125		
	Allocation 1	Allocation 2	Allocation 3	Allocation 1	Allocation 2	Allocation 3
Design Parameter						
PSUs (SS Regions)	712	660	642	264	239	230
Segments	5,696	5,280	5,136	2,112	1,912	1,840
Completed screenings	110,854	76,993	74,927	123,300	83,534	80,682
Completed interviews	53,432	49,481	48,153	59,431	53,685	51,852
Field Interviewers (FIs)	558	492	481	485	412	401
Field Supervisors (FSs)	31	27	27	27	23	22
Regional Supervisors (RSs)	6	5	5	5	4	4
Regional Directors (RDs)	2	2	2	2	2	2

PSU = primary sampling unit; SS = State sampling.

3.11.2 Hybrid Designs

The hybrid designs in Table 3.15 (see Section 3.6) are specified to provide an effective sample size of 400 per year based on the design effect for past month marijuana use in the 12 or older population. The corresponding annual effective sample size for small sample States in the current design is 483 (current State sample of 900 divided by allocation 1, cluster size 9.375, with a design effect of 1.861, as shown in Table 3.15). Therefore, to provide State-level small area estimates with precision strictly comparable with the current design will require at least 2.4 years of data from the hybrid designs rather than the current design's 2 years ($483/400 \times 2 \approx 2.4$). Because of the 50 percent overlap of segments from one year to the next, State estimates that use 2 years of survey data combined have one third of their segments with a double or 2-year allocation of respondents. Therefore, the average segment cluster size for a 2-year State sample is 12.5 ($[2/3 \times 9.325] + [1/3 \times 18.75] \approx 12.5$) with the current cluster size and 37.5 with the tripled (28.125) cluster size. This suggests that for the three hybrid designs with annual cluster sizes of 28.125, a combined 3 years of data could be required to match 2 years from the current design.

For substate estimates where 3 years of data are combined with the current design, it will take roughly 3.6 years combined from the first three hybrid designs in Table 3.15 to achieve the same realized sample size. Combining over 3 years causes one half of the sample segments to get double allocations of respondents. Therefore, the current design's 3-year average cluster size of 14.06 expands to 42.19 for the last three hybrids in Table 3.15. Because of correlated responses within a cluster, this suggests that combining at least 4 years of hybrid data may be required for substate estimates if the initial cluster size is tripled.

3.11.3 Impact of Stratum Sizes on Substate Estimates

As mentioned in Section 3.11.1, substate estimates are not practical in the national design. In the hybrid design, the sample size will increase in five States and will decrease in all others. In States with an increased sample, the ability to produce substate estimates will be better than the current design for the target cluster size of 9.375 because of an increased number of strata (SS regions) and a decrease in the geographic size of these strata. For the target cluster size of 28.125, the overall effect will be a decrease in the number of SS regions even in the States with an increased sample size. Seven of the original eight large sample States will show decreases in sample size, but will still have sample allocations exceeding 900 prescribed in the current design for small sample States.

Substate regions are defined by the States, and, generally, their boundaries do not coincide with those of the SS regions. As a result, some variation in achieved sample size arises from the randomness associated with the inclusion of sample segments within substate region boundaries. This variation will increase as the substate regions get larger either because of a reduction in sample size, an increase in cluster size, or both.

3.12 Response Rates

Unweighted overall response rates will decrease for the age allocations that shift the sample distribution toward older age groups; this occurs for age allocations 2 and 3 shown in Table 3.14 in Section 3.5.3. Weighting inversely to selection probabilities will adjust away most of this increase, but some additional decrease in response rates may occur because of selecting more of the sample in pairs under age allocations 2 and 3. Chromy and Penne (2002, Table 4) projected some reduction in response rates across all age groups for designs that incorporate more selection of respondents within pairs. Age allocations 2 and 3 are likely to increase the number of pair selections because they also decrease the number of required screenings. The selection of pairs can be controlled to some extent by the specification of a parameter in the current sample selection algorithm. With some adjustment in this parameter, the overall effect of alternative designs on response rates should be minimal.

3.13 Conclusions

A preliminary analysis included several solutions to an age group optimization among seven age groups (12 to 14, 15 to 17, 18 to 25, 26 to 34, 35 to 49, 50 to 64, and 65 or older). The major findings from this analysis were as follows:

- By relaxing all sample size requirements, SAMHSA precision requirements can be met with a much smaller sample than is currently used.
- By relaxing the requirement of equal sample to the three major age groups, the solution suggested that the optimal age group distribution is 22, 27, and 51 percent to the 12 to 17, 18 to 25, and 26 or older age groups, respectively.
- By requiring that the total sample of 67,500 persons be allocated equally to the 12 to 17, 18 to 25, and 26 or older age groups, the solution puts more sample in the 15 to 17, 50 to 64, and 65 or older age groups than does the current design.

Using these results and findings from other Sample Issues Methods Study investigations, SAMHSA subsequently refined the design specifications and several alternative sample designs were evaluated. These designs included a national design of 40,000 persons, a hybrid design that preserves minimal sample sizes for some States, and a 50-State design with the current sample size. For each of these designs, three age group allocations and two cluster size options were evaluated.

In summary, the national design with a sample of 40,000 offers the greatest cost reductions, in large part to the reduction in field staff required. Hybrid designs still produce major cost reductions while preserving minimal sample sizes for some State estimates. On the other hand, the precision analysis favors a 50-State design with a smaller cluster size. State and substate estimates are not practical under the national design, and under the hybrid design, State estimates will require pooling data over more years than are currently required. Finally, unweighted overall response rates are expected to decrease for the age allocations that shift the sample distribution toward older age groups.

4. Continuous Design Assessment

4.1 Introduction

4.1.1 Purpose

Since 1999, the National Survey on Drug Use and Health (NSDUH) has collected interview data from about 67,500 persons annually. The sample is allocated approximately equally to four calendar quarters. Under the current data collection plan, the sample is released at the beginning of each calendar quarter,[20] and most of the data are collected in the first 4 to 6 weeks of each quarter (i.e., in January, April, July, and October). As a result, most of the past month drug use references correspond with the months of December (of the prior year), March, June, and September. If there are no seasonal effects related to drug use, collecting data in this manner paints an accurate portrait of drug use across the year. On the other hand, if there are peaks or troughs in past month drug use, the current design may not be accurately portraying annual averages in drug use. In addition, evaluations of drug use across specific months of the year (e.g., an evaluation of summer drug use among college students) may not be accurate. Unlike the current design, a more continuous design is one that collects data on a continuous basis, with little or no downtime within data collection periods.

4.1.2 Background

Under the current NSDUH design, the sample is released at the beginning of each calendar quarter. To achieve the desired sample in each State, an additional 20 percent of the quarterly sample is selected, and the 120 percent sample is randomly partitioned into smaller subsamples. Typically, 80 to 120 percent of the sample is released at the beginning of the quarter, then small reserve samples of 5 to 20 percent are released midway through the quarter as needed in each State. Figure 4.1 shows the distribution of completed interviews by week of the quarter for the 2009 NSDUH. Figure 4.2 displays the cumulative distribution. In every quarter, most of the cases (approximately 80 percent) are completed in the first 6 weeks of the quarter, with weeks 2 and 3 yielding the most interviews. The remainder of the quarter is used to clean up all remaining pending cases. Typically, these pending cases are more challenging than those completed earlier in the quarter and may require extra effort or special skills for a field interviewer (FI) to complete. The following are some of the typical tasks completed during each quarters' cleanup phase:

- Transfer pending refusal cases to strong refusal converters.
- Identify the most viable pending cases on which to focus efforts.
- Identify cases that need additional visits.
- Hold small group conference calls to coach on refusal conversion techniques that will work best for remaining pending work.

[20] As part of the sample size management plan, each State's quarterly sample is partitioned into random subsamples. Although most of the sample is released at the beginning of the quarter, additional subsamples often are released 4 to 6 weeks into the quarter as needed in each State.

Figure 4.1 Percentage of Interviews Completed per Week in 2009

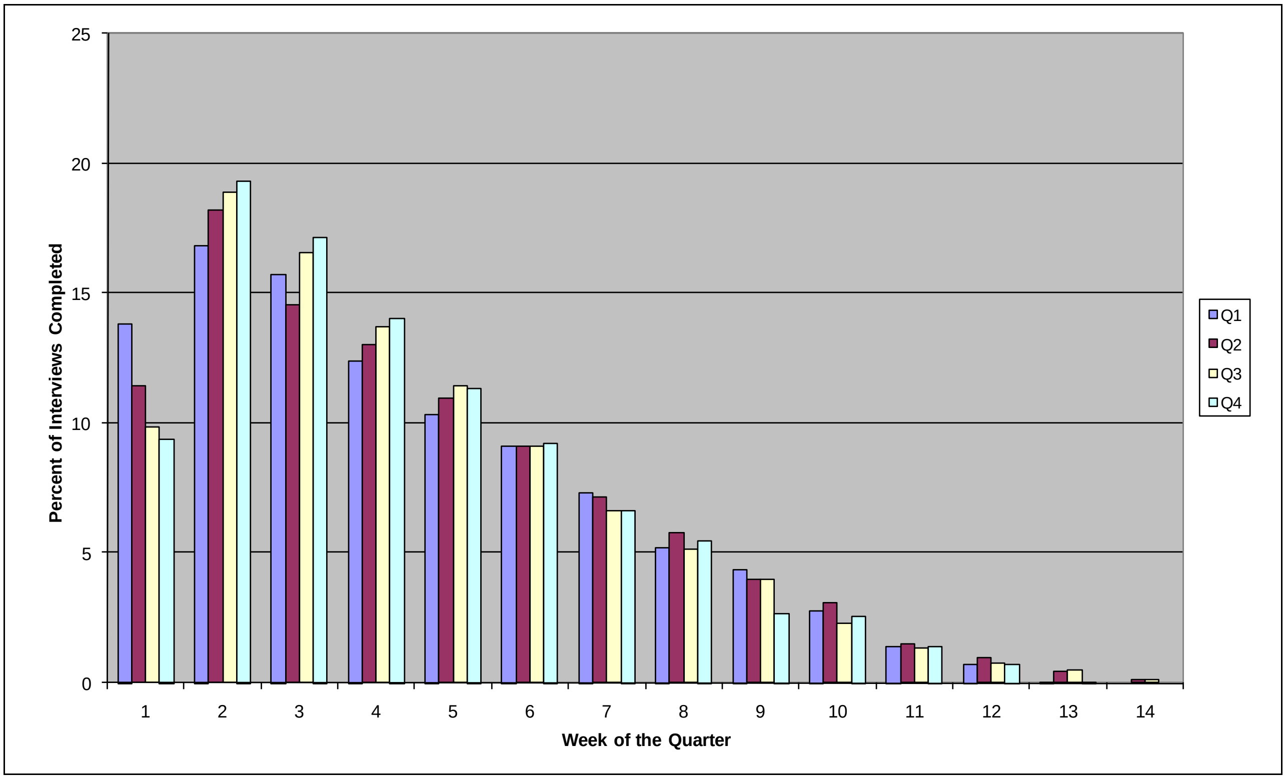
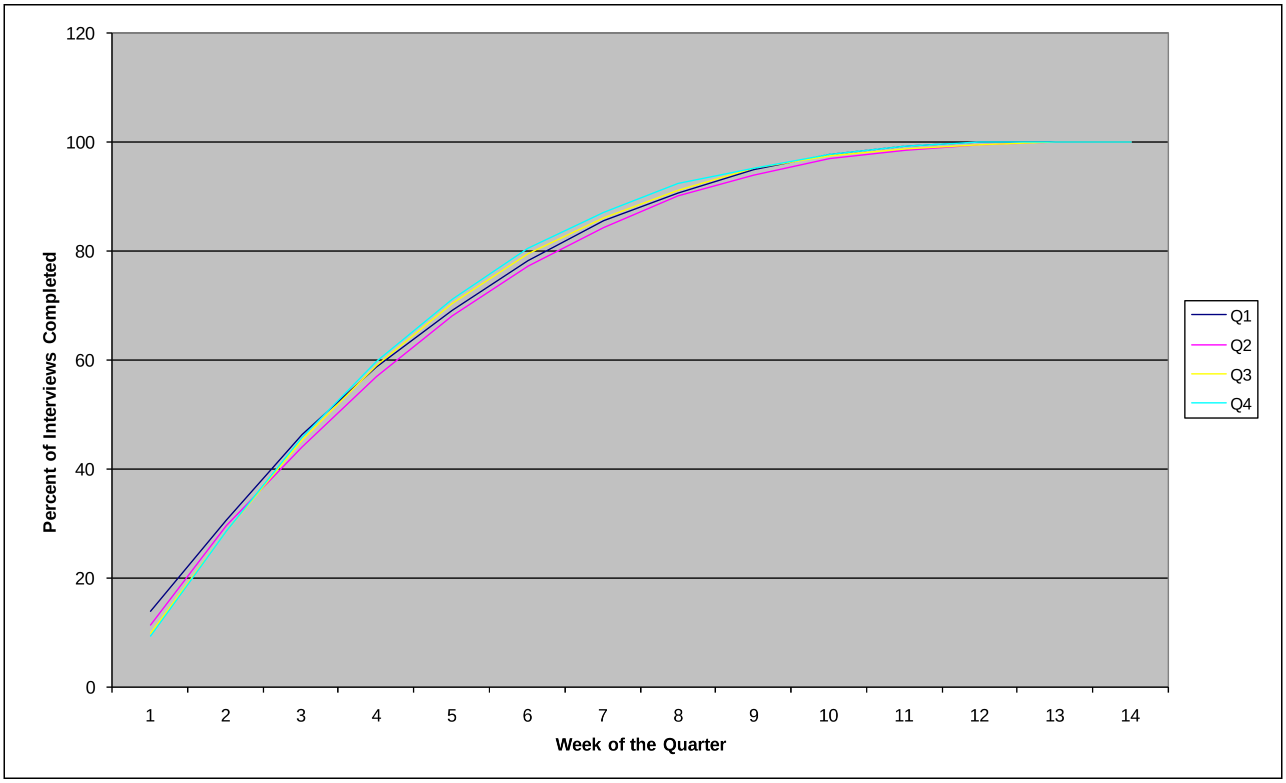


Figure 4.2 Cumulative Percentage of Interviews Completed per Week in 2009



- Consolidate remaining work with FIs who will have the most impact and are the most cost efficient.
- Schedule travel for bilingual FIs, where indicated, to complete Spanish interviews.
- Make decisions about and plan cleanup travel trips with traveling field interviews (TFIs) and borrowed field interviewers (BFIs).
- Mail call-me/meet-me letters as a last-ditch effort to gain cooperation in controlled access situations.

Some efficiency would be lost in attempting to spread out cases throughout the quarter. It is more cost efficient for FIs to work as many cases as possible and longer hours to avoid unnecessary back-and-forth travel to the segment.

In addition to the quarters being front heavy, there may be temporal differences in prevalence estimates between quarters. A more continuous design would allow the analyst to study seasonal differences in drug use and mental health measures with greater accuracy. Table 4.1 displays 2009 estimates by quarter for several substance use, substance dependence, substance treatment, and mental health measures. At the national level, there were small differences between the quarterly measures. However, only past month use of alcohol (quarter 1 vs. quarter 4 and quarter 3 vs. quarter 4) produced statistically significant differences. Seasonal differences for certain subgroups may be more pronounced. For example, Table 4.2 displays estimates by quarter for persons aged 18 to 25. For this age group, significant differences were found for past month use of cigarettes (quarter 1 vs. quarter 4), past month use of alcohol (quarter 1 vs. quarter 4 and quarter 3 vs. quarter 4), and past month serious psychological distress (SPD) (quarter 1 vs. quarter 2, quarter 2 vs. quarter 3, and quarter 2 vs. quarter 4). Table 4.3 displays these same measures by quarter for full-time college students aged 18 to 22. For this group, quarter 1 had a higher rate of past month alcohol use than quarters 3 and 4 (68.8 vs. 62.9 and 59.4 percent, respectively), and quarter 2 (65.7 percent) also had a higher rate than quarter 4. Rates of past month SPD were lower in quarter 1 than in quarter 2 (4.9 vs. 7.4 percent), higher in quarter 2 than in quarter 3 (7.4 vs. 4.1 percent), and lower in quarter 3 than in quarter 4 (4.1 vs. 5.9 percent).

Prior studies have shown seasonal patterns in substance use among youths aged 12 to 17. As summarized in Gfroerer, Wu, and Penne (2002), a few studies have found increased rates of marijuana use during the summer months, suggesting that patterns in marijuana use may be related to the cycles in school attendance. On the other hand, an analysis of data from the 1992 to 1996 National Household Surveys on Drug Abuse (NHSDAs) (NSDUH's name prior to 2002) found a lower prevalence of current marijuana use during July to September (Huang, Schildhaus, & Wright, 1999). This analysis found seasonal differences in past month use of alcohol, an illicit drug except marijuana, marijuana only, an illicit drug, and heavy drinking among youths aged 12 to 17. Finally, using 1999 and 2000 NHSDA data, Gfroerer et al. (2002) found that 21 percent of all recent initiates of marijuana use reported initiating in the months of June and July.

Table 4.1 Substance Use, Substance Dependence, Substance Use Treatment, and Mental Health Measures, by Quarter of Data Collection: Percentages, 2009

Measure	Total	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Past Month Substance Use					
Cigarettes	23.3	23.0	23.0	23.6	23.6
Alcohol	51.9	53.5[c]	51.4	52.9[c]	49.8
Illicit Drugs[1]	8.7	8.8	9.0	8.5	8.4
Illicit Drugs Other Than Marijuana[1]	3.6	3.7	3.8	3.4	3.6
Cocaine	0.7	0.7	0.7	0.7	0.5
Past Year Dependence[2]					
Alcohol	3.5	3.3	3.3	4.0	3.5
Illicit Drugs[1]	1.9	2.0	2.1	1.7	1.8
Past Year Substance Treatment[3]					
Alcohol	1.2	1.1	1.3	1.4	1.2
Illicit Drugs[1]	0.9	1.1	0.9	0.9	0.8
Mental Health Measures[4]					
Serious Psychological Distress in the Past Month[5]	4.6	4.8	4.8	4.4	4.4
Major Depressive Episode in the Past Year[6]	6.5	6.6	6.9	6.6	6.1

*Low precision; no estimate reported.

[a] Difference between estimate and quarter 2 estimate is statistically significant at the 0.05 level.

[b] Difference between estimate and quarter 3 estimate is statistically significant at the 0.05 level.

[c] Difference between estimate and quarter 4 estimate is statistically significant at the 0.05 level.

[1] Illicit Drugs include marijuana/hashish, cocaine (including crack), heroin, hallucinogens, inhalants, or prescription-type psychotherapeutics used nonmedically. Illicit Drugs Other Than Marijuana include cocaine (including crack), heroin, hallucinogens, inhalants, or prescription-type psychotherapeutics used nonmedically. The estimates are based on data from original questions not including methamphetamine items added in 2005 and 2006.

[2] Dependence is based on the definition found in the 4th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV).

[3] Received Substance Use Treatment refers to treatment received in order to reduce or stop illicit drug or alcohol use, or for medical problems associated with illicit drug or alcohol use. It includes treatment received at any location, such as a hospital (inpatient), rehabilitation facility (inpatient or outpatient), mental health center, emergency room, private doctor's office, self-help group, or prison/jail.

[4] These mental health measures are among adults aged 18 or older.

[5] Serious Psychological Distress (SPD) is defined for this table as having a score of 13 or higher on the Kessler-6 (K6) scale during the past 30 days.

[6] Major Depressive Episode (MDE) is defined as in the 4th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV), which specifies a period of at least 2 weeks when a person experienced a depressed mood or loss of interest or pleasure in daily activities and had a majority of specified depression symptoms.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2009.

Table 4.2 Substance Use, Substance Dependence, Substance Use Treatment, and Mental Health Measures among Persons Aged 18 to 25, by Quarter of Data Collection: Percentages, 2009

Measure	Total	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Past Month Substance Use					
Cigarettes	35.8	37.4[c]	36.6	35.1	34.2
Alcohol	61.8	64.0[c]	61.3	62.6[c]	59.5
Illicit Drugs[1]	21.2	20.8	22.7	21.5	19.9
Illicit Drugs Other Than Marijuana[1]	8.3	8.1	9.4	8.1	7.6
Cocaine	1.4	1.7	1.5	1.5	0.9
Past Year Dependence[2]					
Alcohol	6.6	6.8	6.9	7.0	5.7
Illicit Drugs[1]	5.4	5.7	6.1	5.1	4.7
Past Year Substance Treatment[3]					
Alcohol	2.0	1.9	1.9	1.8	2.3
Illicit Drugs[1]	1.8	2.0	1.7	1.5	1.9
Mental Health Measures					
Serious Psychological Distress in the Past Month[4]	7.6	7.2[a]	8.9[b,c]	7.0	7.2
Major Depressive Episode in the Past Year[5]	8.0	8.0	8.8	7.3	7.8

*Low precision; no estimate reported.

[a] Difference between estimate and quarter 2 estimate is statistically significant at the 0.05 level.

[b] Difference between estimate and quarter 3 estimate is statistically significant at the 0.05 level.

[c] Difference between estimate and quarter 4 estimate is statistically significant at the 0.05 level.

[1] Illicit Drugs include marijuana/hashish, cocaine (including crack), heroin, hallucinogens, inhalants, or prescription-type psychotherapeutics used nonmedically. Illicit Drugs Other Than Marijuana include cocaine (including crack), heroin, hallucinogens, inhalants, or prescription-type psychotherapeutics used nonmedically. The estimates are based on data from original questions not including methamphetamine items added in 2005 and 2006.

[2] Dependence is based on the definition found in the 4th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV).

[3] Received Substance Use Treatment refers to treatment received in order to reduce or stop illicit drug or alcohol use, or for medical problems associated with illicit drug or alcohol use. It includes treatment received at any location, such as a hospital (inpatient), rehabilitation facility (inpatient or outpatient), mental health center, emergency room, private doctor's office, self-help group, or prison/jail.

[4] Serious Psychological Distress (SPD) is defined for this table as having a score of 13 or higher on the Kessler-6 (K6) scale during the past 30 days.

[5] Major Depressive Episode (MDE) is defined as in the 4th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV), which specifies a period of at least 2 weeks when a person experienced a depressed mood or loss of interest or pleasure in daily activities and had a majority of specified depression symptoms.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2009.

Table 4.3 Substance Use, Substance Dependence, Substance Use Treatment, and Mental Health Measures among Full-Time College Students Aged 18 to 22, by Quarter of Data Collection: Percentages, 2009

Measure	Total	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Past Month Substance Use					
Cigarettes	27.1	27.5	28.8	26.2	26.2
Alcohol	63.9	68.8[b,c]	65.7[c]	62.9	59.4
Illicit Drugs[1]	22.7	22.4	25.7	23.5	20.1
Illicit Drugs Other Than Marijuana[1]	8.6	7.5	11.1	8.6	7.7
Cocaine	1.7	1.4	2.2	2.3	1.0
Past Year Dependence[2]					
Alcohol	6.9	7.2	7.7	7.0	5.8
Illicit Drugs[1]	4.8	5.2	5.6	4.9	3.9
Past Year Substance Treatment[3]					
Alcohol	1.1	1.6	0.9	1.1	1.0
Illicit Drugs[1]	0.8	0.8	0.9	0.7	0.8
Mental Health Measures					
Serious Psychological Distress in the Past Month[4]	5.5	4.9[a]	7.4[b]	4.1[c]	5.9
Major Depressive Episode in the Past Year[5]	7.8	7.9	8.6	6.8	8.2

*Low precision; no estimate reported.

[a] Difference between estimate and quarter 2 estimate is statistically significant at the 0.05 level.

[b] Difference between estimate and quarter 3 estimate is statistically significant at the 0.05 level.

[c] Difference between estimate and quarter 4 estimate is statistically significant at the 0.05 level.

[1] Illicit Drugs include marijuana/hashish, cocaine (including crack), heroin, hallucinogens, inhalants, or prescription-type psychotherapeutics used nonmedically. Illicit Drugs Other Than Marijuana include cocaine (including crack), heroin, hallucinogens, inhalants, or prescription-type psychotherapeutics used nonmedically. The estimates are based on data from original questions not including methamphetamine items added in 2005 and 2006.

[2] Dependence is based on the definition found in the 4th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV).

[3] Received Substance Use Treatment refers to treatment received in order to reduce or stop illicit drug or alcohol use, or for medical problems associated with illicit drug or alcohol use. It includes treatment received at any location, such as a hospital (inpatient), rehabilitation facility (inpatient or outpatient), mental health center, emergency room, private doctor's office, self-help group, or prison/jail.

[4] Serious Psychological Distress (SPD) is defined for this table as having a score of 13 or higher on the Kessler-6 (K6) scale during the past 30 days.

[5] Major Depressive Episode (MDE) is defined as in the 4th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV), which specifies a period of at least 2 weeks when a person experienced a depressed mood or loss of interest or pleasure in daily activities and had a majority of specified depression symptoms.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2009.

Although there may be concerns with the current method of data collection, a more continuous method of data collection may have its own consequences. This chapter describes the pros and cons of switching NSDUH to a more continuous design. Several options are evaluated, including collecting data either every 2, 4, or 6 weeks. In addition, the option of extending case assignments into the following data collection period is explored. Table 4.4 summarizes the statistical and operational implications of a more continuous design, while Section 4.2 provides greater detail on these items. Finally, the pros and cons are briefly examined of a less continuous design (Section 4.3), a middle ground solution of collecting data every 2 months (Section 4.4), and the American Community Survey (ACS) continuous design model (Section 4.5). Section 4.6 provides a summary of the different design methods discussed.

4.2 Evaluation of Continuous Design Options

4.2.1 Survey Planning

4.2.1.1 Sampling

Releasing the sample and collecting data on a more continuous basis would provide a greater ability to adjust the sample and achieve annual State and national targets. To compensate for quarterly variations in response rates and yields, the quarterly sample is currently partitioned into smaller subsamples. To achieve the desired sample in each State, 80 to 120 percent of the sample is released at the beginning of the quarter, then small reserve samples of 5 to 20 percent are released midway through the quarter, as needed. The use of sample partitions provides a means of managing the sample size to meet quarterly and annual goals by State. With samples released every 2, 4, or 6 weeks, there would no longer be a need for the sample partitions that are currently used to manage the NSDUH sample. Instead, sample sizes would be adjusted when preparing the sample for each data collection period. However, the efficiency gained by not having to prepare and release sample partitions would be negated by having to prepare more frequent samples and adjusting the sample sizes at each iteration.

A more continuous design also would simplify the NSDUH eligibility rule. The current eligibility rule requires that a person reside at the sampled dwelling unit for the majority (i.e., more than half) of the calendar quarter. For more mobile populations (e.g., college students), the question of eligibility may become clearer if the length of the data collection period was reduced from 13 weeks to 2, 4, or 6 weeks.

Finally, shorter data collection periods would improve the ability to adjust the sample in response to natural disasters and major national events. Supplementing the sample in the impacted areas would be timelier because the sample would be released more frequently than 4 times a year. These samples could be changed or supplemented as frequently as every 2, 4, or 6 weeks, depending on the length of the data collection period.

Although, from a design perspective, there are benefits for moving to a more continuous design approach, the downside is that this approach would increase costs because samples would have to be prepared more frequently. The costs of the various continuous design options are estimated in Section 4.2.5.

Table 4.4 NSDUH: Comparison of Continuous Designs to the Current Design

Design	Description	Survey Planning (Questionnaire Changes, Sample Expansion, etc.)	Training Programs	Field Operations	Equipment and CMS	Analysis and Reporting	Annual Cost
Current	Collect data quarterly	Sample managed through release of subsamples Questionnaire changes and sample expansion limited to next calendar quarter	Veteran and NTP training sessions held during downtime at the end of each quarter	Downtime used for management activities and staff vacations Sufficient time for overcoming problems associated with access barriers, weather, etc.	Systems configured to support parallel activities associated with data collection in consecutive quarters	Preliminary analyses conducted after 2 calendar quarters Annual data processing begins in January Aggressive schedule for analysis and reporting to meet September press release	No change in cost
Continuous	Collect data every 2, 4, or 6 weeks	Greater ability to adjust sample and respond to major events Simplifies implementation of eligibility rule Less impact if a problem is found during early data review Increased cost of selecting more frequent samples	No downtime creates a challenge for scheduling training sessions	Greater need for TFIs Possible FI attrition and burnout due to loss of downtime Greater impact of weather and staff being out for vacation/personal reasons Difficult to overcome access barrier situations due to time constraints No time to follow up on falsified work Potential decline in response rates Increased costs due to loss in efficiency	Basic reworking of systems required Slightly more effort to manage	Minimal impact on editing, imputation, and weighting Ability to conduct preliminary analyses earlier	Increase in cost associated with releasing sample on a more continuous basis Costs increase as the duration of the data collection period decreases
Continuous with Overlapping Data Collection Periods	Collect data every 2, 4, or 6 weeks Allow case assignments to extend into the following data collection period	Possibility of using two questionnaires in January	No downtime creates a challenge for scheduling training sessions	Additional challenges associated with managing two samples	Basic reworking of systems required Slightly more effort to manage	Adjustment to analysis and reporting schedule required	Increase in cost associated with releasing sample on a more continuous basis Costs increase as the duration of the data collection period decreases

CMS = Case Management System; FI = field interviewer; NTP = new-to-project; TFIs = traveling field interviewers.

From a sampling perspective, the pros and cons of moving to a more continuous design are the same regardless of whether all case assignments are completed within the data collection period or if cases are allowed to extend into the following data collection period.

4.2.1.2 Instrumentation

Early data review and fixing instrumentation problems are the primary NSDUH instrumentation advantages from adopting a continuous design. Currently, the instrumentation team undertakes early data review early in the first quarter of annual data collection. If inconsistencies or errors are found in the instrument, the instrumentation team reviews the programming and specifies, programs, and tests the correction. Although it is not necessary to wait, patches are not typically released before the beginning of the next quarter's data collection, meaning a lag of 2 months or more may be observed before the instrument is revised. This practice is observed to maximize the likelihood that all interviewers are using the same instrument to collect data within a data collection period.

With a continuous design, regardless of the 2-week, 4-week, or 6-week schedule, interviews would be spread out more evenly across the quarter. Because data would not be so concentrated in the early weeks of data collection, an early data review may identify problems with the instrument using a smaller number of interviews. If less work is completed before the problem is found, fewer potential effects from questionnaire problems will be realized.

This advantage of a continuous design holds if overlapping data collection periods are implemented, regardless of the schedule. The only additional complication of the overlapping data collection periods would be the possibility of using two questionnaires in January. If the instrument underwent major changes, this could be confusing to field staff. However, major instrument changes are unlikely because the questionnaire is frozen for many years at a time to maintain consistency of estimates. For years in which the instrument is redesigned, special trainings may have to be implemented to cover the new procedures. One possible approach to overcome complications related to redesigning the questionnaire is to modify the schedule at the end of the calendar year. For example, December cases could be finalized prior to the start of the following data collection period (i.e., no overlapping), or the data collection period could be shortened.

4.2.2 Training Programs

A major challenge of a 2-, 4-, or 6-week data collection period is the loss of downtime during the quarter. Currently on NSDUH, this down period (typically the last 2 to 3 weeks of the quarter) is used to complete refresher training for veteran FIs (via a quarterly refresher iLearning course) and conduct New-to-Project (NTP) training sessions prior to the start of data collection in the next quarter. Although the data collection period would change depending on the design adopted, the overall construct of the veteran and NTP FI training programs would remain similar to the current NSDUH design, as data collection needs and staffing would be continuous.

In addition to refresher training conducted during the year, all veteran FIs and field supervisors (FSs) would also complete training prior to the start of each new survey year, which would consist of reviewing the FI manuals and completing several iLearning courses at home, followed by a 1-day, in-person FS team meeting in January. All new FIs and FSs would be required to attend NTP training and fulfill all training requirements, including passing certification, to successfully graduate. The NTP training program would remain very similar to the current NTP sessions conducted on NSDUH, incorporating any changes as a result of the new design. Thus, the challenge becomes scheduling training for both new and veteran FIs, which maintains staffing levels and FI knowledge, while not interfering with data collection given less downtime under a more continuous design.

One potential method to mitigate this staffing challenge would be to hold smaller, more frequent NTP training sessions to train new staff for regions with immediate needs, perhaps on a 6-week or bimonthly schedule. This change would lessen the number of project staff serving as trainers at each session and allow for more focused recruiting. However, an increase in the overall number of NTP training sessions would result in a larger number of project staff needed as trainers for the sessions throughout the year, including FSs and other management staff who would have to manage their current assignments and workloads along with their trainer role. With more frequent NTP trainings and as staffing levels change, one advantage is that FSs would have more frequent opportunities to recruit staff in regions that are experiencing staffing shortfalls. Following the NTP session, FSs would have a smaller number of new FIs to mentor and manage, but this process would still require a great amount of time and effort to provide oversight and guidance. Finally, although the NTP training sessions would be smaller, conducting additional training sessions throughout the year to keep staffing levels constant would result in an increase in training costs as compared with the current design.

With the loss of downtime between quarters, a similar difficulty would be present in scheduling a time to conduct the end-of-year training for all veteran FI staff to review changes for the upcoming survey year, complete several iLearning courses, and hold a 1-day, in-person FS team meeting in January. The training schedule would have to be adjusted to allow adequate time for FIs to complete these tasks and attend their FS team meeting, along with conducting data collection during this time.

In addition to the challenges in scheduling end-of-year trainings, the loss of down periods between quarters may result in needing more TFIs and FIs because this down period is typically a time when FIs take time off, attend other training sessions, work other part-time jobs, and so on. Section 4.2.3.4 provides additional information on how staffing levels might be impacted with the continuous design option.

4.2.3 Field Operations

A continuous design with 2-, 4-, or 6-week data collection periods would significantly affect field operations and would possibly have a negative impact on response rates, costs, staffing, and data quality. The 2-week model would produce the most challenges, and a 4- or 6-week model would still have some of the same issues, but to a lesser degree. Likewise, implementing an overlapping data collection period, where there would be a 2-week cleanup period after the data collection deadline, would allow some flexibility and would be less challenging than the 2-week design.

4.2.3.1 Impact on Field Preparations

Once the sample is released to the field and before staff begin collecting data, a number of tasks are performed. For example, segment kits and lead letters are shipped to all FIs with assignments and any additional materials needed. FIs sign their lead letters and send them to selected households. The preparatory time needed and logistics required to provide the field staff with these materials in each of these continuous design options needs to be well thought out and taken into consideration.

4.2.3.2 Impact on Response Rates

Over the years, the incidence of controlled access situations has increased on field surveys and will likely continue to increase in future years. The continuous design model would reduce the amount of time available to adequately work through some of the steps involved in communicating with building management personnel and university institutional review boards (IRBs) and attending board meetings to explain the study. Based on project experience, some controlled access negotiations can take up to 4 to 6 weeks to work with layers of management and/or IRBs to obtain the necessary permissions to work in a building, neighborhood, boarding school, or on a college campus. If staff do not have ample time to work through these steps, they will be forced to finalize the cases as nonresponses, resulting in lower response rates. Understandably, this would be less of an issue in the 6-week design option, but would be a greater problem for samples that need to be worked in a 4-week or 2-week period.

To illustrate this point, Table 4.5 provides the number of times controlled access situations were encountered in colleges, universities, and boarding schools and the number of lines affected between 2006 and 2010. Although staff were not required to go through the IRB approval process in all of these situations, it is important to note that staff had to overcome some sort of access barrier at each of these institutions, with some perhaps requiring only a few days of negotiations, while others requiring several weeks of negotiating with personnel. It is also important to realize that even after permission is granted by an institution, FIs are often given limited time to be on the school's campus to meet with students.

Table 4.5 Controlled Access Situations in Colleges, Universities, and Boarding Schools, by Year, 2006-2010 NSDUHs

Year	Number of CA Situations Encountered at Colleges, Universities, and Boarding Schools	Total Number of Lines Affected
2010	56	953
2009	73	1,034
2008	54	787
2007	80	845
2006	74	911

CA = controlled access.

Each year, some regions are affected by adverse weather conditions, making it dangerous, and even impossible, for FIs to work for weeks at a time. Over the years, some regions have been plagued with harsh winter weather conditions, tornadoes, hurricanes, flooding, and widespread fires, just to name a few. Again, the 6-week design option, and perhaps the 4-week option, would likely allow enough time for weather conditions to improve in most regions so staff can safely work their assigned cases in areas affected by these unfortunate acts of nature. Alternatively, if the design allows for case assignments to overlap into the next data collection period, this delay may allow enough time for weather conditions to improve. However, response rates would likely suffer if staff in these areas are limited to a 2-week data collection period and could even result in some staff taking dangerous risks and feeling pressure to work in unsafe weather conditions.

Another factor to take into consideration would be the challenges faced in the field and the level of effort needed to adequately work refusal cases and hard-to-reach respondents. Depending on the type of refusal encountered, an FI may decide to finalize the case immediately (i.e., hostile refusal), or if it is a softer refusal, the FI may ask his or her FS to send a refusal letter and conduct a follow-up at a later date. FSs typically allow refusal cases a cooldown period before they send a letter, or in some instances, the FS may want to send a letter and a different FI. Either way, ample time is needed to adequately work these types of cases. Despite the best efforts, some respondents are simply hard to reach. In these instances, FIs have their FSs send unable-to-contact letters and/or call-me letters in an effort to make contact at some point during the data collection period. Again, the 6-week design would allow more flexibility in converting cases and/or finding someone at home than with the 4-week or 2-week design, but the 4-week and 2-week design options would make these follow-up letters and contacts very challenging and would negatively impact response rates.

Completion of Spanish cases would also be impacted by the continuous design option. NSDUH currently employs a limited number of bilingual FIs because of the relatively small number of cases completed in Spanish. FSs often borrow FIs from other regions as an efficient way to work the bilingual cases and improve response rates. Borrowing FIs for both Spanish cases and/or regions that simply need additional help would be more challenging because FIs would be committed to working their own assignments and would be unavailable to assist in other regions. Even if staff could be available to assist in other regions, the continuous design option does not allow much flexibility in making timely travel arrangements, handling the case transfers, and ample time to work the cases.

FI vacations could be devastating to a region if more than one or two FIs wanted to schedule time off at the same time. If multiple FIs in a region are sick or have personal emergencies at the same time, this too, could be devastating to the results in an FS's region. A shorter data collection period could greatly impact urban areas where currently three to four FIs work in multiple segments, yielding positive results by following behind previous interviewers.

The impact on response rates for each of the above-mentioned discussion points varies depending on which design option is considered. The 6-week continuous data collection design is most similar to the current design and its effects on response rates would likely be less than the other two designs. It is anticipated that response rates would suffer more with the 4-week design option and the most with the 2-week design option. Implementing an overlapping data collection period, where there would be a 2-week cleanup period after the data collection deadline, would allow some flexibility and would likely have less of an impact on response rates.

4.2.3.3 Impact on Costs

Working a shorter data collection period would increase costs because of a loss in efficiency. For example, depending on which option is selected, an FI assigned to a segment would need to travel to the same segment multiple times and each time new cases are released in that segment. For remote segments, one or possibly two FIs typically are sent on one trip to complete all the cases within the segment. Releasing cases in these remote areas every few weeks would require multiple trips and would be very costly.

Because of the condensed time frame for working cases in a continuous design option, FIs would need to focus on working only their assigned cases and, therefore, would not have as much flexibility to assist in other regions. Hence, to provide ample coverage, there would be a greater need to hire more TFIs to travel at a moment's notice. Not having the flexibility to schedule trips well in advance when airfares are usually lower could increase costs for the project. It should also be noted that because of the skill set required for TFIs, their hourly wages are higher than average FI wages, and, because of being on travel status, their expenses are typically 8 times that of a regular FI. Setting up TFI trips and managing TFIs while in the field require additional time from the FSs, which would likely result in incurring overtime for them as well.

4.2.3.4 Impact on Staffing

The continuous design option would most likely create some staffing issues, ultimately affecting costs. Currently, most NSDUH FIs work part time and use their downtime to attend training for other research companies, work other part-time jobs, take vacations, and so on. The current quarterly schedule allows time to push and clean up cases, then a short period of downtime to get refreshed and motivated for the next quarter. A tighter data collection schedule with the continuous design would require FIs to work continuously with little to no downtime and would most likely require many FIs to work full time rather than part time. All of these factors may contribute to increased burnout among FIs, possibly resulting in higher turnover rates. As noted previously, many FIs like the flexibility offered by part-time work for a variety of reasons. Recruiting and hiring FIs willing to devote full time to one project would likely be very challenging. It should be noted that some States may require that anyone employed full time must also be offered health benefits, which again would increase costs. On the other hand, offering full-time work and benefits to FIs might make the position more attractive for some. In the FS survey report, FSs suggested that offering health benefits and increasing the pay rate would assist with FI recruitment and retention (Peytcheva & Currivan, 2009). If benefits are offered to full-time FIs, most likely those same benefits would need to be provided to the FSs.

Some advantages in employing full-time field staff versus part-time field staff in terms of data quality and turnover rates may include the following:

- It is more efficient to monitor data quality across a smaller group of full-time FIs.
- FIs employed full time may feel more invested in the study and, because of a steadier income, may be more reluctant to leave the project.

- Some States require an employer to provide subsidized fringe benefits if the employee works, on average, over a certain number of hours per week. NSDUH would abide by these laws, and, therefore, some FIs would likely receive subsidized or free fringe benefits. As a result, these FIs may be reluctant to leave the project.

These advantages are further evidenced in the 2005 to 2008 turnover report (Peytcheva, Currivan, & Cunningham, 2010), where reasons that FIs left the project were examined based on responses to the exit interview. Not enough steady work and lack of benefits were two reasons cited in the report as to why FIs left the NSDUH FI position. On the other hand, in the FI survey report, Peytcheva, Murphy, Cunningham, and Currivan (2009) found that FIs had no clear preference for work distribution throughout the quarter (working stable hours throughout the quarter or working more hours in the beginning of the quarter). These FIs noted that their only preference would have been for the variable work hours to be clearly stated in the job description before they accepted the position.

Some disadvantages in full-time versus part-time employment and the effect it may have on data quality and turnover rates might include the following:

- Many FIs are retired, elderly staff and are limited in how much they can earn because of Social Security benefits. Therefore, a large portion of the current veteran part-time workforce may be lost.
- Requiring FIs to work full time may cause staff to burn out at a higher rate, resulting in higher turnover rates.
- Based on a finding in the FI turnover report (Peytcheva et al., 2010), the most common reason given among staff who leave the project is that they simply cannot work the required hours. So, requiring FIs to work full time would result in replacing a large portion of the current staff with newly hired FIs. Hiring a large number of new FIs could ultimately result in an increase in data quality issues and lower response rates.
- Wage costs would possibly be higher because of having to subsidize or provide fringe benefits for FIs working over a certain number of hours per week on average.

One staffing option to consider is a combination of full-time and part-time FIs. In this scenario, some FIs would work up to 40 hours a week, and part-time FIs would work 20 to 25 hours per week. This would enable staff who can only commit to part-time hours to work less than full-time hours, yet it would also provide them with a steady stream of income. Although offering steady work may be desirable for some FIs, it will take away the opportunity for FIs to use downtime to attend training for other research companies, take vacations, work other part-time jobs, and so on. FIs accustomed to the current NSDUH data collection schedule would need to adjust.

There would also be implications on the FS's job with the continuous design option. By no longer having downtime, FSs would lose a short period at the end of each quarter when motivational management becomes a priority. FSs use this time to conduct FS team meetings, conduct brainstorming and motivational sessions with their teams, train at NTP training sessions, and so on. Motivational sessions conducted at the end of a particularly difficult quarter can help to reenergize an FS's team, which often results in a good start for the next quarter. The time involved in planning motivational activities varies depending on whether the FS plans an in-person team meeting or simply a group conference call. In general, 1 to 2 weeks is an adequate amount of time for FSs to plan and implement motivational management and complete other end-of-quarter duties. If FS teams are in a continuous data collection mode, FSs would have less time to plan for these other activities and would no longer have the flexibility to be out of the office for a week at a time to attend training sessions.

Furthermore, because a higher turnover rate is anticipated due to FI burnout, FSs would have to increase recruiting activities. It is important to note that all newly trained FIs are sent into the field with a mentor, and mentors are typically the best and most productive FIs. Therefore, to continue mentoring newly trained FIs, the best and most productive FIs would need to be taken out of the field to complete this task, which again would negatively impact production and response rates. Also, because the continuous design requires more frequent training sessions, taking the better skilled, higher producing staff out of the field to mentor new FIs would occur more frequently.

Another consideration is the difficulty that might be encountered in staffing and conducting any special field test requests. FSs and FIs would be overloaded with activities in the 2-, 4-, or 6-week design, making it difficult, or even impossible, to use the same FIs and FSs on any special field tests. Thus, it would likely require staffing a field test with new FSs and FIs, which could be problematic. These new personnel would need to go through the entire NTP training and the field test training programs. In addition, new FSs would need to become familiar with the NSDUH case management system (CMS) and any special CMSs needed for the field test. By using staff less experienced with the many challenges on NSDUH, the results of a field test may be skewed.

The current management structure would likely be affected depending on which design option is chosen. Because the 6-week design is somewhat similar to the current data collection period, it is anticipated that the management structure would be similar to the current structure. However, with the 4-week and certainly the 2-week design options, more drastic changes would be expected. For example, in the 2-week design option, fewer FIs would be needed, but most would be working full time. Because of the many logistical challenges noted previously of managing cases being released every 2 weeks, more FSs would be needed to manage the FIs and their cases, thus reducing the FI-to-FS ratio. For the 4-week design, more FIs would be needed than are needed for the 2-week option, but fewer than are needed for the 6-week design option. In the 4-week design, FIs would likely need to work up to 30 hours a week, and again because of the logistics involved in managing cases released every 4 weeks, the FI-to-FS ratio would need to be adjusted. The need to adjust FI-to-FS ratios to prevent overloading FSs is further justified by a finding in the FS survey report (Peytcheva & Currivan, 2009). The report noted that about half of the FSs were satisfied with the number of hours they work in a typical workweek. Among those who were not satisfied with their work hours, most expressed a preference for working fewer hours per week. In addition, the majority of the FSs did not feel that their pay rate was appropriate for the job.

In addition to the logistical challenges and the continuous workflow that needs to be managed, FSs also will be required to manage counting and listing. Thus, the FS workload would need to be carefully assessed to ensure appropriate work-life balances.

Currently, fewer than 10 percent of NSDUH FIs work 30 hours or more per week on average. About 50 percent average between 10 and 19 hours per week. As noted in the data collection simplification and interviewer retention report (DCSIR) (Cunningham et al., 2009), the majority of FIs preferred flexible, part-time work because they care for children, grandchildren, a spouse, or an aging parent; because they have some other family commitments or responsibilities; or because making too much income on NSDUH reduces their retirement income from other sources. In urban areas, FIs are commonly hired who have full-time jobs that provide more steady income and offer fringe benefits and who bring a strong skill set to the FI position. The vast majority of these FIs contribute significantly to NSDUH's success. Therefore, regardless of which design option is chosen, eliminating all part-time FI positions involving a commitment of 15 to 20 hours per week is not recommended. However, to address the problems noted in the FI exit interview among the subset of FIs who need more income and have extensive availability, an approach in which a higher proportion of FIs is working 30 to 40 hours per week on average might decrease the FI turnover rate. Perhaps offering more work and adding fringe benefits for some FIs may be a way to attract and retain a larger group of these qualified FIs who can work more hours per week.

4.2.3.5 Effect on Data Quality

Data quality would be affected because of the short amount of time available to push cases through the verification process and allow ample time to follow up on FIs whose work appears to be questionable. This is especially true for the 2-week design option, but it would also be problematic in the 4-week design option. Field-verifying an FI's work within the 6-week design option would create some challenges, but it would likely be feasible. It is important to note that when field-verifying an FI's work, the best FIs are used. So, taking better FIs out of the field to conduct field verifications would again have an impact on production and response rates.

4.2.3.6 Field Management and Training Structure of the National Survey of Family Growth

To better understand the impact of a continuous design on field operations, the technical reports and publications from several in-person continuous surveys (National Survey of Family Growth [NSFG], Current Employment Statistics [CES] Survey, National Health Interview Survey [NHIS], National Crime Victimization Survey [NCVS], American Community Survey [ACS], and National Health and Nutrition Examination Survey [NHANES]) were reviewed for field management and training models. This information was only found for the NSFG. Table 4.6 provides the field management and training structure for NSDUH and Cycle 7 of the NSFG, which was conducted from 2006 through 2010. Each year of the NSFG consists of four replicate samples, introduced at the beginning of each quarter. The quarter is split into two phases with the second subsampling phase occurring during the last 2 weeks of the quarter. The full data collection period lasts 48 weeks, with a 4-week break for end-of-year holidays and training of new interviewers. The new interviewers are introduced as part of a rotation of the sample's primary sampling units (PSUs) across non-self-representing areas each year. At any one point,

Table 4.6 Field Management and Training Structure of NSDUH and NSFG

Survey	Completed Interviews per Year	Hours per Interview	RSs	FSs	FIs	Average FI Hours per Week	Attrition Rate	New-to-Project FI Training (days)	Veteran FI Training (days)
2010 NSDUH	67,500	6.9	7	40	650	15	19.83%	7	1
NSFG (Cycle 7)	5,000	9.0	1	2	40	30	—	5[a]	—

FI = field interviewer; FS = field supervisor; RS = regional supervisor.

[a] The National Survey of Family Growth (NSFG) required FIs new to interviewing to attend an additional 1.5 days of training.

the sample consists of 25 non-self-representing areas and 8 self-representing areas, with about 38 to 40 interviewers working in Cycle 7. These interviewers were required to work 30 hours per week to complete a total of 5,000 interviews each year. The employment model for Cycle 6 required over 200 interviewers to complete 12,571 interviews in 2002 (Groves, Mosher, Lepkowski, & Kirgis, 2009; Lepkowski, Mosher, Davis, Groves, & Van Hoewyk, 2010).

Kirgis and Lepkowski (2010) reported that selecting the right staff was critical for the Cycle 7 employment model, given that in most PSUs only one interviewer was staffed to keep the design small and controlled. Therefore, the risk of attrition had to be minimized to avoid having unstaffed areas. The changes in recruitment focus between Cycles 6 and 7 resulted in a staff of interviewers who were significantly more likely to have interviewing experience and a higher education and were more likely to say they like approaching the household based on an interviewer characteristic questionnaire completed by the interviewer.

An example of the increase in field efficiency between Cycle 6 and Cycle 7 of the NSFG was provided by Kirgis and Lepkowski (2010) in the distribution of interviewer hours across various tasks. In Cycle 6, about 40 percent of interviewer hours were spent on screening and interviewing. The remainder of the hours were spent on nonproduction activities, such as administration time, travel, and computer problems. In Cycle 7, where FIs were required to work 30 hours per week, about 55 percent of time was spent in screening and interviewing. The hours per interview also decreased from 11.3 hours per interview in Cycle 6 to 9.0 hours per interview in Cycle 7. However, it was noted that other operational refinements to the data collection protocols were put into place for continuous interviewing, which also contributed to the efficiency improvement in Cycle 7. Also, the NSFG response rate declined from 78 percent in Cycle 6 to 76 percent in Cycle 7.

An aspect of the NSFG model that could be applied to the NSDUH continuous design is having an end-of-year break for holidays and training of new interviewers. The NSFG model also emphasizes the need to attract and retain highly qualified interviewers in a continuous design.

4.2.3.7 Field Operations Summary

To summarize, the continuous data collection design would impact the day-to-day field management and would alter the current job expectations of field staff, both FSs and FIs. The most significant changes would be in managing the preparatory work required each time cases are released, keeping staff motivated and engaged in continuous work with little to no downtime, and the logistics involved in managing multiple sample releases.

As noted above, the 2-week design option would be the most challenging option. Some of the more negative aspects of data collection that would be greatly impacted by a 2-week period are as follows:

- not having enough time to adequately negotiate controlled access situations;
- little flexibility in using borrowed field staff to complete work when staff are out unexpectedly;
- little flexibility in borrowing bilingual FIs from adjacent regions;
- relying heavily on increased use of TFIs, which would increase costs;
- not enough time to use unable-to-contact letters and refusal letters effectively;
- compromising NTP mentoring, an important training component (i.e., observing at least one screening and an interview ties up veteran staff from working their own assignments); and
- not enough time to follow up on data quality issues.

The 4-week time period may be more tenable for FIs regarding organization, and it could help them visualize data collection goals better. However, controlled access situations would still be challenging and in some cases impossible in a 4-week period. And, although the 6-week design option would also have its challenges, the additional time would make some of the previously outlined issues less problematic.

Implementing an overlapping data collection period, where there would be a 2-week cleanup period after the data collection deadline, would allow some time to work on challenging cases and controlled access cases. However, the logistical management of two overlapping samples would create some field management challenges. For example, when a new sample is released, FIs may be tempted to work on the fresh cases rather than on some of the more challenging cleanup cases that remain from the previously released sample. Thus, FSs would need to prioritize these aging cleanup cases to ensure FIs put forth the effort in working these cases so response rates do not suffer.

4.2.4 Other Survey Operations

4.2.4.1 Equipment and Case Management Systems

Currently, NSDUH server-side computer systems are configured to support parallel activities associated with two consecutive data collection quarters. This type of overlapping operation is required to support in-field data collection for quarter Q, while simultaneously supporting startup activities (e.g., recruiting, case release, and initial case assignment) for quarter Q+1. The data processing team accommodates this need for parallel operation by maintaining separate relational databases for each quarter. This allows data from overlapping quarters to be processed by running automatic data processing software against the databases for quarter Q, repointing the software at the databases for quarter Q+1, then rerunning the same software a second time. The database repointing and software rerunning all happen under program control and without a need for manual intervention. This mechanism evolved out of necessity because of the need to support activities from two consecutive quarters and because the quantity of data involved precluded holding data from multiple quarters in a single relational database. The performance penalties for doing so were simply too great at the time the current systems were designed and built. An advantage of the quarter-based system design is that the size of the resulting relational databases is now limited, and the maximum size of the quarterly databases is well within the operational capacity of the current relational database servers. Database queries execute quickly, and the CMS Web site performance is excellent. Another important advantage is that the data processing team only needs to maintain a single version of the automatic data processing and Web site source code, which can be easily repointed at the different quarter-specific databases to support data processing for any single quarter.

Transition to a continuous design would dictate some basic reworking of the automatic data processing strategy described previously. The systems would be reconfigured to use relational databases containing information spanning multiple data collection periods. Additional logic would be incorporated into the data processing and Web site software to select appropriate subsets of data, as required by each different operation or Web request. Doing this is not expected to result in problematic performance degradation because relational database performance and capacity have improved greatly over the past several years since the current system was originally designed and implemented. A reworking of this type would, however, be a large and challenging job. It would require significant effort by the software development team. Once the new scheme was in place, it might also require a slightly higher level of effort to maintain. This would depend on the length of the data collection periods and on the frequency with which new sample lines were added. Currently, sample lines are selected and relational databases are preloaded with new lines 4 times per year. This is a manual process, requiring oversight by an experienced database programmer. If the sample preload operation and all the corresponding tasks (QCing of sample lines, generating lead letters, creating hard-copy lists of selected dwelling units, etc.) become more frequent, then the level of ongoing programming support will be increased. Switching to a continuous design would require only minimal changes to the in-field data collection hardware and software configurations. However, if data collection were truly continuous, if data collection periods overlapped, and if the overlaps were allowed to span the end of a year, the question of when to make the yearly updates to the screening and interviewing software would need to be considered. It may be that a small annual break would still be needed in the ongoing cycle of data collection to load new versions of the data collection software and to make sure all of the interviewers started out each new data collection year with a consistent set of software and associated system configurations.

A continuous design with overlapping data collection periods would require many of the same changes as the continuous design option described previously. It is only slightly more complicated in that period indicators would need to be included in all of the data structures so that the Web sites and data processing software could distinguish data from different data collection periods. From the data processing, data management, and software development perspectives, this option is quite similar to the continuous design option.

4.2.4.2 Analysis and Reporting

For both the continuous design and continuous design with overlapping data collection period alternatives, the changes in data collection are not expected to have substantial impact on analysis and reporting activities unless data collection activities extend into the subsequent survey year. For example, analysis and reporting activities for each survey year currently begin in January of the year following data collection, now scheduled for January through December. If data collection activities move into January of the next year, the schedule for analysis and reporting activities would need to be adjusted accordingly.

Current analysis and reporting activities are based on a quarterly data collection design, and most activities begin in January of the year following data collection, which now occurs between January and December. If a continuous data collection period with data collection conducted at 6-week intervals or more frequently were implemented, many activities would not be substantially affected if data collection remains within the January to December time period. There would be no or minimal effect expected for weighting, editing, and imputation. The 6-month preliminary analysis that is now conducted after two quarters could be conducted after three 6-week periods if the sample size was sufficient to yield nonsuppressed estimates for analysis variables of interest. Thus, the continuous design would allow preliminary analyses to be conducted earlier than the current 6-month option yet later than the first quarter alternative option.

Table production and reporting activities would not be affected by a continuous design if data collection were still completed so that data collection periods did not overlap. The one exception might be using data from embedded field tests. Field tests and special studies, such as the Mental Health Surveillance Study (MHSS), may require follow-up interviews. When the data collection period is shortened, less time is available to complete follow-up interviews within the data collection period, which might necessitate follow-up interviews spilling into the next data collection period. Although this would not impact use of the main study data, it would impact use of the follow-up interview data related to the main analysis. For example, in the MHSS, if the follow-up interviews had been completed later than the end of the data collection period, there would have been less time to analyze these data to develop measures of mental illness for use in the main analysis.

If a more frequent data collection period were implemented, with data collection ending in the subsequent data collection period, the major impact would be on the schedule. The weighting, editing, imputation, analysis, and reporting schedule for production of detailed tables and the national findings reports would need to be adjusted to accommodate this change in schedule or the numbers of tables and size and complexity of the reports would need to be decreased. Alternatively, press release dates later than early September would need to be considered if the current requirements for detailed tables and reporting remain. The schedules for State and substate estimates' reports and other analytic studies also would be impacted.

4.2.5 Costs

For the purpose of estimating survey costs, it is assumed that the current 50-State annual design will be preserved for the 2012 and 2013 NSDUHs. Therefore, a continuous sample design would be implemented for subsequent survey years. Table 4.7 shows estimated costs for the 2014 through 2018 NSDUHs for the current design and sample releases every 2, 4, and 6 weeks. By applying a standard increase of 1 percent (an estimate of the increase because of salary increases and inflation) to all costs for each calendar year starting from the 2009 survey year cost-to-complete forecasts, the annual costs have been estimated for the 2014 through 2018

Table 4.7 Comparison of Costs for Continuous NSDUH Designs: 2014-2018

	2014	2015	2016	2017	2018	Total
Current Design						
Estimated Costs	$53,439,032	$51,800,614	$52,318,620	$52,841,806	$53,370,224	$263,770,295
Total Completed Interviews	67,500	67,500	67,500	67,500	67,500	337,500
Average Cost per Interview	$792	$767	$775	$783	$791	$782
Sample Release Every 6 Weeks						
Estimated Costs	$56,500,799	$54,755,952	$55,303,512	$55,856,547	$56,415,113	$278,831,923
Total Completed Interviews	67,500	67,500	67,500	67,500	67,500	337,500
Average Cost per Interview	$837	$811	$819	$828	$836	$826
Sample Release Every 4 Weeks						
Estimated Costs	$57,400,309	$55,698,719	$56,255,706	$56,818,263	$57,386,445	$283,559,442
Total Completed Interviews	67,500	67,500	67,500	67,500	67,500	337,500
Average Cost per Interview	$850	$825	$833	$842	$850	$840
Sample Release Every 2 Weeks						
Estimated Costs	$58,299,818	$56,641,485	$57,207,900	$57,779,979	$58,357,778	$288,286,960
Total Completed Interviews	67,500	67,500	67,500	67,500	67,500	337,500
Average Cost per Interview	$864	$839	$848	$856	$865	$854

NSDUHs. Higher costs are assumed with more frequent sample releases based on these major assumptions:

- More frequent sample releases would place more pressure on FIs to complete work in a short period of time and would not allow as much downtime as a design with less frequent sample releases, which might lead to higher FI turnover. Higher FI turnover would require additional FI recruiting, training, and mentoring resources.
- More frequent sample releases would require additional charges for shipping segment materials to FIs and FIs returning the segment materials to RTI.
- More frequent sample releases would compress the schedule for producing 6-month summary tables, annual summary tables, and the national findings reports following the end of data collection (assuming cases released in June would be completed in July and cases released in December would be completed in January), which would likely increase labor costs.
- More frequent sample releases would require additional sampling and data management labor to prepare, release, and manage case assignments because of repeating the activity more times during a given survey year. Also, overlapping data collection periods (e.g., completing June and July cases in July) would require a more complex set of databases, Web functions, and CMS than are currently used.
- More frequent sample releases would likely be accompanied by shorter data collection periods for each wave. If so, additional travel resources would likely be needed in some areas for borrowed FIs (BFIs) and TFIs to complete work because of less advance notice for making travel arrangements for situations in which a local FI

is temporarily unavailable for personal reasons or because an FI cannot work in an area due to weather obstacles or limited access. Under a design with longer data collection periods, a local FI would have more time to complete the work without outside help, and travel expenses would be generally lower because of more opportunities to take advantage of discounted advance fares.

- It is anticipated that the management structure would be similar to the current structure for the 6-week design. For the 2-week design option, fewer FIs would be needed, but many would likely be working full time. Because of the logistical challenges of managing cases being released every 2 weeks, more FSs would be needed to manage the FIs, thus reducing the FI-to-FS ratio from the current 17:1 to perhaps a 13:1 ratio. For the 4-week design, more FIs would be needed than are needed in the 2-week option, but fewer than are needed for the 6-week design option. In the 4-week design, many FIs would likely need to work up to 30 hours a week, and again because of the logistics involved in managing cases released every 4 weeks, the FI-to-FS ratio would need to be adjusted, perhaps to 15:1.

The continuous sample design assumes a 50 percent overlap in PSUs. In addition, it is assumed that field enumeration rather than address-based sampling (ABS) would be used. The costs in Table 4.7 assume listing 7,200 segments for the 2014 NSDUH. For all other survey years, 3,600 segments would be listed.

4.3 Less Continuous Design Option

Although the focus of this chapter is on describing the pros and cons of a more continuous design, one option the Substance Abuse and Mental Health Services Administration (SAMHSA) may choose to consider is a single, longer data collection period (e.g., conducting the survey over a 6- to 9-month period rather than across four quarterly samples). This model is similar to the data collection period of related surveys, such as Monitoring the Future. There are benefits and challenges to the extension of the data collection period. Providing FSs with more of their sample at once would allow them to more efficiently manage the workload. FSs could reduce the number of trips to the field because they would know where the sample is and help FIs plan accordingly. FSs also could make the best possible use of TFIs by giving them more segments to work while in the field.

The longer time frame also allows FSs and FIs to address issues that are currently hindered by time constraints. For example, when dealing with controlled access situations, FSs and FIs would have more time to correspond with gatekeepers and building managers to gain entry. More time also would be available to wait out field issues. For example, staff could work around holidays and bad weather in regions where adverse weather conditions are a factor and could time data collection at universities during periods when school is in session.

However, this proposed design is not without challenges. The eligibility of respondents would need to be modified to address the longer data collection period. Determining eligibility of residents within a 6-month period becomes more complicated because there is a higher probability that someone may move from one residence to another within a 6-month period as opposed to a 3-month period, increasing the likelihood of a respondent being selected twice.

Students on university campuses also could be an issue because of the high probability that their residence will change between semesters and the summer months.

FSs would need to adapt their goal structures to maintain FI motivation and goals, and strategies would need to be modified to ensure that FIs do not procrastinate. FSs also would have to strategize to ensure that response rates do not diminish because of cases being closed out too early and to make sure that FIs return to refusals within a reasonable time frame.

With the 9-month design, down periods could be scheduled to occur around the holidays, which could potentially be beneficial in retaining staff who typically want to take that time off. Overall, the increase in control that this design affords FSs could allow them to manage their sample more efficiently; however, it would take time to adapt the current protocols and procedures to the new design.

Also with this design, collecting the same amount of data within a shorter time period may result in needing more FIs who are willing to work full time. This may become a challenge because many FIs prefer their part-time status and enjoy having periods of downtime to pursue other interests (see Section 4.2.3.4). Likewise, some FIs may not like having to complete all their work in a 6- or 9-month period and then having another 6 or 3 months with potentially no income. This would be particularly problematic for FIs who like to work on more than one study at a time as it would be difficult to find other studies to work on during these 3- or 6-month periods.

If data collection were completed in a 6-month or 9-month period rather than the current 12-month period, the schedule for the completion of existing analysis and reporting activities would need to be adjusted accordingly, but this change would not have a significant impact on analysis and reporting activities. It would be difficult to complete the existing activities within the calendar year of even the 6-month data collection. However, estimates could be made available earlier than in the current 12-month data collection period.

If a smaller number of cases were collected in the reduced time period, this could result in more suppression of data. Also, data collected in a 6-month or 9-month period would be representative of the reduced time period and not the full year. Thus, some ability to examine seasonality of estimates would be lost.

4.4 A 2-Month Continuous Design Option

This chapter presents some of the pros and cons of NSDUH moving to a more continuous design (2-, 4-, and 6-week data collection periods) and a less continuous design (6- to 9- month data collection period). Another "middle-of-the-road" design option to consider is one that entails six 2-month data collection periods during each survey year. This would allow 2 weeks at the end of each quarter for final cleanup, completion of refresher training for veteran FIs (via a quarterly refresher iLearning course), and conduct of NTP training sessions prior to the start of the next quarter. However, similar to the 2-, 4-, and 6-week options, condensing the data collection period could be problematic for regions that experience adverse weather conditions and for the timely completion of challenging controlled access situations.

Regarding staffing and training, the 2-month option also may result in an increased need to recruit and maintain staffing levels across the six data collection periods. One potential method to mitigate this staffing challenge would be to hold NTP training sessions at the end of each data collection period (for a total of six sessions per year) to train new staff for regions with immediate needs for the upcoming period. Although these NTP training sessions would likely be smaller, conducting additional training sessions to keep staffing levels constant would result in an increase in training costs as compared with the current design.

Other costs would likely be about the same as in the current design. For example, although less FS labor might be required to manage the work during the third month of each quarter, assuming the same number of cases, FI labor per case would likely be similar, too, but travel expenses might be higher for the reasons stated in Section 4.2.5. Also, more FS overtime might be incurred during the first 2 months of the quarter because of having to help FIs with more field problems within a compressed period of time.

4.5 American Community Survey Continuous Design Option

4.5.1 Overview of the American Community Survey Sample Design

Another continuous design option that has not been discussed is modeled after the American Community Survey (ACS). This model is similar to the continuous design with overlapping data collection periods that was discussed in Section 4.2. However, the ACS classifies respondents by month based on the date of data collection. The sample is worked for 3 months. Mail-only data collection is used in the first month of each panel. During the second month of each panel, computer-assisted telephone interviewing (CATI) data collection is used for nonrespondents, but mail responses continue to be accepted. During the third month, a subsample of nonrespondents is followed up by a personal visit using computer-assisted personal interviewing (CAPI); mail responses continue to be accepted. The procedure is illustrated in Table 4.8.

Table 4.8 American Community Survey Data Collection in Three Overlapping Phases

	Month of Data Collection					
Sample Panel	**November**	**December**	**January**	**February**	**March**	**April**
November	Mail	Phone	Visit			
December		Mail	Phone	Visit		
January			Mail	Phone	Visit	
February				Mail	Phone	Visit

Source: Based on Exhibit 7.1, U.S. Census Bureau (2009).

Base weights are computed based on the sampling and subsampling rates. To ensure that each month is correctly represented, an initial variation in monthly sample (VMS) factor is computed to adjust the achieved base weights over three monthly samples back to the base weight that properly represents the current month, as illustrated in Table 4.9.

Table 4.9 Computation of Variation in Monthly Sample Factor

	Month of Data Collection		
	March	**April**	**May**
Base Weight	**100**	**100**	**100**
Total Weight after Three Phases of Sampling			
First Month	55 (March sample)	45 (April sample)	40 (May sample)
Second Month	30 (February sample)	25 (March sample)	30 (April sample)
Third Month	30 (January sample)	25 (February sample)	20 (March sample)
Total	115	95	90
VMS Factor	100 ÷ 115	100 ÷ 95	100 ÷ 90

VMS = variation in monthly sample.

Source: Based on Exhibit 11.5, U.S. Census Bureau (2009).

4.5.2 Application of ACS-Like Continuous Sampling to NSDUH

A continuous design similar to the ACS model could be applied to NSDUH. The potential NSDUH design could involve monthly interview assignments of one new segment each month or 12 segments per year. The average sample size per segment would be 10 to 12 interviews per segment, so that an FI would complete 120 to 144 interviews per year. During each month, each FI would work as many cases as possible in the current month's sample and follow up on pending cases from the previous 2 months' samples. The more productive FIs would have few prior month pending cases. Most of the prior month pending cases should be associated with access problems.

When applied to NSDUH, one benefit of the ACS model is that it would require fewer FIs than are currently used and each FI's workload would be more evenly spread across months. Under favorable field circumstances, most FIs could complete a large proportion of the monthly sample assignment during the first month. On the other hand, weather conditions, road conditions, and access limitations would be more disruptive and might require more use of TFIs. Language barriers (Spanish-only speakers), escort requirements for high-crime areas, FIs' personal circumstances (illness, accidents, death in the family, and unreliable transportation) all take time to recognize and adapt. The need to reassign FIs to adjust for these conditions would reduce the first month's production in unpredictable ways. To cover all of these field circumstances, it may be advisable to plan for 1.25 or 1.5 FIs per segment.

In addition to monitoring the first month's sample, FSs would need to closely monitor the cases pending from prior months to make sure FIs put forth the effort to work them. Managing field operations would certainly be more complex. Control systems for managing the simultaneous fielding of three samples at varying stages of completion would also need to be developed.

Adjustments for achieving annual yield targets by State could be achieved solely by adjusting the monthly releases before fieldwork begins rather than the current procedure of holding back part of each segment's sample and releasing additional lines during each quarterly data collection period. However, because of poor advance size measures, currently assigned segment sample sizes may vary quite widely. This would be more of a problem with a monthly release system, and more creative sharing of assignments and use of TFIs might be required. Because the monthly sample sizes would vary, additional assumptions and weight adjustment procedures would be required.

Additional complications would arise with the first and last month's samples. During the first month of implementing an ACS-like continuous design, there would be no pending cases from the prior 2 months, so the achieved sample would either be smaller than a typical month's yield or the sample sizes would need to be adjusted in advance. In December, high variation in completed monthly interviews is expected to be exacerbated by holiday vacation schedules, for both respondents and FIs. Thus, a smaller sample may need to be released. In January of a continuing year, cases pending from November and December would be administered a questionnaire version consistent with the month of completion rather than the month in which they were sampled.

Defining analysis samples in terms of the date of completion would allow estimates for any period to be defined in terms of successive whole months, including calendar years, quarters, and single months. More variability in the weights due to the monthly variation in achieved sample is expected. Crossing monthly control totals with State and/or other dimensions would increase the variability of the weights. Thus, the overall weighting and weight adjustment methodology would need to be redesigned.

Finally, similar to the 4-week continuous design option (with and without overlapping samples), the ACS-like continuous design would allow no downtime for training, other management activities, and planned vacations. The loss of downtime for these activities further supports the need to have sufficient numbers of FIs working at any given time. In addition, because of the added complexity and monitoring requirements placed on the FSs, a lower ratio of FIs to FSs is advised.

In summary, the ACS continuous design model would provide better coverage of the calendar year and would allow analysts to study seasonal variations in drug use and mental health measures. It would also have the benefit of reducing the number of FIs required to field NSDUH. Although the model would add some additional complexities in managing three overlapping samples and the weighting and analysis procedures, the design could result in a net cost savings.

4.6 Conclusions

Moving to a more continuous design from the current quarterly design of NSDUH would more accurately portray annual drug use if there are peaks and troughs in drug use throughout the year. There is little evidence to suggest seasonal differences in drug use at the national level, but there may be small differences in drug use among subgroups throughout the year. Moving to a more continuous design would have its own consequences. The largest impact of a more continuous design would be on training and field operations. There would be little downtime for training veteran and new FIs, which could overburden FSs and thereby negatively affect response rates. In addition, negative impacts on response rates, costs, staffing, and data quality would be expected because of having shorter data collection periods. A more continuous design could be beneficial for survey planning (i.e., implementing questionnaire changes, responding to major events, and achieving sample targets). However, additional work would be required to prepare more frequent samples. Finally, a more continuous design would have minimal impacts on analysis and reporting unless case assignments extend into the following data collection period (and the following survey year).

Another alternative that SAMHSA may wish to consider is a less continuous design similar to the single 6-month data collection period of the Monitoring the Future survey. A single, longer data collection period allows more time for overcoming challenges, such as controlled access and weather. This design would simplify most other aspects of the survey and would have minimal impact on analysis and reporting. Finally, the data would be representative of the data collection period and not the full year.

A middle-ground solution involves collecting data every 8 weeks. Although this solution mitigates some of the problems with the other continuous design options (e.g., provides 2 weeks at the end of the data collection period for cleanup, training, and other management activities), the compressed time period would still pose problems with overcoming adverse weather conditions and controlled access situations. Otherwise, the 2-month option is the most similar to the current design and provides better coverage of the calendar year.

Finally, a continuous design similar to that of the ACS could be applied to NSDUH. The ACS model releases its sample on a monthly basis and allows 3 months for completion. Rather than counting respondents in the month they are sampled, the ACS design counts respondents in the month of data collection. When applied to NSDUH, the ACS model has the benefit of reducing costs by requiring fewer FIs. However, the model is very complex and provides some challenges for sampling, field management, weighting, and analysis.

5. Biennial Design Assessment

5.1 Introduction

5.1.1 Purpose

Since 1999, the National Survey on Drug Use and Health (NSDUH) has collected interview data from approximately 67,500 persons annually, and the design has provided for State-level estimates in all 50 States and the District of Columbia by requiring minimum sample sizes in each State. Under the current annual design, direct national estimates are produced each year, and trends can be established relatively quickly when either a new baseline is established or new questions or estimates are introduced. The consistency of the NSDUH questionnaire, as well as other data collection and design aspects, allows data to be pooled over several years to produce more reliable estimates, especially for smaller subpopulations. National-level estimates presented in the annual detailed tables and national findings reports are based predominantly on single years of data. Exceptions are made for estimates among rare populations; in these cases, multiple years of data are pooled together to increase sample sizes and enhance precision. Annual State estimates produced using small area estimation (SAE) procedures are based on 2-year moving averages to increase precision for States with smaller samples. State estimates also are produced using direct estimation based on 3 or more years of pooled data, and substate estimates produced using SAE are based on 3 years of data. Many special and ad hoc analyses also utilize multiple years of combined data.

This chapter describes some of the implications of switching NSDUH from an annual survey of 67,500 persons to a biennial survey of 67,500 persons (i.e., collect data from 67,500 persons every other year). In addition, the following two alternatives to a biennial survey are described: (1) an annual survey of 33,750 persons (i.e., half of the current annual sample of 67,500), referred to as the "half-sample design"; and (2) an annual survey of one half of the States one year and the other half the following year, referred to as the "half-State design." Table 5.1 provides a general overview of the four designs. Section 5.2 presents cost estimates for these four designs over the 5-year cycle from 2014 to 2018. Section 5.3 summarizes the advantages, disadvantages, and impact of switching to a biennial design. Section 5.4 and 5.5 provide the pros and cons of the half-sample and half-State design alternatives, respectively. Section 5.6 provides a summary of the conclusions.

5.1.2 Other Biennial Surveys

A search of periodic national surveys similar in scope to NSDUH revealed only a handful of surveys moving from more to less frequent data collection. Most of the surveys identified were longitudinal panel surveys, where the annual sample was composed mostly of units involved in prior survey waves. Here are a few examples of the panel surveys identified:

- The American Housing Survey (AHS) went from annual to biennial data collection beginning in 1981.

Table 5.1 National Survey on Drug Use and Health: Comparison of Biennial Design with the Current Design and Two Alternatives

Design	Description	Sample Representation	National Estimates	State Estimates and National Outcome Measures	Substate Estimates	Survey Planning (questionnaire changes, sample expansion, etc.)	Interviewer and Field Management Staffing	Startup and Shutdown Activities	Ability to Measure Impact of Major Events (e.g., Hurricane Katrina)	Annual Cost
Annual	Current design with n = 67,500 respondents surveyed every year	Representative of the NSDUH target population	Published annually and mostly based on 1 year of data	Published annually and based on 2 consecutive years of data combined	Published every other year and based on 3 consecutive years of data combined	Ability to respond to emerging issues using split or expanded samples	Consistency in staffing, with no impact on data quality	None	Able to respond to major events by conducting special data collection and analyses	Comparable with current costs
Biennial	Current design except n = 67,500 respondents surveyed every other year with no data collection in the "off years"	Representative of the NSDUH target population	Published every 2 years	Published every 2 years and based on data spanning 3 calendar years	Published every 4 years and covering data collected 2 to 6 years earlier (impractical)	Less ability to respond to emerging issues quickly	(a) Change in field management and interviewing staff from one survey year to the next could affect estimates, response rates, and quality; (b) could significantly increase training and recruiting costs	(a) Required for computer equipment and database storage each survey year; (b) some loss of efficiency due to retraining experienced staff or training new staff to replace attrited staff every other year	Severely limited ability to respond to major events	(a) Significant savings from conducting the survey every other year; (b) increased overall cost per interview due to annual startup and shutdown activities; (c) some loss of efficiency resulting from noncontinuous staffing
Annual with Half Sample	Current design with n = 33,750 respondents (one half of current sample) surveyed every year; all States in the sample each year	Representative of the NSDUH target population	Published annually and based mostly on 1 year of data	Published every 2 years and based on 4 consecutive years of data combined	Published every 4 years and covering data collected 2 to 6 years earlier (impractical)	Limited ability to respond to emerging issues	Reduction in staff, with no impact on data quality	None	Limited ability to respond to major events due to smaller sample than current design	(a) Significant savings from reducing the sample size; (b) increased cost per interview due to additional travel to a more dispersed sample and fewer cases to amortize fixed costs

(continued)

Table 5.1 National Survey on Drug Use and Health: Comparison of Biennial Design with the Current Design and Two Alternatives (continued)

Design	Description	Sample Representation	National Estimates	State Estimates and National Outcome Measures	Substate Estimates	Survey Planning (questionnaire changes, sample expansion, etc.)	Interviewer and Field Management Staffing	Startup and Shutdown Activities	Ability to Measure Impact of Major Events (e.g., Hurricane Katrina)	Annual Cost
Annual in Half of States	Survey 33,750 respondents in one half of States each year, so 67,500 respondents surveyed in all States over a 2-year period	Single-year estimates will not be nationally representative	Published every 2 years; however, interpretation of results may differ because not all States are surveyed at the same time	Published every 2 years using data spanning 4 calendar years; comparing results between first half and second half State samples may be difficult because not all States are surveyed at the same time	Published every 4 years and covering data collected 2 to 6 years earlier (impractical)	Less ability to quickly respond to emerging issues	(a) Consistency in field management, but change in field interviewer mix from year to year; (b) change in interviewing staff from one survey year to the next in a given State could affect estimates, response rates, and quality and significantly increase recruiting and training costs	Required for interviewers, computer equipment each year in half of the States	Ability to respond to major events dependent on location of sample in that survey year	(a) Cost savings from reduction in annual sample; (b) significant increase in cost per interview due to startup and shutdown in the active States each year, including recruiting, as well as fewer cases to amortize fixed costs

- The National Longitudinal Survey of Youth (NLSY) went from annual to biennial beginning in 1979.
- The Panel Survey of Income Dynamics (PSID) went from annual to biennial beginning in 1997.

Few presentations or publications described the experiences of these surveys in moving from an annual to a biennial design. Nelson and McGough (1980) presented a number of policy considerations at a 1980 American Statistical Association (ASA) meeting. Their paper focused primarily on the capabilities of the survey design to address key policy issues. For example, the authors cited the important need for annual data on non-new construction additions to the housing stock and inventory losses. They noted that "less frequent information would miss important turning points and hide relationships" (Nelson & McGough, 1980, p. 112). Concerns about biennial versus annual data collection also led the authors to maximize sample size over each 4-year period (i.e., two data collection waves). This paper did not fully address other issues in moving to biennial data collection (e.g., staffing and other operational concerns). In addition, a number of other design changes took place concurrently, such as the use of computer-assisted telephone interviewing (CATI) beginning in 1983. This change and others make it difficult to estimate the effect of changing the frequency of data collection on indicators such as response rates.

One repeated cross-sectional survey with a documented shift from annual to biennial data collection was identified—the General Social Survey (GSS). The GSS went from annual to biennial in 1994, whereupon the biennial data collection actually comprised two surveys. Packaging two surveys into one every other year appeared to be motivated primarily by cost concerns. Entwisle (2007) specifically pointed out substantial savings in training costs and listing expense. An important concern Entwisle (2007) mentioned was that the redesigned GSS is a rotation of core survey items. To maximize the number of questions asked across each wave of respondents, each biennial survey has three ballots. She indicated that this feature makes the GSS data more difficult to use for standard analyses, such as crosstabulations and regression analysis, because the completed interviews do not include the same complete set of responses.

Looking at one indicator of data quality, outcome rates, the switch from annual to biennial data collection did not appear to have a significantly adverse effect on the GSS. Of course, a number of factors affect survey outcomes, such as response rates, so firm conclusions cannot be drawn from these experiences. In the waves following the move to biennial data collection, GSS response rates steadily fell (Davis, Smith, & Marsden, 2009). It should be noted that the 1990s were a period of steadily declining response rates for most surveys. Interestingly, while refusal rates steadily increased for most waves in the late 1990s and 2000s, eligibility rates also declined (Davis et al., 2009). This latter outcome suggests making changes to the protocol, which might further complicate drawing conclusions from these experiences.

Personal communication with a former National Opinion Research Center (NORC) employee who managed the GSS during the transition time provided further context on the experiences following the move to a biennial design. A key challenge of this design change was the major impact on office staff—that is, project staff who do not actually collect the data. Given that GSS field staff moved in and out of projects on a regular basis, he did not recall greater difficulty in recruiting field interviewers (FIs) for the second biennial survey wave. The doubling of the sample size (from about 2,200 to 4,440 norm) was actually helpful in that more resources were available for interviewer recruitment. For office staff, on the other hand, significant problems were created with respect to the continuity of staff and methods. The NORC researcher suggested a key outcome of this factor was cost overrun rather than shortfalls in the completion rate.

The 1988 and 1990 National Household Surveys on Drug Abuse (NHSDAs)[21] also were considered in this research, but these biennial survey years featured significant differences in design and complexity from the current NSDUH in sampling, data collection methods (paper vs. computerized survey instruments), schedule, staffing, and respondent sample size, among other design elements. Therefore, experiences from the 1988 and 1990 NHSDAs are not beneficial for comparison because of these differences.

The experiences of the National Survey of Family Growth (NSFG) or other surveys moving to more frequent data collection do not seem relevant to a biennial design, unless the biennial design included a much larger sample size. The current list of sample options does not include increasing the sample size. If the biennial sample size were to be increased significantly, as the GSS did, similar challenges could be expected, as was the case in 1999 when NHSDA moved to a 50-State design.

5.2 Estimating Survey Costs

For the purpose of estimating survey costs, it is assumed that the current 50-State annual design will be preserved for the 2012 and 2013 NSDUHs. Then a biennial design will be implemented for subsequent survey years. Table 5.2 shows the estimated costs for the 2014 through 2018 NSDUHs for the annual, biennial, and two alternative design options (i.e., half-sample and half-State designs). The annual costs for the 2014 through 2018 NSDUHs have been estimated by applying a standard increase of 1 percent (an estimate of the increase due to salary increases and inflation) to all costs for each calendar year starting from the 2009 survey year cost-to-complete forecasts. Also, for the biennial design, the cost of shutting down operations following the 2014 survey and for starting up and shutting down operations for the 2016 and 2018 surveys has been estimated. For the half-State design alternative, the cost of starting up and shutting down operations in half of the States each year has been estimated.

The annual, biennial, and two alternative designs assume a 50 percent overlap in primary sampling units (PSUs). In addition, it is assumed that field enumeration rather than address-based sampling (ABS) is to be used. The costs in Table 5.2 for the annual and biennial designs assume listing 7,200 segments for the 2014 NSDUH. For all of the other survey years, 3,600 segments will be listed. The half-sample design assumes listing 3,600 segments for the 2014 NSDUH and 1,800 segments in subsequent years. Finally, the half-State design assumes listing 3,600 segments each for the 2014 and 2015 NSDUHs and 1,800 segments each for the 2016 through 2018 survey years.

[21] Prior to 2002, NSDUH was known as the National Household Survey on Drug Abuse (NHSDA).

Table 5.2 Comparison of Costs for Annual and Biennial NSDUH Designs: 2014-2018

	2014 Survey	2015 Survey	2016 Survey	2017 Survey	2018 Survey	Total
ANNUAL DESIGN						
Estimated Costs	$53,439,032	$51,800,614	$52,318,620	$52,841,806	$53,370,224	$263,770,295
Total Completed Interviews	67,500	67,500	67,500	67,500	67,500	337,500
Average Cost per Interview	$792	$767	$775	$783	$791	$782
BIENNIAL DESIGN						
Estimated Costs	$54,063,829		$57,251,976		$59,564,956	$170,880,760
Total Completed Interviews	67,500		67,500		67,500	202,500
Average Cost per Interview	$801		$848		$882	$844
ANNUAL WITH HALF SAMPLE						
Estimated Costs	$38,201,203	$37,250,036	$37,622,536	$37,998,761	$38,378,749	$189,451,286
Total Completed Interviews	33,750	33,750	33,750	33,750	33,750	168,750
Average Cost per Interview	$1,132	$1,104	$1,115	$1,126	$1,137	$1,123
ANNUAL IN HALF OF STATES						
Estimated Costs	$38,201,203	$41,589,556	$42,005,452	$42,425,506	$42,849,761	$207,071,479
Total Completed Interviews	33,750	33,750	33,750	33,750	33,750	168,750
Average Cost per Interview	$1,132	$1,232	$1,245	$1,257	$1,270	$1,227

5.3 Advantages, Disadvantages, and Impact of Switching to a Biennial Design

5.3.1 Survey Costs

Cost savings will be experienced primarily because segment listing and data collection will not occur every year. However, there will be cost inefficiencies in nearly all tasks because of the noncontinuous surveys, resulting in a higher overall cost per interview. These inefficiencies result from annual startup and shutdown costs associated with not conducting the survey every year, including the need to regularly recruit, train, and monitor large numbers of temporary and field staff because it may be difficult to retain staff during the "off year." See Section 5.2 for cost details.

5.3.2 Survey Planning

Switching NSDUH from an annual to a biennial design will reduce the utility of the survey for planning and resource allocation purposes; data will not be available on an annual basis to inform budget decisions or performance measurement initiatives, such as the National Outcome Measures (NOMs) developed by the Substance Abuse and Mental Health Services Administration (SAMHSA). In addition, the ability to respond quickly to directed changes in scope, such as the 2007 Mental Health Feasibility Study, will be greatly diminished, and such changes will likely require an extended time line.

The ability to continuously identify and address new issues associated with obtaining cooperation from respondents will be reduced. If some event that occurs in an off year or between survey years affects respondent attitudes toward the government or the survey (e.g., 9/11, Hurricanes Katrina and Rita, the Deepwater Horizon oil spill), it will not be identified until data collection resumes instead of as the change is occurring. This delay may have negative effects on response rates and comparability of estimates across years.

Question revisions, inserts, and deletions, as well as methodological changes that require testing in a split-sample experiment prior to implementation, will need to be identified about 1 year prior to the experiment and will not be fully incorporated into the survey until 2 years after the split-sample implementation. Thus, such revisions will take about 3 years to be fully implemented. Inability to quickly obtain data for a desired need will diminish the value of NSDUH. On the positive side, this will allow more time and a full year's sample to analyze a split-sample experiment before decisions about survey changes are made.

5.3.3 Survey Operations and Data Quality

A major detriment to switching to a biennial design is that most NSDUH staff will be working only every other year. It will be difficult to retain experienced and talented professional staff who are reassigned to work on other projects or who seek new employment in an off year. This will result in more new staff each year than with the annual data collection design.

In addition, a biennial design will have a major impact on field staffing. Increased costs will result from recruiting and training new listing staff, FIs, and field supervisors (FSs), as well more extensive refresher training for returning veteran staff after an off year and a greater level of quality control work for less experienced staff. In addition to cost implications, there will likely be an impact on data quality because employing inexperienced listers may lead to less efficient listing and more listing errors. Prior methods studies have shown the potential impact of interviewer experience on both response rates and the reporting of sensitive substance use practices (Chromy, Eyerman, Odom, McNeeley, & Hughes, 2005). This may affect trend lines because of possible interviewer effects between 2014 and 2016 and because 2016 will be the first year of data collection following an off year.

Greater detail on the impact of a biennial design on field staffing, training, data quality, and startup and shutdown activities is provided in the next sections.

5.3.3.1 Field Staffing

A biennial design will negatively affect the staffing of both the FS and FI positions in a number of ways. Actual retention rates are difficult to predict; however, it is anticipated that many FIs likely will find other work during the off years and therefore may not be available to work on NSDUH again during the "on" years. Because networking and referrals are a common practice in recruiting and hiring top-performing FIs, it is quite probable that the better-performing FIs will find more suitable work on other projects that may offer more continuity and steady employment.

The recruiting and staffing of experienced FIs largely depends on the surveys actively being conducted in the field at the time, as well as when these other surveys are beginning and ending data collection. For example, if a biennial NSDUH were fielded at the same time as the U.S. Census Bureau's decennial census, recruiting might prove to be more challenging than during other survey years. On the other hand, if NSDUH were fielded the year after a decennial census, there might be more experienced Census Bureau interviewers available for work.

Because the current NSDUH design offers consistent employment, a large number of FIs have been successfully retained who have worked on this study for many years and have made a career out of working on NSDUH. It is anticipated that with a biennial design, many of the more experienced and more loyal FIs will choose to retire rather than continue working on the project. Many of these same committed FIs are also some of NSDUH's best refusal converters. Losing some of the strongest refusal converters to retirement or to other studies or jobs (because work on NSDUH will no longer offer job stability) certainly will have a negative effect on response rates. Moreover, by not being able to offer consistent employment, NSDUH ultimately will attract FIs who frequently switch studies; thus, NSDUH will become their "on-the-side" job as opposed to their main source of part-time income.

Because increased attrition of experienced and loyal staff is anticipated, more FIs will be needed to complete the work, likely resulting in the need to recruit and hire more inexperienced FIs than under the current annual design. In addition to increasing recruiting and hiring costs, inexperienced FIs also require additional training and mentoring, resulting in a significant increase in training time for in-house staff, FSs, and experienced FIs used for mentoring. After training, FSs will need to invest a great deal of time working with these newly hired FIs to ensure their understanding of NSDUH's procedures and coaching them on working efficiently. Thus, requiring FSs to hire large numbers of new staff every other year will be very taxing on the FSs' time. Current experience shows that it typically takes about 2 full quarters to get a new FI fully trained and proficient on the many aspects of the FI job and NSDUH procedures. With a large number of newly hired FIs requiring more time from the FSs, more FSs will be needed to train and mentor new FIs who will be brought in each survey year. It is important to keep in mind that even experienced, returning FIs likely will get "rusty" with an entire year off the project and will require additional training and FS time to rebuild their skills.

Because NSDUH has offered consistent work, job satisfaction, and a steady paycheck, NSDUH has the added benefit of developing a very loyal, dedicated team of FSs who have acquired a wealth of experience and expertise on the study. Under a biennial design, FSs will be forced to leave the project to seek other employment opportunities. The better-performing FSs will certainly be recruited by other studies during the inactive years. Thus, new FSs will need to be hired and trained, which will require increased time investment by in-house staff.

Another concern is that NSDUH's strong team concept and cohesiveness will be diminished. NSDUH's current FSs have worked extremely hard over the years to build strong teams by developing the skills of their FIs, nurturing a team spirit, and working together to reach project goals. The turnover associated with a biennial design certainly will damage team cohesiveness. In the field, continuity will be lost with controlled access situations where staff have made progress with different State, county, and university representatives and building coordinators, possibly harming the project's ability to gain entry year after year.

In summary, all of these issues will have an initial negative impact on response rates and staffing. Through the consistent work offered with the annual design, FI and FS loyalty has been achieved and NSDUH's reputation has been built. The staffing makeup will possibly change from predominantly strong veteran FIs (who use NSDUH as a primary source of income) to predominantly "in-and-out" FIs who often switch studies or work on multiple projects at a time. Therefore, loyalty to NSDUH will have to be achieved through other means to attract staff back after a year off, such as increased pay or counting-and-listing work during off years. In addition, the number of FIs working at any given time will likely need to be increased because NSDUH might become a lower priority job to many FIs, who will devote less time to the project than the average FI does under the current design.

5.3.3.2 Training

As mentioned above, a biennial design will pose significant challenges in retaining experienced FIs and FSs, who more than likely will seek new employment in the off years. Because it will have been at least a year since data collection was last conducted, returning veteran FIs and FSs will need extensive refresher training to ensure familiarity with and understanding of all project procedures, in addition to reviewing any updates or changes for the upcoming survey year. The refresher training will include a detailed review of the FI manuals and completion of several iLearning courses at home in preparation for an in-person training session (to be conducted in the month preceding, or at the start of, data collection). The in-person training will be scheduled to last approximately 3 days and will comprise an abbreviated review of all of the topics covered during New-to-Project (NTP) FI training, including mock exercises and a modified certification process to ensure each FI's compliance with all project protocols and procedures. All returning FSs also will complete the at-home refresher training tasks and serve as trainers at the in-person training sessions.

All new FIs and FSs will be required to attend NTP FI training and fulfill all training requirements, including passing certification, to successfully graduate. The NTP FI training program will remain very similar to the current NTP sessions conducted on NSDUH, incorporating any changes as a result of a biennial design. However, training a very large number of new FIs to ensure NSDUH is fully staffed prior to or at the start of data collection will be very challenging and will require additional staff, multiple training locations, and other resources from outside the project to assist with this effort. Additionally, after NTP FI training, FSs will have a large number of new FIs to mentor and manage, which will require a great amount of time and effort to provide oversight and guidance, as mentioned in the previous section. Finally, as with the current NSDUH design, NTP FI training sessions will be held as needed in March, June, and September to maintain staffing levels.

The need for a longer in-person refresher training program for veteran FIs and larger NTP FI training sessions to train new FIs at the start of data collection and throughout the year will result in a significant increase in training costs as compared with the current NSDUH design.

5.3.3.3 Data Quality

A biennial design will mean having a wave of new FIs at the beginning of the survey year, which may negatively affect data quality. As discussed in Section 5.3.3.1, a substantial amount of time is needed to teach new FIs the importance of collecting quality data and the potential implications for failing to follow proper procedures. At the beginning of each biennial survey year, a jump in data quality errors can be expected, necessitating a good deal of time to be spent training and retraining new FIs.

Some data quality issues take time to reveal themselves (e.g., patterns indicating shortcutting or falsification). If a large group of new FIs were to be hired at once, a spike in data quality issues can be expected as the data collection survey year progresses. Additionally, data quality staff also take time to hone and develop their skills and could potentially lose some of their skill set in an off year. Moreover, some of the data quality staff might not be able to return to the study every other year, which will require identifying and training inexperienced data quality staff.

Over the years, superior data quality has been established based on a system that works most effectively with the current study design. If this were amended to a biennial design, it is suspected that an immediate decline in data quality will occur that can be rebuilt as processes are adapted. An initial remedy to this situation will be to increase verification on all FIs at the beginning of a biennial survey year to identify patterns as quickly as possible. To keep up with the monitoring of individual FI data quality, having more data quality staff working to identify potential problems is also recommended. Increasing the percentage of verification calls completed for all FIs will require hiring, training, and managing more telephone interviewers, which also will increase costs.

5.3.3.4 Startup and Shutdown Activities

Additional time and cost considerations will need to be built in for startup and shutdown activities for computer equipment and database storage. All project computer equipment will have to be returned, checked, cleared of files, inventoried, and stored every biennial survey year. This process will need to be repeated in reverse order as equipment is taken out of storage and then prepared and assigned to field staff for each successive biennial survey.

For a biennial design, replacing the equipment every 5 years is recommended, as is done with the current design. Although the equipment might receive less physical wear and tear, technological obsolescence will still be the major issue. Operating systems, software development tools, antivirus programs, and network infrastructure will all change significantly in a 5-year time span. Also, spare parts become increasingly difficult to obtain. An attempt to stretch the equipment purchase cycle beyond 5 years will introduce additional risk and additional costs associated with maintaining software for obsolete legacy platforms. It is expected that the effect of inactivity on the equipment every other year will be relatively minor. Given the downtime, cycling the charges on laptop batteries at least one time during an inactive year is suggested. At a minimum, the batteries from the laptops will be removed while they are warehoused.

Under a biennial design, programming staff will have to focus on keeping all systems operational when starting a new survey year, ensuring the systems remain operational during the survey year, then shutting down systems at the conclusion of the survey year, leaving little time for developing enhancements that improve survey quality and efficiency. Project databases will need to be archived to avoid continually incurring significant computer storage charges, which may result in lengthening the response time for ad hoc reports and analyses. Additionally, there will be some loss of efficiency due to having to retrain experienced staff or train new staff to replace attrited staff every other year.

5.3.4 Impact on Estimates

Estimates produced using NSDUH data may be affected by the switch from an annual to a biennial data collection for several reasons. As described in Section 5.3.3, the levels of interviewer experience (resulting from considerable restaffing of FIs and field management) can affect response rates and reporting of sensitive behaviors, which might affect trends between the end of the current design and the first two data collection years of a biennial design. Additionally, many estimates at the State or substate level or among rare populations at the national level are computed using 2 years, 3 years, or sometimes more years of data pooled together. Estimates using 2 survey years of pooled data can still be computed, but these estimates will be less timely because it will take 3 calendar years to produce an estimate, 5 calendar years to compute measures of change based on 2-year moving averages, and 7 calendar years to compute measures of change based on independent 2-year samples. Estimates computed using 3 or more survey years of pooled data may no longer be practical.

5.3.4.1 Single-Year Estimates

With a few exceptions, national-level estimates are based on single years of data. For this reason, there should be minimal impact from a biennial design on the estimates and standard errors (SEs), although the ability to make direct annual estimates for off years will be lost, and detailed tables, national findings reports, and public use data files will be produced only every other year. Additionally, the utility of the survey to monitor trends in substance use and mental health problems will be diminished because the information will be less timely; changes in the use of specific drugs or drug combinations or among subpopulations may not be apparent for 2 years. Also, because of the new "baseline" that may be required with a NSDUH redesign, more time will elapse before any trends can be confirmed.

5.3.4.2 Pooled-Year Estimates

Analyses of small populations or rare occurrences that require pooling of data across 2 or more years of NSDUH data will be compromised because data will be available only every 2 years rather than every year. If these estimates continue to be based on pooled data, then they will be less timely by combining years that cover a larger span of time; using only a single year of data risks having fewer reliable estimates and more suppression.

Annual State estimates produced using SAE procedures require 2 years of pooled data to increase precision for States with smaller samples. Substate estimates produced using SAE require 3 years of data. Tables 5.3 and 5.4 display the publishing schedule for the State and substate SAE reports, respectively, under the different designs.

Table 5.3 Publishing Schedule for State Small Area Estimation Reports, by Survey Year

State Reports	2014 and 2015	2014 and 2016	2014 to 2017	2015 and 2016	2016 and 2018	2016 to 2019
Annual Design	March 2017			March 2018		
Biennial Design		March 2018			March 2020	
Annual with Half Sample			March 2019			March 2021
Annual in Half of States			March 2019			March 2021

Table 5.4 Publishing Schedule for Substate Small Area Estimation Reports, by Survey Year

Substate Reports	2014, 2015, and 2016	2014, 2016, and 2018	2014 to 2019	2016, 2017, and 2018	2018, 2020, and 2022	2018 to 2023
Annual Design	July 2018			July 2020		
Biennial Design		July 2020			July 2024	
Annual with Half Sample			July 2021			July 2025
Annual in Half of States			July 2021			July 2025

Moving averages in the State SAE reports are each based on 2 years of pooled survey data with 1 year in common (e.g., [2003 + 2004] ÷ 2 vs. [2004 + 2005] ÷ 2). With data collection occurring every other year in a biennial design, State trend measures will span 5 calendar years (i.e., [2014 + 2016] ÷ 2 vs. [2016 + 2018] ÷ 2). With this option, 5 calendar years are needed before change measures can be computed using 3 survey years, which diminishes the value of change estimates for States. In addition, assuming a new baseline in 2014, measures of change will not be available until the second State SAE report is published (i.e., in March 2020 according to Table 5.3). For substate estimates requiring 3 years of pooled data, it will take 5 years to get a single estimate, then an additional 4 years before change measures can be computed. For this reason, trends for substate estimates no longer will be practical because they will be less timely and their value to each State will be diminished.

State estimates also are produced using direct estimation based on 2 or more years of pooled data. Table 5.5 shows direct State estimates for Ohio—a large State with midrange prevalence rates for select estimates—and compares annual estimates with assumed biennial estimates pooling both 2 and 3 years of data. Past month use of tobacco products, past month use of cigarettes, and illicit drug and alcohol dependence or abuse all had significantly higher estimates for 3-year pooled estimates under the assumed biennial design compared with the annual design. Tables 5.6 and 5.7 show these same comparisons for Maryland, a small State with low prevalence rates, and Rhode Island, a small State with high prevalence rates. Neither of these States had any significant differences between 3-year pooled estimates computed under the two designs.

Table 5.5 Comparison of Estimates for Ohio (Large State) for Annual and Biennial Designs, by Number of Years of Pooled Data: Percentages and Standard Errors of Percentages

Estimate	2-Year Pooled Estimates		3-Year Pooled Estimates	
	Biennial (2004-2006)	Annual (2005-2006)	Biennial (2002-2004-2006)	Annual (2004-2005-2006)
PAST MONTH USE				
Illicit Drugs	7.8 (0.43)	7.8 (0.45)	7.9 (0.32)	7.8 (0.36)
Marijuana	6.2 (0.37)	6.1 (0.39)	6.4 (0.30)	6.1 (0.31)
Cocaine	1.0 (0.18)	1.0 (0.17)	1.0 (0.14)	1.0 (0.14)
Nonmedical Use of Psychotherapeutics	2.3 (0.19)	2.3 (0.20)	2.3 (0.15)	2.3 (0.16)
Tobacco Products	33.6 (0.85)	33.5 (0.94)	34.7[y] (0.71)	33.5 (0.73)
Cigarettes	28.6 (0.84)	28.3 (0.92)	29.5[y] (0.67)	28.3 (0.72)
Alcohol	50.8 (1.04)	51.6 (1.15)	51.0 (0.91)	51.1 (0.92)
Binge Drinking	24.2 (0.75)	24.5 (0.78)	24.2 (0.63)	23.9 (0.63)
Heavy Drinking	7.7 (0.44)	7.5 (0.41)	7.9 (0.38)	7.5 (0.34)
DEPENDENCE/ABUSE				
Illicit Drugs	3.2 (0.29)	3.1 (0.29)	3.3 (0.24)	3.0 (0.23)
Alcohol	7.2 (0.43)	7.8 (0.45)	7.5 (0.35)	7.4 (0.36)
Illicit Drugs *or* Alcohol	9.1 (0.51)	9.7 (0.55)	9.4 (0.40)	9.2 (0.43)
Illicit Drugs *and* Alcohol	1.3 (0.14)	1.2 (0.14)	1.4[y] (0.13)	1.2 (0.11)

*Low precision; no estimate reported.

[a] Difference between 2004-2006 and 2005-2006 estimates is statistically significant at the 0.05 level.

[b] Difference between 2004-2006 and 2005-2006 estimates is statistically significant at the 0.01 level.

[y] Difference between 2002-2004-2006 and 2004-2005-2006 estimates is statistically significant at the 0.05 level.

[z] Difference between 2002-2004-2006 and 2004-2005-2006 estimates is statistically significant at the 0.01 level.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2002, 2004, 2005, and 2006.

Currently, national-level estimates presented in the detailed tables based on pooled estimates are those among 15- to 44-year-old women by pregnancy status (2 years of data combined) and those for detailed reasons for not receiving substance use treatment among persons who needed but did not receive and felt the need for treatment (4 years of data combined). Table 5.8 shows comparisons between pooled estimates of selected substance use by pregnancy status using the current annual design (i.e., adjacent years of data are combined) and single-year and pooled estimates created under the assumption of a biennial design (i.e., excluding data collected in odd years). There is only one significant difference when comparing biennial pooled data with annual pooled data for "previous year" prevalence rates (2002-2004 vs. 2003-2004), and no differences when comparing single-year biennial rates with annual rates for the previous year. There are also no differences comparing either the biennial alternative with the annual pooled estimate for "current year" estimates. The SEs also look comparable.

Table 5.6 Comparison of Estimates for Maryland (Small State, Low Prevalence) for Annual and Biennial Designs, by Number of Years of Pooled Data: Percentages and Standard Errors of Percentages

Estimate	2-Year Pooled Estimates		3-Year Pooled Estimates	
	Biennial (2004-2006)	Annual (2005-2006)	Biennial (2002-2004-2006)	Annual (2004-2005-2006)
PAST MONTH USE				
Illicit Drugs	5.4 (0.60)	5.6 (0.74)	6.2 (0.61)	5.9 (0.59)
Marijuana	3.8 (0.38)	4.2 (0.57)	4.2 (0.39)	4.3 (0.43)
Cocaine	0.8 (0.26)	0.9 (0.33)	0.7 (0.13)	1.1 (0.28)
Nonmedical Use of Psychotherapeutics	2.0 (0.36)	1.6 (0.30)	2.5 (0.47)	1.8 (0.27)
Tobacco Products	26.6 (1.57)	25.3 (1.71)	26.7 (1.26)	26.2 (1.37)
Cigarettes	23.7 (1.62)	22.3 (1.76)	23.8 (1.30)	23.1 (1.43)
Alcohol	51.0 (2.53)	52.9 (2.29)	53.7 (1.62)	52.0 (2.05)
Binge Drinking	19.0 (1.21)	19.8 (1.46)	20.4 (1.19)	19.5 (1.09)
Heavy Drinking	5.3 (0.62)	5.8 (0.74)	5.8 (0.71)	5.6 (0.57)
DEPENDENCE/ABUSE				
Illicit Drugs	2.0 (0.34)	2.3 (0.33)	2.1 (0.33)	2.2 (0.27)
Alcohol	5.8 (0.87)	6.4 (0.77)	6.1 (0.64)	6.3 (0.71)
Illicit Drugs *or* Alcohol	7.2 (0.88)	7.7 (0.87)	7.6 (0.64)	7.6 (0.74)
Illicit Drugs *and* Alcohol	0.6[a] (0.14)	1.0 (0.19)	0.7 (0.11)	0.9 (0.13)

*Low precision; no estimate reported.

[a] Difference between 2004-2006 and 2005-2006 estimates is statistically significant at the 0.05 level.

[b] Difference between 2004-2006 and 2005-2006 estimates is statistically significant at the 0.01 level.

[y] Difference between 2002-2004-2006 and 2004-2005-2006 estimates is statistically significant at the 0.05 level.

[z] Difference between 2002-2004-2006 and 2004-2005-2006 estimates is statistically significant at the 0.01 level.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2002, 2004, 2005, and 2006.

In general, the results presented in Tables 5.5 to 5.8 suggest that there will be minimal differences in estimates under a biennial design. It should be noted that these results are not necessarily representative of what will be seen under a true biennial design. First, the biennial estimates presented here were calculated using nonconsecutive NSDUH years, which ignores the 50 percent overlap in PSUs that will be present in an actual biennial design. In addition, based on prior experience, having a higher proportion of new FIs is expected to increase the estimates and their SEs.

Table 5.7 Comparison of Estimates for Rhode Island (Small State, High Prevalence) for Annual and Biennial Designs, by Number of Years of Pooled Data: Percentages and Standard Errors of Percentages

Estimate	2-Year Pooled Estimates Biennial (2004-2006)	2-Year Pooled Estimates Annual (2005-2006)	3-Year Pooled Estimates Biennial (2002-2004-2006)	3-Year Pooled Estimates Annual (2004-2005-2006)
PAST MONTH USE				
Illicit Drugs	12.3 (1.37)	11.3 (1.09)	11.7 (1.09)	11.4 (0.99)
Marijuana	9.9 (1.12)	9.4 (0.94)	9.8 (0.92)	9.4 (0.83)
Cocaine	1.6 (0.41)	1.6 (0.33)	1.6 (0.35)	1.4 (0.29)
Nonmedical Use of Psychotherapeutics	2.8 (0.46)	2.4 (0.37)	2.9 (0.39)	2.6 (0.33)
Tobacco Products	30.3 (1.81)	28.6 (1.83)	30.3 (1.23)	29.0 (1.51)
Cigarettes	25.9 (1.73)	24.2 (1.46)	26.1 (1.07)	24.8 (1.33)
Alcohol	61.3 (2.46)	61.4 (2.27)	59.5 (1.86)	61.2 (1.92)
Binge Drinking	30.0 (2.30)	27.8 (1.78)	29.3 (1.67)	28.9 (1.75)
Heavy Drinking	9.9 (1.28)	10.8 (1.20)	9.5 (0.89)	10.1 (1.06)
DEPENDENCE/ABUSE				
Illicit Drugs	4.4 (0.65)	4.2 (0.57)	4.6 (0.62)	4.2 (0.49)
Alcohol	10.0 (1.14)	9.6 (1.19)	9.9 (1.03)	9.3 (0.94)
Illicit Drugs *or* Alcohol	12.0 (1.20)	11.8 (1.26)	11.9 (1.14)	11.3 (1.00)
Illicit Drugs *and* Alcohol	2.4 (0.51)	2.0 (0.42)	2.6 (0.43)	2.1 (0.37)

*Low precision; no estimate reported.

[a] Difference between 2004-2006 and 2005-2006 estimates is statistically significant at the 0.05 level.

[b] Difference between 2004-2006 and 2005-2006 estimates is statistically significant at the 0.01 level.

[y] Difference between 2002-2004-2006 and 2004-2005-2006 estimates is statistically significant at the 0.05 level.

[z] Difference between 2002-2004-2006 and 2004-2005-2006 estimates is statistically significant at the 0.01 level.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2002, 2004, 2005, and 2006.

Table 5.8 Comparison of National Pregnancy Estimates Requiring Pooled Years of Data among Females Aged 15 to 44, by Number of Years of Pooled Data: Percentages and Standard Errors of Percentages

Pregnancy Status and Past Month Substance Use	Biennial Estimate Alternatives				Annual Estimates	
	Single-Year Data		2-Year Pooled Data			
	2004	2006	2002-2004	2004-2006	2003-2004	2005-2006
PREGNANT						
Illicit Drugs	4.0 (0.78)	4.2 (0.74)	3.6 (0.44)	4.1 (0.54)	4.6 (0.60)	4.0 (0.49)
Marijuana	3.2 (0.72)	3.5 (0.72)	3.0 (0.40)	3.3 (0.51)	3.6 (0.51)	3.0 (0.44)
Cocaine	0.2 (0.12)	0.2 (0.10)	0.2 (0.11)	0.2 (0.08)	0.3 (0.11)	0.2 (0.09)
Nonmedical Use of Psychotherapeutics	1.2 (0.37)	1.0 (0.26)	1.0 (0.23)	1.1 (0.23)	1.4 (0.32)	1.3 (0.26)
Tobacco Products	17.8 (1.76)	17.3 (1.80)	17.9 (1.20)	17.5 (1.26)	18.7 (1.24)	16.8 (1.18)
Cigarettes	17.4 (1.75)	17.2 (1.80)	17.3 (1.19)	17.3 (1.26)	18.0 (1.22)	16.5 (1.18)
Alcohol	11.9 (1.67)	11.4 (1.65)	10.6 (1.05)	11.7 (1.17)	11.2 (1.12)	11.8 (1.15)
Binge Drinking	4.0 (0.82)	2.2 (0.54)	3.6 (0.52)	3.1 (0.49)	4.5 (0.63)	2.9 (0.47)
Heavy Drinking	0.3 (0.16)	0.2 (0.13)	0.5 (0.19)	0.3 (0.10)	0.5 (0.15)	0.7 (0.25)
NOT PREGNANT						
Illicit Drugs	9.8 (0.31)	10.1 (0.30)	10.0 (0.22)	10.0 (0.22)	10.2 (0.22)	10.0 (0.21)
Marijuana	7.2 (0.27)	7.0 (0.24)	7.3 (0.19)	7.1 (0.18)	7.5 (0.18)	7.0 (0.17)
Cocaine	1.0 (0.10)	1.2 (0.11)	1.0 (0.07)	1.1 (0.08)	1.0 (0.07)	1.2 (0.08)
Nonmedical Use of Psychotherapeutics	3.9 (0.19)	4.1 (0.19)	3.9 (0.14)	4.0 (0.14)	3.9 (0.14)	4.0 (0.14)
Tobacco Products	30.7 (0.51)	30.8 (0.52)	31.4 (0.36)	30.8 (0.36)	31.0 (0.39)	30.7 (0.39)
Cigarettes	29.6 (0.50)	29.5 (0.51)	30.4 (0.35)	29.6 (0.36)	30.0 (0.38)	29.5 (0.38)
Alcohol	53.1 (0.54)	52.8 (0.61)	53.3 (0.38)	52.9 (0.41)	52.8 (0.39)	53.0 (0.42)
Binge Drinking	23.5 (0.42)	24.1 (0.49)	23.4 (0.32)	23.8 (0.32)	23.3 (0.31)	23.6 (0.34)
Heavy Drinking	5.4 (0.21)	5.6 (0.23)	5.2[a] (0.15)	5.5 (0.15)	5.6 (0.15)	5.4 (0.16)

*Low precision; no estimate reported.

[a] Difference between estimate and 2003-2004 estimate is statistically significant at the 0.05 level.

[b] Difference between estimate and 2003-2004 estimate is statistically significant at the 0.01 level.

[y] Difference between estimate and 2005-2006 estimate is statistically significant at the 0.05 level.

[z] Difference between estimate and 2005-2006 estimate is statistically significant at the 0.01 level.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2002, 2003, 2004, 2005, and 2006.

5.3.5 Other Analyses

Other analyses using NSDUH data include NSDUH's ad hoc requests and special reports, as well as the NOMs for SAMHSA's Center for Substance Abuse Prevention (CSAP) and Center for Substance Abuse Treatment (CSAT). A majority of these other analyses are conducted on NSDUH data at the State level or among small populations; thus, multiple years of data are pooled together.

Switching to a biennial design will lead to fewer available and less timely estimates. Less data accumulation over time may reduce some opportunities for special studies and special analyses that benefit from combining data across years. In the 2008 NSDUH special request cycle, the States of Hawaii, Maine, Pennsylvania, and Texas and the District of Columbia all requested updates to previous ad hoc analyses in order to pool more data together (e.g., update 2003-2006 combined estimates to 2003-2007 estimates) or to create more current, pooled information (e.g., update 2003-2006 estimates to 2004-2007 estimates). Updates to previous ad hoc requests also were made by the Office of National Drug Control Policy (ONDCP), New Mexico State University, and *Reason Magazine*. Additionally, single-year updates to source tables for the *Surgeon General's Report on Youth Tobacco Use and the Underage Drinking Report to Congress* have been provided annually starting with the 2006 NSDUH special request cycle. In addition to ad hoc analyses being less timely, emerging trends will be more difficult to identify in a timely manner. Finally, use of a biennial design may severely limit the ability to conduct special data collection and analyses of the impact of special national events and disasters, such as 9/11, Hurricanes Katrina and Rita, and the Deepwater Horizon oil spill. See Section 5.3.4.2 for more details on the impacts on pooled estimates.

Methodological studies also are routinely conducted for NSDUH. Conducting NSDUH every other year will allow more time, and a full year's sample, to analyze split-sample experiments before decisions need to be made about how to proceed with future surveys. However, the downside of this is that methodological changes to the survey suggested by methods studies may not be able to be implemented for a year or two, at which point the findings may not be fully relevant or another set of issues may have arisen. Additionally, it will be more difficult to isolate the effect of methodological changes to the survey. The real change in estimates between, for example, 2014 and 2016 will be more difficult to control for with no data available from 2015. Also, breaks in data collection may result in methodological anomalies, such as "new interviewer" effects or decreases in the response rate that may affect the quality of the estimates.

5.3.6 Other Considerations

A final consideration unrelated to the cost of conducting NSDUH or the quality of the data is that switching from an annual to a biennial survey may send a signal that estimates obtained from NSDUH are less important than they once were or that the substance use problem is not as urgent.

5.4 Advantages, Disadvantages, and Impact of Switching to a Half-Sample Design

A half-sample design preserves the current 50-State annual design for all years but yields a smaller sample size per year (n = 33,750). The reduction is achieved by collapsing pairs of nearby State sampling (SS) regions to form half of the current number of SS regions in each State and in the Nation as a whole. This design will require listing 1,800 new segments each year except 2014. In 2014, 3,600 segments will be listed. Segment listings will remain fairly fresh by being used within 2 years of being listed, as is done in the current design.

5.4.1 Survey Costs

Reducing the annual sample from 67,500 to 33,750 persons will result in large cost savings, primarily because of the smaller number of completed interviews planned annually. Compared with the current design, however, a half-sample design will result in a higher cost per interview (see Section 5.2). The segments in each State will be more dispersed across fewer and geographically larger SS regions, which will increase the number of long-distance trips by interviewers, with some requiring an overnight stay and meal per diem. Also, there will be fewer cases over which to amortize fixed costs.

5.4.2 Survey Planning

Reducing the annual sample by one half will diminish the utility of NSDUH for planning and resource-allocation purposes; reliable data on some measures will not be available on an annual basis to inform budget decisions or performance measurement initiatives, such as SAMHSA's NOMs. For some measures, the ability to make direct annual and national estimates will be retained. Thus, reports with national findings will still be published every year.

5.4.3 Survey Operations and Data Quality

Although a half-sample design will result in a reduction in project field staff, there will be more stability in professional and technical staff than under a biennial design. The team of professional and technical staff will remain relatively consistent from year to year. Staff turnover can be handled in an ongoing manner that promotes smoother transition and shorter learning curves for new staff. Greater detail on the impact on field staffing, training, and data quality is provided in the Sections 5.4.3.1 through 5.4.3.3.

5.4.3.1 Field Staffing

In a half-sample design alternative, the majority of the staffing issues associated with a biennial design will be eliminated. With a reduction in needed field staff, the best FSs and FIs can be retained, thus continuing to improve efficiency in the field. However, because the sample will be more spread out, good staff may be lost because of lack of work in their area. Additionally, new FIs will need to be hired to fill these positions and to staff areas without local FIs or veteran FIs willing to travel to complete the work. Because a half-sample design will be continuous and require fewer FIs, recruiting challenges are not expected given that NSDUH FSs have been successful in recruiting replacement FIs under the current annual design, even during the U.S. Census Bureau's decennial census years.

5.4.3.2 Training

A reduction of the sample size for this design will result in a smaller field staff required for data collection. However, the design of the veteran and new FI training programs will remain similar to the current NSDUH design because data collection and staffing will be continuous.

All veteran FIs and FSs will complete refresher training consisting of a review of the FI manuals and completion of several iLearning courses at home, followed by a 1-day in-person FS team meeting in January. New FIs and FSs will be required to attend an NTP FI training session and fulfill all requirements, including passing certification, to successfully graduate from training.

The largest difference with a half-sample design will be the reduction in the number of staff needing training (both new and veteran FIs and FSs), which will result in decreased training costs as compared with the biennial, half-State, and current NSDUH designs.

5.4.3.3 Data Quality

There are no data quality concerns with a half-sample design. Data quality can actually be improved by allowing the best-performing FIs to be retained. Reducing the staff to only the top-performing FIs will likely improve field costs, efficiency, data quality, and costs for recruiting.

5.4.4 Impact on Estimates

Under a half-sample design, annual direct national estimates will be available every year, but precision will be reduced for many statistics. Analyses of small populations or rare occurrences that require pooling across 2 or more years of NSDUH data will require pooling over twice as many years to maintain equivalent precision.

5.4.4.1 Single-Year Estimates

Under the current design, national-level estimates are based on single years of data (with a few exceptions). Under a half-sample design, the ability to produce annual national-level estimates will be maintained; however, the precision of these estimates will be reduced. SEs of national-level estimates are projected to increase by a multiplicative factor of 1.41. Detailed tables and national findings reports will continue to be produced every year. However, some detailed tables currently published may require accumulation of data across 2 or 3 survey years. In addition, published tables will be subject to a higher level of cell suppression because of a decrease in precision. Some subgroups, such as Native Hawaiians or Other Pacific Islanders (a standard subgroup in demographic analysis tables), already exhibit large amounts of suppression under the current design and will most likely either need to be eliminated or combined with other categories in order for any estimates to be available. Entire analysis tables, such as those that show estimates broken down by individual ages, may need to be dropped or revised.

As noted in Section 5.4.3.1, a reduction in sample size will result in a reduction in FI and FS staffing numbers. This may create a different interviewer and field management experience mix, which can affect annual estimates and trends. However, the ability to continuously identify and address any issues associated with obtaining cooperation from respondents, which can affect annual estimates, will be maintained.

5.4.4.2 Pooled-Year Estimates

Combining data over several years may be required for analyses of small populations or rare occurrences. Trend lines for direct State estimates, which are based on 3-year averages, will require pooling over more years and may not be as useful to policymakers.

State-level estimates produced using SAE procedures will require combining 4 years of data and will be available only every other year (see Table 5.3). State-level trend measures currently contrast moving averages each based on 2 years of pooled survey data with 1 year in common (e.g., [2003 + 2004] ÷ 2 vs. [2004 + 2005] ÷ 2). With a half-sample design option, trend measures will require 6 calendar years of data (i.e., [2014 + 2015 + 2016 + 2017] ÷ 4 vs. [2016 + 2017 + 2018 + 2019] ÷ 4). Substate estimates will require pooling 6 years of data (10 years for trend estimates) and therefore will no longer be practical.

5.4.5 Other Analyses

Other analyses using NSDUH data will be affected by the reduction in sample size. Although annual estimates of trends for CSAP's and CSAT's NOMs will be possible, sample size may be an issue for some analyses. Less data accumulation over time may reduce some opportunities for special studies and ad hoc analyses that benefit from combining data across years. Additionally, emerging trends will be more difficult to identify in a timely manner. Finally, a half-sample design limits the ability to conduct meaningful special data collection and analyses in specific areas to assess the impact of major national events or disasters (e.g., 9/11, Hurricanes Katrina and Rita, and the Deepwater Horizon oil spill).

Methodological studies also will be affected by a half-sample design. The smaller annual sample size may limit the ability to embed experimental studies and maintain the option to base trends on a subsample if the methods tested affect trend measures. It also will be more difficult for studies of rare phenomena (e.g., interviewer falsification) to draw meaningful conclusions.

5.4.6 Other Considerations

Similar to a biennial design, a reduction in sample size may send a signal that estimates obtained from NSDUH are less important than they once were or that the substance use problem is not as urgent.

5.5 Advantages, Disadvantages, and Impact of Switching to a Half-State Design

A half-State design involves surveying half of the States one year and the other half the following year. Thus, this design is very similar to a biennial design in that national estimates will be available only every other year. Over a 2-year period, 67,500 persons will be interviewed.

5.5.1 Survey Costs

Conducting NSDUH in half of the States one year and the other half the following year will reduce the number of completed interviews planned annually and therefore will result in large cost savings. However, compared with the current design, a half-State design will result in a higher cost per interview (see Section 5.2). Similar to a biennial design, there will be startup and shutdown activities in half of the States each year, and there will be fewer interviews over which to amortize fixed costs.

In addition, there will be an increase in costs associated with staffing and training because different States will be included each year, requiring additional FIs to be hired and trained or an increase in the number of long-distance trips by FIs. For example, assuming that the half-State design begins in 2014, there will be an increase in costs (a) after releasing a large portion of FIs in 2014 because the sample was not released in their States and (b) rehiring or replacing these FIs when the sample returns to their States in 2015. Even veteran FIs who were off the project for a year will require a more extensive retraining than an actively working veteran FI. These factors will lead to a significant increase in recruiting, training, and mentoring costs and will likely lead to an increase in data collection per interview costs because new FIs tend to be less efficient than veteran FIs. The increased cost estimate also assumes an increase in labor and shipping costs for returning equipment from FIs being released in one year and subsequently shipping equipment back out the following year to returning and new FIs in half of the States.

5.5.2 Survey Planning

Conducting the survey in half of the States each year will diminish the utility of the survey for planning and resource allocation purposes. Reliable data on some measures will not be available on an annual basis to inform budget decisions or performance measurement initiatives, such as SAMHSA's NOMs. Single-year estimates will not be nationally representative, and 2-year pooled estimates will be difficult to interpret because the survey will not be fielded in all of the States at the same time.

A half-State design also will limit questionnaire changes and the ability to respond to emerging issues. To maintain comparability, the questionnaire will need to be frozen over each 2-year period. There are schedule development implications to freezing the questionnaire over a 2-year period. Split-sample experiments will not be as efficient. Question revisions, inserts, and deletions, as well as methodological changes requiring testing in a split-sample experiment prior to implementation, will need to be identified about 1 year prior to the experiment. With a frozen questionnaire, it may be possible for results to be analyzed after the first year, depending on specifications of the question testing, but full incorporation into the survey will be delayed another 2 years. Thus, it will take about 3 years before such revisions needing pretesting can be fully implemented.

The current 50-State annual design provides the ability to respond to national crises (e.g., 9/11, Hurricanes Katrina and Rita, and the Deepwater Horizon oil spill) by expanding the sample in the impacted areas and conducting special analyses. The ability to respond to such crises will be limited by a half-State design because the impacted area may or may not be in the sample at the time of the major event or natural disaster.

5.5.3 Survey Operations and Data Quality

A half-State design provides the ability to retain FSs and professional and technical staff because the survey will be conducted on an ongoing basis. However, the FI staff likely will change from year to year because the sample will be in different States each year. The following sections provide greater detail on how a half-State design will affect field staffing, training, data quality, and other survey operations.

5.5.3.1 Field Staffing

Conducting the survey in half of the States each year will minimize some of the staffing challenges presented in a biennial design. For example, some stability of the FS staff could be maintained because work will be available each year. However, because data will be collected in only half of the States, the number of FSs will need to be reduced by half.

The FI staffing makeup will experience a shift, with the emphasis now being on FIs available to travel from State to State. Depending on which States are selected and their proximity to where an FI lives, some FIs living in bordering States could potentially travel to segments in one State in the first year and to a bordering State in the second year. However, travel costs will increase substantially. Staff retention for FIs working local assignments will remain the same as with a biennial design.

To further complicate matters, some FSs likely will be managing data collection in States about which they are not as knowledgeable. This sort of shift will require an adjustment period for FSs who need to become familiar with both a new region as well as a new team of local FIs. Although these sorts of adjustments can be accomplished, they will present challenges for FSs and ultimately raise data collection costs.

Some of the same challenges described in a biennial design will be similar with a half-State design. For example, unsteady employment for local FIs likely will result in high turnover rates, therefore requiring FSs to recruit and hire a large number of new FIs each year. In the first year of data collection, an FS's team likely will be made up of a small group of local FIs and a few FIs who can travel from State to State. In the second year of data collection, the FS team will be made up of a different group of local FIs and possibly this same group of FIs who are available to travel. Thus, team cohesiveness also will be diminished with a half-State design. With this design, NSDUH essentially will have two groups of FIs: (a) veteran FIs who will travel from State to State and (b) local FIs who will likely experience a high turnover rate because of the unsteady nature of the work.

Similar to a biennial design, it is anticipated that these half-State design issues will have an initial negative impact on response rates. Among the local FIs, this design will lead to hiring FIs who often switch studies or work on multiple projects at the same time. Thus, similar to a biennial design, loyalty to NSDUH will need to be built through other means to attract staff back after a year off, such as increased pay or count-and-listing work during "off years." As discussed in Section 5.3.3.1, the recruitment of experienced FIs will largely depend on other surveys active in the field at the time, as well as when these surveys begin and end data collection. For example, recruiting will present some challenges during the U.S. Census Bureau's decennial census given that a half-State design does not offer field staff continuous employment, which has traditionally made working on NSDUH more desirable than working on the Census Bureau's census. On the other hand, FIs located in the States where NSDUH data collection was inactive during the census can work on the census and then return to NSDUH the following year.

5.5.3.2 Training

As mentioned in the previous section, a half-State design will result in FI staffing changes, depending on which States are surveyed each year. Although some veteran FIs who travel from State to State will remain on staff from year to year, there still will be a need to hire a number of new FIs each year to maintain staffing levels in the States being surveyed each year.

Returning veteran staff working each year (both FIs and FSs) will complete refresher training consisting of an at-home review of the FI manuals and completion of several iLearning courses detailing any changes for the upcoming year, along with a general review of NSDUH protocols and other refresher training topics. Because data collection will be continuous and FI staffing will vary from year to year, an in-person FS team meeting will most likely not be held or at most will be only 1 day long.

New FIs and FSs will be required to attend an NTP FI training session and fulfill all requirements, including passing certification, to successfully graduate from training. The NTP sessions will be held in January, March, June, and September as needed to maintain staffing levels.

Although the overall number of FIs needed for data collection will be reduced, there will be an increase in the number of new FIs needed to fill vacancies from year to year in the States being surveyed. Additionally, veteran FIs who have not worked for a year or more and wish to return (e.g., an FI who worked the first year, did not work the second year, then wanted to return the following year) will need to complete a more intensive veteran training program, similar to the refresher training specified for a biennial design. These FIs will complete the at-home refresher portion (FI manual review and iLearning), followed by an approximately 3-day session, including mock exercises and a modified certification, to ensure each FI's understanding and compliance with all project protocols and procedures. The staffing needs for a half-State design can result in higher training costs (to train new FIs along with veteran FIs who have been off the project for more than a year). However, these training costs will be less overall than those incurred by a biennial design.

5.5.3.3 Data Quality

The design for a half-State alternative does not alleviate many of the data quality concerns posed by a biennial design. NSDUH still will have a large number of newly hired FIs at the beginning of each of the survey years, resulting in a spike in data quality issues as mentioned in a biennial design's implications.

The risk of data quality staff losing some of their skill set in an off year is alleviated by this design. However, overall, an immediate decline in data quality may be expected as staff adapt to this design. As with a biennial design, these challenges could be met by increasing verification protocols and adding additional data quality staff, which will increase costs substantially.

5.5.3.4 Startup and Shutdown Activities

Similar to a biennial design, a half-State design will require additional time and costs for startup and shutdown activities for computer equipment. Because the composition of the field staff is expected to change each year, project computer equipment will have to be returned, checked, cleared of files, and inventoried annually. This process then will be repeated in reverse order as equipment is prepared and assigned to new field staff.

5.5.4 Impact on Estimates

Under a half-State design, single-year estimates will not be nationally representative. Detailed tables and national findings reports can be published every other year; however, national estimates will be difficult to interpret because persons from all of the States will not be interviewed in the same year.

Analyses of small populations or rare occurrences that require pooling across 2 or more years of NSDUH data will require pooling over twice as many years to maintain equivalent precision. Trend lines for direct State estimates, which are based on 3-year averages, will require pooling over more years and may not be as useful to policymakers.

State-level estimates produced using SAE procedures will require combining 4 years of data and will be available only every other year (see Table 5.3). There also will be a conceptual problem with comparisons between State estimates based on data collected in even years and State statistics from odd-year data collection. These State comparisons can be biased by confounding even-year versus odd-year temporal effects. These potential even-year versus odd-year confounding effects also can bias the assignment of State small area estimates to the quintile ranks and associated color codes used to depict the State small area estimates on maps. State-level trend measures currently contrast moving averages, with each based on 2 years of pooled survey data with 1 year in common (e.g., [2003 + 2004] ÷ 2 vs. [2004 + 2005] ÷ 2). With a half-State option, trend measures will require 6 calendar years of data (i.e., [2014 + 2015 + 2016 + 2017] ÷ 4 vs. [2016 + 2017 + 2018 + 2019] ÷ 4). Substate estimates will require pooling 6 years of data (10 years for trend estimates) and therefore will no longer be practical.

The change in the mix of FIs expected with a half-State design can affect annual estimates and trends. Prior research has shown an increase in estimates and their SEs among new interviewers (Chromy et al., 2005). Additional methodological work will be required to fully understand the impact of the changed FI mix on the survey.

5.5.5 Other Analyses

A half-State design will affect other analyses using NSDUH data. Estimates of trends for CSAP's and CSAT's NOMs will be produced every other year. Emerging trends will be more difficult to identify in a timely manner. Finally, a half-State design will limit the ability to conduct special data collection and analyses to those areas that are in the sample at the time of a major event or natural disaster.

Methodological studies that are routinely conducted for NSDUH also may be affected by a half-State design. Methodological studies of national scope may be difficult to interpret if they are spread across 2 survey years. It also will be difficult to isolate the effect of methodological changes to the survey from real change in the estimates. Finally, a half-State design may result in methodological anomalies, such as "new interviewer" effects or decreases in response rate that may affect the quality of the estimates.

5.6 Conclusions

Switching NSDUH from the current annual data collection and analysis to a biennial design will result in large cost savings. Although the average cost per interview is expected to increase under a biennial design because of increased administrative and training activities and a "learning curve" for new and returning staff, it still will be less expensive to conduct the survey only every other year. However, there may be implications to the quality and timeliness of NSDUH data under a biennial design. The inability to retain field staff can contribute to an inconsistency in interviewer experience levels, which can affect response rates and respondent reporting of sensitive items. NSDUH no longer will have the ability to quickly respond to substantive issues or to conduct special analyses on the impact of national events. Finally, estimates produced using combined years of data, such as those for smaller geographic regions (e.g., States and substate areas) or for rare subpopulations (e.g., pregnant women), no longer will be timely and in some cases will become impractical.

The design alternatives of reducing the annual sample size by one half or conducting the survey in half of the States one year and the other half the next year also will result in significant cost savings. Similar to a biennial design, there are concerns with each of these design alternatives. A half-sample design will maintain consistency of field and professional and technical staff after the first year and therefore will eliminate any data quality concerns. However, the precision of annual estimates will be reduced, and analyses requiring pooling across 2 or more years of NSDUH data will require pooling over twice as many years to maintain equivalent precision. As a result, some analyses will not be timely enough to be useful to policymakers.

Although a half-State design will maintain the consistency of professional and some field management staff, the mix of FIs will change from year to year. This change in FI composition can affect prevalence estimates and response rates. Similar to a biennial design, a half-State designed survey's ability to respond to emerging issues or conduct special analyses following major national events will be limited because not every State will be included in the sample each year. Estimates will be published every other year, and results will be difficult to interpret because persons from all of the States will not be interviewed at the same time.

6. Expanded Target Population Assessment: Including Children Under Age 12

6.1 Introduction

There are two compelling reasons to consider expanding the target population for the National Survey on Drug Use and Health (NSDUH) to include children under 12 years of age: (1) some children under 12 do use drugs, and (2) a potential undercoverage of substance use initiation in the past year among persons aged 12 or older needs to be addressed. As shown in Appendix B of the 2007 NSDUH national findings report (Office of Applied Studies [OAS], 2008, p. 128), the estimates of undercoverage in past year initiation show the following:

> The overall adjusted estimate for past year initiates of alcohol use by persons 11 years of age or younger on the date of the interview was 279,949, or about 6.1 percent of the estimate based on past year initiation by persons 12 or older only (279,949 ÷ 4,558,760 = 0.0614).
>
> Based on similar analyses, the estimated undercoverage of past year initiates was 5.6 percent for cigarettes, 1.1 percent for marijuana, and 22.7 percent for inhalants.
>
> The undercoverage of past year initiates aged 11 or younger also affects the mean age at first use estimate. An adjusted estimate of the mean age at first use was calculated using a weighted estimate of the mean age at first use based on the current survey and the numbers of persons aged 11 or younger in the past year obtained in the aforementioned analysis for estimating undercoverage of past year initiates. Analysis results showed that the mean age at first use was changed from 16.8 to 16.3 (or a decrease of 2.8 percent) for alcohol, from 16.9 to 16.4 (or a decrease of 3.1 percent) for cigarettes, from 17.6 to 17.5 (or a decrease of 0.6 percent) for marijuana, and from 17.1 to 15.6 (or a decrease of 8.9 percent) for inhalants.

Based on literature reviews and NSDUH data from recent survey years, this chapter describes survey design issues for examination when considering collecting NSDUH data from children under 12. In Section 6.2, these considerations are presented, including issues related to sampling, nonresponse, questionnaire issues, data quality, and costs. Section 6.3 summarizes the key issues and challenges for consideration.

6.2 Considerations for Expanding Coverage

6.2.1 Sampling

Technically, the sample can be expanded to include children under the age of 12, but it is important to understand how this expansion will affect the sample size, variance, estimation, and other operational issues related to sampling, such as cost. These sample design and operational issues are described in the following sections.

6.2.1.1 Sample Design

The current annual national sample size for 12 to 17 year olds is 22,500, or 3,750 per each single year of age. Applying the same sampling rates to persons aged 10 and persons aged 11 will add 7,500 interviews. From an operational sampling perspective, the sample can be extended below age 10, but other considerations discussed in the remaining sections will demand extensive changes to the survey to make it practical.

Precision requirements and sample allocation also are driven by the prevalence rates. Table 6.1 and Figure 6.1 show how selected prevalence rates decline by age for those 12 to 17 (i.e., older individuals under the age of 18 have higher prevalence for most substances). Table 6.2 shows age at first use data reported by 12 year olds.

The expected lower prevalence rates for 10 and 11 year olds argue for a lighter sample allocation for the purpose of estimating overall substance use by broader age categories (e.g., 10 to 18, 10 or older). This means that an additional sample of fewer than 7,500 may be feasible. Estimates by single years of age may only be required at the national level. A process similar to what is done each year in allocating to those 26 or older will be needed to determine annual allocations to those aged 17 or younger.

Based on a previous examination of optimal cluster size (see Chapter 2), the additional sample requirements can be achieved efficiently by increasing the cluster size (persons sampled per segment) and maintaining the current sample of 7,200 segments.

6.2.1.2 Operational Issues

The current sampling algorithm is based on separate sampling rates by five age groups. In 1999, a version of the algorithm was applied to select the sample based on five age groups and three racial/ethnic categories for use with the paper-and-pencil interviewing (PAPI) mode on interviewing in a subsample of sample dwelling units (SDUs). As demonstrated by the 1999 PAPI modification, altering the sampling algorithm and the screening software to select from a younger age group at a specified rate is certainly feasible. The screening software will need to be modified to include rostering of all eligible persons, including the newly eligible younger age group.

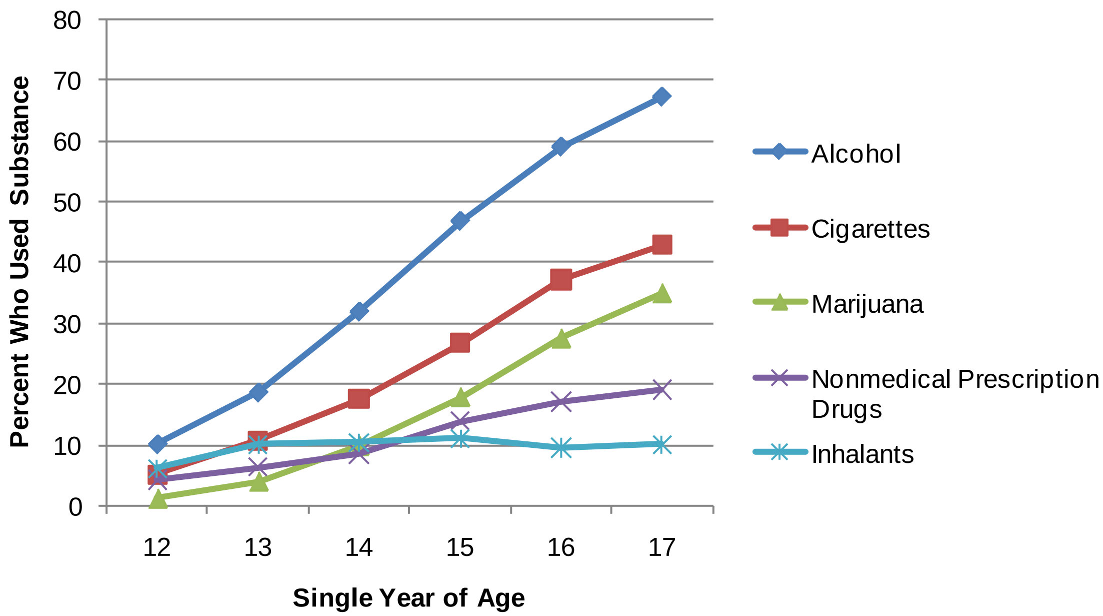
Under the current design, the person sampling rate for 12 to 17 year olds drives the screening sample size. If younger persons were sampled at about the same rate as 12 to 17 year olds, then more screening will be required. Previous investigations also have explored the possibility of selecting up to three respondents per SDU. This approach will reduce the need for increased screening and can be incorporated into the rostering and sampling application of the computerized screener. Based on experience with pair sampling, it is expected that sampling three persons per household will slightly reduce the interview response rates. However, the actual impact of selecting triads will need to be tested in a field study.

Table 6.1 Lifetime Use of Selected Substances, by Single Years of Age: Percentages, 2007 NSDUH

Substance	Age (in Years) 12	13	14	15	16	17
Cigarettes	5.1	10.7	17.5	26.8	37.2	42.9
Any Tobacco	6.4	13.2	21.7	31.1	42.5	49.7
Alcohol	10.1	18.6	31.9	46.8	59.0	67.2
Marijuana	1.3	4.1	9.9	17.9	27.5	34.9
Inhalants	6.1	10.1	10.4	11.1	9.5	10.1
Nonmedical Prescription Drugs	4.3	6.4	8.6	13.9	17.1	19.1
Pain Relievers	3.4	4.7	7.5	11.4	14.7	15.9

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007.

Figure 6.1 Lifetime Use of Selected Substances, by Single Years of Age: Percentages, 2007 NSDUH



Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007.

Table 6.2 Unweighted and Weighted Number of Persons Reporting an Age at First Use of 12 Years Old or Younger among Persons Aged 12, by Substance and Age at First Use: Annual Averages Based on 2005-2008 NSDUHs

Age at First Use (in Years)	Substance									
	Cigarettes		Any Tobacco		Alcohol		Marijuana		Inhalants	
	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted
1	*	*	5	849	7	2,096	*	*	6	1,256
2	*	*	5	1,025	11	2,660	*	*	*	*
3	7	1,992	11	2,806	19	4,162	*	*	4	1,366
4	27	6,211	32	7,476	18	5,585	4	876	13	3,978
5	27	5,920	42	10,522	43	12,547	*	*	22	7,198
6	27	5,139	34	6,976	40	10,496	*	*	21	5,873
7	52	11,929	60	13,724	54	11,911	*	*	34	10,679
8	63	14,444	98	21,647	86	23,612	6	1,431	44	11,192
9	106	25,669	134	33,569	121	31,706	8	1,601	79	21,038
10	173	45,528	196	51,934	363	90,352	49	10,037	234	63,036
11	217	51,838	253	62,080	437	116,150	75	18,287	359	101,315
12[1]	147	42,489	179	50,672	288	74,021	70	20,422	185	51,580

*Low precision; no estimate reported.

[1] The age 12 row represents approximately half of the number of persons initiating use at age 12 because 12-year-old respondents have, on average, half a year's experience at age 12 assuming uniform distribution of birth dates over the calendar year.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2005-2008.

6.2.2 Nonresponse

Two types of nonresponse have the potential to bias estimates among youths and adolescents, including any children under age 12 who might be selected for an interview using an expanded sample design. These types of nonresponse are (a) interview nonresponse, in which a youth is selected for an interview but does not complete one; and (b) item nonresponse, in which a sampled youth completes the interview but has missing data for an analysis variable of interest. Based on recent NSDUH data, assumptions about the occurrence of these types of nonresponse for sample members younger than age 12 are discussed in the following paragraphs.

6.2.2.1 Interview Nonresponse

Under the current design, younger persons tend to have higher response rates than older sample members. However, parents often refuse for their children, especially if the parent refuses to take the survey. Parents might have stronger objections if asked for permission to interview children under age 12 in the household. An analysis of parental refusal rates of youth interview respondents (respondents aged 12 to 17) was completed to test the hypothesis that parental refusal rates are higher among younger respondents.

Table 6.3 displays parental refusal rates from survey years 2005 to 2008 among respondents aged 12 to 17. The percentages represent the weighted proportion of incomplete interviews due to parental refusals among all eligible selected interviews in a single year of age. The parental refusal rate in each individual age category increased over time. Additionally, parental refusal rates for younger adolescents were consistently higher than those for older adolescents. When combining data from all 4 years, as displayed in the Total (All Years) column, it is evident that, overall, the parental refusal rate increased as the age of the youth respondents decreased. Based on these findings, it is reasonable to assume that the parental refusal rate for individuals under 12 years of age will be higher than for individuals 12 to 17 years old.

To determine whether there were significant differences in parental refusal rates among different ages within a survey year, an overall Wald F-test was performed for each survey year. Results from the overall Wald tests showed statistically significant values, which indicate there were notable differences among the single year ages within the 12 to 17 age group. To test for statistically significant differences among the parental refusal rates of each individual age category, pairwise comparisons were run between single year of age within each survey year. The findings of these tests, reported in Table 6.4, indicate that younger adolescents (i.e., 12 or 13 year olds) had significantly higher parental refusal rates than older adolescents (i.e., 16 or 17 year olds). Based on these results, the inference can be made again that respondents under age 12 are somewhat likely to have higher parental refusal rates than respondents aged 12 to 17.

The increase in parental refusal rates among younger ages is counteracted somewhat by a slight increased willingness for younger persons to agree to participate (if parents allow it). Although the refusal rates among 12 year olds are historically around 1 to 2 percent, the same rates for 17 year olds are approximately 3 to 4 percent. The overall refusal rates presented in

Table 6.3 Weighted Single-Year Parental Refusal Rates among Persons Aged 12 to 17: 2005 to 2008 NSDUHs

Respondent Age (in Years)	Parental Refusal Rate									
	2005		2006		2007		2008		Total (All Years)	
	Count of Parental Refusals	Weighted Percent	Count of Parental Refusals	Weighted Percent	Count of Parental Refusals	Weighted Percent	Count of Parental Refusals	Weighted Percent	Count of Parental Refusals	Weighted Percent
Total (All Ages)[1]	1,737	6.80	2,041	8.10	1,985	8.06	2,192	8.71	7,955	7.91
12	321	7.58	380	9.58	370	9.35	377	9.46	1,448	8.99
13	325	7.63	354	7.83	349	8.76	422	10.64	1,450	8.69
14	289	6.43	349	8.66	361	8.45	378	8.45	1,377	7.99
15	290	6.81	361	8.24	313	7.80	347	8.92	1,311	7.94
16	261	6.27	305	7.18	301	7.04	338	7.64	1,205	7.03
17	242	5.85	279	6.90	284	7.03	317	7.05	1,122	6.71

[1]Total (All Ages) includes respondents who were unspecified 12 to 17 years of age.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2005-2008.

Table 6.4 Parental Refusal Trend Tests among Persons Aged 12 to 17: 2005 to 2008 NSDUHs

Tests	Parental Refusal Rate							
	2005		2006		2007		2008	
	Test Statistic	*P* Value	Test Statistic	*P* Value	Test Statistic	*P* Value	Test Statistic	*P* Value
Overall	2.76	0.01**	3.96	0.00**	3.20	0.00**	5.39	0.00**
Pairwise Comparison								
12 versus 13	-0.07	0.94	2.34	0.02*	0.70	0.48	-1.38	0.17
12 versus 14	1.78	0.08	1.19	0.23	1.14	0.25	1.30	0.20
12 versus 15	1.20	0.23	1.69	0.09	1.96	0.05*	0.59	0.55
12 versus 16	1.89	0.06	3.21	0.00**	2.96	0.00**	2.50	0.01**
12 versus 17	2.59	0.01**	3.60	0.00**	3.10	0.00**	3.27	0.00**
13 versus 14	1.75	0.08	-1.21	0.23	0.39	0.69	2.58	0.01**
13 versus 15	1.23	0.22	-0.58	0.56	1.20	0.23	1.87	0.06
13 versus 16	2.08	0.04*	0.94	0.35	2.52	0.01**	3.77	0.00**
13 versus 17	2.78	0.01**	1.38	0.17	2.22	0.03*	4.40	0.00**
14 versus 15	-0.59	0.55	0.56	0.58	0.84	0.40	-0.55	0.58
14 versus 16	0.25	0.80	2.17	0.03*	2.03	0.04*	1.17	0.24
14 versus 17	0.95	0.34	2.58	0.01**	1.98	0.05*	1.99	0.05*
15 versus 16	0.88	0.38	1.43	0.15	1.04	0.30	1.55	0.12
15 versus 17	1.59	0.11	2.02	0.04*	1.07	0.28	2.21	0.03*
16 versus 17	0.69	0.49	0.43	0.67	0.01	0.99	0.85	0.39

* Significant at the 0.05 level.

** Significant at the 0.01 level.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2005-2008.

Table 6.5 are the sums of both the respondent refusal and parental refusal rates and indicate a virtual "cancelling out" effect of increased parental refusals versus the decreased refusal rate among younger age groups.

To further examine the distribution of the final interview results for persons aged 12 to 17, the completion and refusal rates were analyzed along with interview outcomes of the parental figures living in the same household (see Table 6.6). Without the availability of complete household relationship information at the person-selection stage of data collection, certain assumptions had to be made in the analysis summarized in the table. The subset of the entire selected sample of youths aged 12 to 17 used in the analysis consists of those youths from households where two persons were selected to complete the interview and their "relationship to screener" information indicated they were either (1) a son or daughter (including step) or son-in-law/daughter-in-law of the householder or the householder's spouse who also was selected, or (2) a sibling of the screener whose parent or guardian also was selected to complete the NSDUH interview.

Among adolescents whose parents completed the interview, the majority also successfully completed the interview (approximately 97 percent). Among adolescents whose parents refused to complete the interview, more than half of them were not given the opportunity to complete the interview because of parental refusal (58.09 percent on average in the 2005 through 2008 NSDUHs). The unweighted interview response rates among the adolescents whose parents refused themselves to complete the interview were similar across both the younger (12 to 14) and older (15 to 17) adolescent age groups—annual average of 24 to 25 percent. These results coincide with the results in Table 6.5, which show that younger adolescents were less likely than older adolescents to refuse once given the opportunity to participate, but their parents were more likely to refuse to allow them to participate. Again, it is difficult to assess how this phenomenon will play out with individuals under 12 years of age. The refusal rate is not expected to drop much below what it is for 12 to 14 year olds. Therefore, either the "cancelling out" phenomenon will occur for individuals under 12 years of age, or increased parental refusal rates for respondents under 12 years of age will have a negative impact on the response rate for these individuals.

6.2.2.2 Item Nonresponse

NSDUH respondents who are routed to a given question can have missing data for the variable(s) corresponding to that question if they do not know how to answer it ("don't know" or DK) or if they refuse to answer it ("refuse" or REF). A number of issues are particularly relevant to the occurrence of item nonresponse (i.e., missing data) and potential effects on data quality if the NSDUH sample were expanded to include youths under the age of 12. First, the rate of missing data for respondents younger than 12 can be higher than the rates for other youths aged 12 to 17 if they have difficulty understanding the meaning of certain questions and therefore are more likely to answer the question as DK. Second, the burden of completing a 60-minute interview for respondents younger than 12 can cause rates of missing data to increase for later sections of the interview. Third, younger respondents' perception of the sensitive nature of some questions can affect whether they refuse to answer.

Table 6.5 Weighted Interview Result Counts and Frequencies among Persons Aged 12 to 17: Percentages, 2005 to 2008 NSDUHs

Year	Interview Result Code	Respondent Age (in Years)					
		12	13	14	15	16	17
2005	70 - Interview Complete	87.45	87.29	88.90	86.71	87.25	85.21
	Overall Refusal	**9.19**	**9.96**	**8.58**	**9.78**	**9.29**	**10.25**
	77 - Refusal	1.61	2.33	2.15	2.97	3.02	4.39
	78 - Parental Refusal	7.58	7.63	6.43	6.81	6.27	5.85
	Other Nonresponse	3.35	2.76	2.52	3.51	3.46	4.55
2006	70 - Interview Complete	84.66	86.42	85.74	85.19	86.00	85.22
	Overall Refusal	**11.49**	**10.20**	**10.82**	**11.36**	**10.21**	**10.47**
	77 - Refusal	1.92	2.37	2.16	3.12	3.02	3.57
	78 - Parental Refusal	9.58	7.83	8.66	8.24	7.18	6.90
	Other Nonresponse	3.85	3.38	3.45	3.45	3.79	4.31
2007	70 - Interview Complete	84.56	85.16	85.93	86.20	85.46	84.83
	Overall Refusal	**11.63**	**11.05**	**10.58**	**10.48**	**10.36**	**10.54**
	77 - Refusal	2.29	2.30	2.13	2.68	3.33	3.52
	78 - Parental Refusal	9.35	8.76	8.45	7.80	7.04	7.03
	Other Nonresponse	3.80	3.78	3.48	3.32	4.17	4.63
2008	70 - Interview Complete	84.32	83.63	86.03	85.15	85.55	83.99
	Overall Refusal	**11.11**	**13.17**	**10.45**	**11.45**	**11.16**	**11.21**
	77 - Refusal	1.65	2.53	1.99	2.54	3.53	4.16
	78 - Parental Refusal	9.46	10.64	8.45	8.92	7.64	7.05
	Other Nonresponse	4.02	2.63	2.86	3.03	2.61	3.95

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2005-2008.

Table 6.6 Unweighted Interview Result Counts and Frequencies among Selected Adolescents Aged 12 to 17 in Households with Selected Parents, by Age Group and Parental Interview Results: 2005-2008 NSDUHs

	Parent Final Interview Results							
	Total		Completed		Refusal		Other Incompletes	
	Count	Unweighted Percent	Count	Unweighted Percent	Count	Unweighted Percent	Count	Unweighted Percent
Total								
Completed	16,345	83.67	15,166	96.72	692	24.39	487	47.84
Refusal	583	2.98	119	0.76	446	15.72	18	1.77
Parental Refusal	1,858	9.51	155	0.99	1,648	58.09	55	5.40
Other	749	3.83	240	1.53	51	1.80	458	44.99
12 to 14								
Completed	8,715	84.06	8,121	96.79	350	23.89	244	47.56
Refusal	229	2.21	43	0.51	178	12.15	8	1.56
Parental Refusal	1,048	10.11	103	1.23	913	62.32	32	6.24
Other	376	3.63	123	1.47	24	1.64	229	44.64
15 to 17[1]								
Completed	7,630	83.23	7,045	96.64	342	24.93	243	48.12
Refusal	354	3.86	76	1.04	268	19.53	10	1.98
Parental Refusal	810	8.84	52	0.71	735	53.57	23	4.55
Other	373	4.07	117	1.60	27	1.97	229	45.35

[1]Selected persons in the 12 to 17 age category with no single year of age specified were included in the 15 to 17 age group.

Note: Parent-child relationships were determined by reported roster information on *relationship to screener*. If a selected person was aged 12 to 17, reported being a screener's son (including step), daughter (including step), or son-in-law/daughter-in-law, and in the same household of a selected person aged 18 or older and reported as the householder or spouse of the householder, both selected respondents were included in this analysis. Additionally, if a selected person was aged 12 to 17, reported being a screener's brother (including step), sister (including step), or brother-in-law/sister-in-law, and in the same household of a selected person aged 18 or older reported as the screener's parent, guardian, or parent-in-law, both selected respondents were included in the analysis.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2005-2008.

Table 6.7 presents unweighted numbers of respondents aged 12 to 17 and weighted percentages of youths with missing data (i.e., any responses of DK or REF) across the 2005 through 2008 NSDUHs for selected lifetime substance use measures in the core NSDUH drug use modules, by single years of age. Missing data for lifetime alcohol use were examined because the alcohol use module occurs relatively early among the core NSDUH modules and because the lifetime alcohol use question includes the instruction for respondents not to include times when they "only had a sip or two from a drink." Missing data for lifetime use of specific hallucinogens were examined because this module introduces the concept of a category of drugs called "hallucinogens," then asks a series of questions about lifetime use of specific hallucinogens. The prescription pain relievers module is the first module that introduces the potentially complex concept of use of prescription drugs "that were not prescribed for you or that you took only for the experience or feeling they caused." These modules include potentially complex terms for children under the age of 12.

In terms of unweighted numbers of respondents, youths aged 12 or 13 comprised more than half of the respondents aged 12 to 17 in 2005 through 2008 who had missing data for each of the hallucinogen and pain reliever questions shown in Table 6.7. Youths aged 16 or 17 comprised about 12 percent or fewer of the respondents aged 12 to 17 with missing data for these questions.

Numbers of respondents and weighted percentages of youths with missing data were low for lifetime alcohol use in 2005 through 2008 and were slightly higher for youths aged 12 or 13 than for older youths. Weighted percentages of youths with missing data also were relatively low (1.6 percent or lower) for the lifetime hallucinogen and nonmedical pain reliever measures shown in Table 6.7. However, the percentages of youths aged 12 with missing data were higher than the corresponding percentages for youths aged 13 or older for all of the hallucinogen questions shown in Table 6.7. The percentages of youths aged 13 with missing data also were higher than those for youths aged 14 or older for all hallucinogen questions.

The percentage of youths aged 12 with missing data also was higher than those for youths of almost all other ages for lifetime nonmedical use of all prescription pain relievers. In addition, the percentage of youths aged 13 with missing data was higher than those for many older ages for nonmedical use of Darvocet®, Darvon®, or Tylenol® with codeine; Vicodin®, Lortab®, or Lorcet®/Lorcet Plus®; any pain reliever "below the red line" on the pain relievers pill card (corresponding to question PR04); and any other pain reliever.

Table 6.7 Unweighted Numbers of Respondents and Weighted Percentages with Missing Data for Selected Core Drug Use Measures among Youths Aged 12 to 17, by Single Years of Age: Annual Averages Based on 2005-2008 NSDUHs

Lifetime Use or Nonmedical Use	Age/Response of "Don't Know" or "Refused"											
	12		13		14		15		16		17	
	Un-weighted Number	Weighted Percent	Un-weighted Number	Weighted Percent	Un-weighted Number	Weighted Percent	Un-weighted Number	Weighted Percent	Un-weighted Number	Weighted Percent	Un-weighted Number	Weighted Percent
Alcohol	24	0.2	20	0.1	8	0.1	4	0.0[a,b]	*	*	5	0.0[a,b]
Hallucinogens												
LSD (Acid)	86	0.6	41	0.3[a]	17	0.1[a,b]	13	0.1[a,b]	*	*	*	*
PCP (Angel Dust)	83	0.6	39	0.3[a]	14	0.1[a,b]	8	0.1[a,b]	*	*	*	*
Psilocybin (Mushrooms)	197	1.5	104	0.7[a]	51	0.3[a,b]	32	0.3[a,b]	16	0.1[a,b]	7	0.1[a,b]
Ecstasy (MDMA)	115	0.9	58	0.5[a]	22	0.1[a,b]	9	0.0[a,b]	5	0.0[a,b]	*	*
Any Other Hallucinogen	116	0.8	70	0.5[a]	34	0.2[a,b]	17	0.1[a,b]	12	0.1[a,b]	6	0.0[a,b]
Prescription Pain Relievers												
Darvocet®, Darvon®, or Tylenol® with Codeine	211	1.6	137	0.9[a]	68	0.5[a,b]	38	0.3[a,b]	31	0.2[a,b]	18	0.2[a,b]
Percocet®, Percodan®, or Tylox®	95	0.6	53	0.4[a]	32	0.2[a]	18	0.2[a]	12	0.0[a,b]	12	0.2[a]
Vicodin®, Lortab®, or Lorcet®/Lorcet Plus®	91	0.7	57	0.4[a]	36	0.2[a]	22	0.2[a]	16	0.1[a,b]	13	0.2[a,b]
Any Pain Relievers below the Red Line on Card A	106	0.7	86	0.6	46	0.3[a,b]	30	0.3[a,b]	16	0.1[a,b]	7	0.1[a,b]
Any Other Prescription Pain Reliever	146	1.0	96	0.7[a]	59	0.4[a,b]	38	0.4[a,b]	22	0.1[a,b]	10	0.2[a,b]

*Low precision; no estimate reported.

[a] Difference between age and 12 year old is significant at the 0.05 level.

[b] Difference between age and 13 year old is significant at the 0.05 level.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2005-2008.

Table 6.8 presents unweighted numbers of respondents aged 12 to 17 and weighted percentages of youths with missing data from the 2005 through 2008 NSDUHs for selected measures from the noncore youth experiences and adolescent depression modules. Examination of missing data by age for these noncore modules can help evaluate the potential for missing data to increase in later self-administered modules for youths. The measures in Table 6.8 also were chosen for examination because of their potential complexity. For example, the question about stealing or trying to steal anything worth more than $50 requires youths to consider whether they *tried* to steal anything in addition to whether they actually stole something and also to consider the value of merchandise. The questions shown in Table 6.8 for the adolescent depression module are the "gate" questions that determine whether youths should be asked additional questions about their lifetime and past year depression symptoms for assessing major depressive episode (MDE). These questions also require youths to consider such concepts as being depressed, or feeling discouraged or hopeless, and also to consider whether they had "a period of time lasting several days or longer" when they had these problems for "most of the day." Consequently, item nonresponse in these gate questions can cause some youths with MDE—and especially younger ones—to skip out of the remaining questions in the adolescent depression module.

As for the core substance use measures presented in Table 6.7, youths aged 12 or 13 comprised more than half of the respondents aged 12 to 17 who had missing data in 2005 through 2008 for most of the measures shown in Table 6.8. For the adolescent depression measures, youths aged 16 or 17 comprised about 15 to 18 percent of the respondents with missing data.

Youths aged 12 or 13 had higher percentages of missing data for any school attendance in the past 12 months compared with youths aged 14 or older. The percentage of youths aged 12 with missing data also was significantly greater than other youths for the frequency of arguing with one or both parents in the past 12 months. For youths aged 13, however, only youths aged 16 or 17 had a significantly different percentage of missing data.

Percentages of youths with missing data did not differ significantly by age for the measures of stealing or trying to steal anything worth more than $50 or for attacking someone with the intent to seriously hurt them. The lack of significant differences for these measures may be related to the relatively low prevalences of these behaviors. In 2009, 4.4 percent of youths aged 12 to 17 stole or tried to steal anything worth more than $50 in the past year, including 2.0 percent of youths aged 12 or 13, 5.4 percent of those aged 14 or 15, and 5.5 percent of those aged 16 or 17 (OAS, 2010). An estimated 7.2 percent of youths attacked someone in the past year with the intent to seriously hurt them, including 6.6 percent of those aged 12 or 13, 7.5 percent of those aged 14 or 15, and 7.4 percent of those aged 16 or 17.

For the lead questions in the adolescent depression module, the percentage of youths aged 12 with missing data was significantly greater than the corresponding percentages for youths at all other ages. Youths aged 13 also had higher percentages of missing data compared with youths aged 14 or older for measures of feeling sad, empty, or depressed and feeling very discouraged or hopeless. For losing interest and becoming bored with most things, youths aged 13 had higher percentages of missing data when compared with youths aged 15 or older.

Table 6.8 Unweighted Numbers of Respondents and Weighted Percentages with Missing Data for Selected Noncore Measures among Youths Aged 12 to 17, by Single Years of Age: Annual Averages Based on 2005-2008 NSDUHs

Module/Measure	Age/Response of "Don't Know" or "Refused"											
	12		13		14		15		16		17	
	Un-weighted Number	Weighted Percent	Un-weighted Number	Weighted Percent	Un-weighted Number	Weighted Percent	Un-weighted Number	Weighted Percent	Un-weighted Number	Weighted Percent	Un-weighted Number	Weighted Percent
Youth Experiences, Past 12 Months												
Attended Any Type of School	224	1.7	149	1.0[a]	86	0.6[a,b]	48	0.3[a,b]	45	0.3[a,b]	30	0.2[a,b]
How Many Times:												
Argued with at Least One Parent	342	2.6	212	1.3[a]	195	1.2[a]	160	1.1[a]	127	0.9[a,b]	99	0.7[a,b]
Stole/Tried to Steal Anything Worth >$50	41	0.3	42	0.2	48	0.3	49	0.3	41	0.3	40	0.3
Attacked Someone with Intent to Hurt	54	0.4	47	0.3	45	0.3	45	0.3	34	0.1[a,b]	30	0.2[a]
Adolescent Depression												
Ever Had Period Lasting Several Days or Longer When Most of the Day:												
Felt Sad, Empty, or Depressed	286	2.1	176	1.2[a]	117	0.9[a,b]	90	0.6[a,b]	66	0.5[a,b]	53	0.4[a,b]
Felt Very Discouraged or Hopeless	257	1.8	153	1.0[a]	89	0.7[a,b]	78	0.6[a,b]	56	0.4[a,b]	49	0.3[a,b]
Lost Interest and Became Bored with Most Things	164	1.1	116	0.7[a]	82	0.5[a]	77	0.6[a]	50	0.4[a,b]	44	0.3[a,b]

*Low precision; no estimate reported.

[a] Difference between age and 12 year old significant at 0.05 level.

[b] Difference between age and 13 year old significant at 0.05 level.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2005-2008.

Taken together, the findings from Tables 6.7 and 6.8 indicate in the unweighted data that youths at the two youngest ages (12 or 13) tend to comprise the majority of respondents aged 12 to 17 with missing data. The percentages of youths at age 12 who had missing data also tended to be higher than corresponding percentages for older adolescents. Although this was not an exhaustive investigation of the rates of item nonresponse among youths, any factors contributing to item nonresponse among youths in the current design will likely lead to similar or higher rates of item nonresponse in youths under age 12.

6.2.3 Institutional Review Board Concerns

The NSDUH questionnaire contains a number of questions about sensitive topics. These topics include questions about the use of alcohol, tobacco, marijuana, cocaine, heroin, methamphetamine, hallucinogens, and inhalants. NSDUH also measures mental health status, mental health service utilization, physical health status, nonmedical use of prescription drugs, family income, and program participation. Considering the sensitive nature of many of the modules, it is possible that respondents under 12 years of age will be more vulnerable to parent coercion than 12 to 17 year olds. Specifically, children younger than 12 might be forced to reveal their answers to sensitive topics, especially if they are too young to fully comprehend confidentiality issues.

In order to determine whether children are capable of assenting, an institutional review board (IRB) must take into account the ages, maturity, and psychological state of the children involved. Based on discussions with a representative of RTI International's IRB, expanding the NSDUH target population to include children under age 12 will call for close examination of the survey design. Data collection procedures should be assessed to ensure they respect the rights of respondents of all ages. The following issues are not all expected to be significant, nor are they exhaustive, but they should be taken into consideration:

- Because of the lower cognitive ability of younger children, confidentiality issues and the voluntary nature of the child's participation will need to be emphasized.
- If a parent were to complete the interview before the child, the parent will have a clear sense of the kinds of questions being asked. This may lead some parents to ask their children how they answered these questions. Emphasis on the confidentiality of the data provided by children under 12 might help to address this concern.
- Providing noncash incentives, which may be more valuable to younger children, can address concerns about parents appropriating cash incentives paid to their children.
- Adults may be hesitant to leave younger children with adult strangers.

In considering the protection of human subjects, IRB representatives felt that the NSDUH team will need to provide some evidence showing that there are not negative outcomes from asking these sensitive questions of younger children. This point is also related to cognitive ability because a lack of question comprehension may trigger unique thoughts among younger children, such as making them feel they should engage in the behaviors mentioned in the survey items. Cognitive testing with respondents both younger than 12 and older than 12 can further inform whether the questions contained within the NSDUH interview have a differential impact upon children of different ages and can provide insight into the IRB's potential concerns.

Existing literature suggests that the NSDUH instrument will need to be modified to adequately address the IRB's concerns with interviewing young children. Knight et al. (2000) summarized the methodological and ethical challenges associated with child self-report of maltreatment. In this article, the authors considered age 12 as an appropriate age for asking children about perceptions of experienced abuse and neglect by citing the Ceci and Bruck (1993) study, which showed that younger children are more vulnerable to potentially suggestive influences in an interview than older children, adolescents, or adults. The authors were concerned that children younger than 12 who had not been abused might be traumatized or confused by direct questions on abusive behaviors. To reduce participant burden related to unnecessary exposure to sensitive questions, they constructed the interview to contain a limited number of key questions and the most salient follow-up items. These findings will need to be explored in the context of NSDUH, such as by conducting cognitive interviews with children under the age of 12 as suggested above. This might lead to different wording of questions or the omission of some questions that are included in the adolescent questionnaire.

6.2.4 Data Quality

As mentioned in Section 6.2.3, there might be greater pressure by parents of children under 12 years of age to reveal survey responses. Even if the pressure is the same as the pressure felt by older adolescents, there might be a perceived pressure by younger children, especially if they do not fully comprehend confidentiality procedures. Because of this, children under 12 years of age might be less likely to acknowledge use and report sensitive behaviors. In addition, younger children might not have the cognitive capacity to accurately recall past use. This leads to a concern that the validity of self-report of sensitive behaviors may be lower for children under 12 years of age than for older children. The research summarized below suggests that the validity of young children's self-reports may only be acceptable when using an age-appropriate instrument.

- Hamby and Finkelhor (2000) recommended collecting self-report data with children as young as age 7, including those aged 10 to 11. This recommendation was based on previous research showing internal consistency, test-retest reliability, and construct validity comparable with those obtained from older children for studying victimization of children. However, these authors did not present any direct evidence on the quality of data from children 10 to 11 years old compared with data from other age groups.
- Schwab-Stone, Fallon, Briggs, and Crowther (1994) showed that children were particularly unreliable in reporting about time factors, such as duration and onset of symptoms, and they concluded that highly structured diagnostic interviews might not be appropriate for use with children of elementary school age. The issue raised by Schwab-Stone et al. is applicable to children aged 10 to 11. This study indicated designing surveys to collect high-quality data from children in this age group can be challenging for some types of questions.
- Borgers, deLeeuw, and Hox (2000) conducted two studies on the influence of cognitive development on response quality in questionnaire research with children and showed that data quality increases with cognitive growth. They concluded that children aged 8 to 11 have enough language and reading skills to competently

complete many common types of survey questions. However, they also found that children of this age could answer only well-designed questions, have problems with negatively phrased questions, and have no tolerance for ambiguity. These findings apply to children aged 10 to 11, who were part of their target population. The authors argued that children can give consistent answers by about age 11, but still their cognitive ability improves with age and then stabilizes around age 14.

- Borgers and Hox (2001) investigated the effect of item and person characteristics on item nonresponse with children aged 8 to 18. They found that younger children do not perform as well as older children with more years of education because education level appears to be related to performance. This finding applies to children aged 10 or 11, who were among the younger children included in their target population. Using years of education as a proxy for age, the authors found that children with 4 or 5 years of education produced greater item nonresponse than children with 6 or more years of education. To reduce item nonresponse among young children, the authors suggested ways to make the instrument appropriate for young children, such as by avoiding ambiguous response scales or avoiding asking sensitive questions.
- Weber, Miracle, and Skehan (1994) indicated that the period between 10 and 14 years of age is a marginal and transitional one because early adolescents achieve competence in communicating thoughts, feeling, and behaviors. This point is similar to the conclusions reached by Borgers et al. (2000) and Borgers and Hox (2001). Combined, this research indicates that 10 or 11 year olds are at a developmental stage that could pose challenges for accurate survey measurement. Although research shows that children aged 10 or 11 have sufficient language and reading skills for various types of interviews, they interpret words very literally and, therefore, have no tolerance for ambiguity in meaning. These findings suggest survey instruments designed for 10 or 11 year old should include very simple instructions and questions and, when possible, use visual stimuli, such as showcards, to achieve data quality (Borgers et al., 2000).
- Stueve and O'Donnell (2000) showed relatively small levels of inconsistencies within baseline surveys and between baseline and follow-up surveys that collected 7th grade students' self-reports on substance use and sexual intercourse. For each behavior examined, less than 2 percent of the sample gave inconsistent answers within a survey and less than 7.5 percent did so over time. Retraction of baseline answers at follow-up was greater for rarer and more socially undesirable behaviors, such as cocaine use versus cigarette use. This finding suggests that the reliability of data from the youngest children currently included in NSDUH (12 and 13 year olds) may be acceptable for many survey items, but not necessarily the most "sensitive" questions. Because Stueve and O'Donnell (2000) did not include younger children in these surveys, their research does not directly demonstrate whether children aged 10 or 11 could answer questions as reliably as 12 or 13 year olds.
- In the 2006 NSDUH Reliability Study, youths aged 12 to 17 appeared less consistent than young adults aged 18 to 25 and older adults aged 26 or older in their reporting of substance use, particularly in the cases of lifetime and past year nonmedical prescription drug use and past year alcohol use (Kennet et al., 2010). Table 6.9 provides the response consistency and kappa estimates for these select substances

Table 6.9 Response Consistency and Kappa Estimates for Selected Substance Use among Persons Aged 12 to 17, by Age: 2006 NSDUH Reliability Study

Substance Use Variable/Age	Number of Respondents[1]	Weighted Percent Missing[2]	Weighted Percent Reporting Use (SE)[3]	Weighted Percent Reporting Consistently (SE)	Kappa (SE)
Nonmedical Use of Prescription-Type Psychotherapeutics in Lifetime[4]					
Aged 12 to 17	962	1.84	16.6 (2.9)	89.4 (2.1)	0.56 (0.08)
12	174	4.59	2.7 (1.5)	97.8 (1.5)	0.32 (0.22)
13	169	2.24	8.9 (4.6)	90.0 (3.9)	0.18 (0.14)
14	173	2.71	19.7 (8.7)	90.2 (4.0)	0.66 (0.13)
15	161	0.34	15.0 (5.8)	84.7 (6.7)	0.30 (0.21)
16	156	1.12	29.0 (10.3)	82.5 (7.2)	0.50 (0.16)
17	129	0.00	27.5 (9.2)	92.7 (3.0)	0.81 (0.09)
Nonmedical Use of Prescription-Type Psychotherapeutics in the Past Year[4]					
Aged 12 to 17	953	4.36	10.6 (2.3)	90.7 (2.0)	0.39 (0.12)
12	174	4.59	1.7 (1.4)	98.1 (1.5)	0.14 (0.15)
13	166	6.99	6.2 (3.5)	93.3 (3.5)	-0.01 (0.01)
14	171	3.22	12.3 (6.7)	96.7 (2.1)	0.85 (0.11)
15	159	0.71	11.7 (5.5)	83.0 (6.8)	-0.05 (0.06)
16	156	1.12	22.2 (7.9)	85.1 (7.1)	0.44 (0.23)
17	127	9.98	10.4 (3.7)	89.4 (4.2)	0.40 (0.16)
Alcohol Use in the Past Year					
Aged 12 to 17	973	0.91	28.8 (2.8)	88.6 (2.1)	0.72 (0.05)
12	183	0.00	2.4 (1.4)	98.9 (1.1)	0.69 (0.26)
13	169	4.28	3.7 (1.4)	94.3 (2.6)	0.38 (0.18)
14	173	0.31	31.7 (8.5)	86.2 (4.5)	0.65 (0.10)
15	162	0.00	32.3 (6.8)	93.2 (2.6)	0.85 (0.06)
16	157	0.00	43.4 (7.8)	76.2 (8.2)	0.51 (0.17)
17	129	0.00	68.6 (8.4)	79.6 (7.9)	0.52 (0.17)

[1] Respondents with nonmissing data at both interviews.

[2] Does not count legitimate skips where a response level could not be assigned based on responses to previous questions.

[3] Reported rates are provided because of the kappa statistic's dependence on prevalence. Kappa statistics should not be compared when their associated prevalence rates are dissimilar.

[4] Nonmedical use of prescription-type psychotherapeutics includes the nonmedical use of pain relievers, tranquilizers, stimulants, or sedatives and does not include over-the-counter drugs.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2006 Reliability Study (persons aged 12 to 17: *n* = 977).

among persons aged 12 to 17. For the nonmedical use of prescription-type psychotherapeutics in the lifetime and the past year, the kappas were lower for the 12 and 13 year olds than for the 16 and 17 year olds. Again, this suggests that younger children might not have the cognitive capacity to accurately report past use, especially when combined with complex questions or words, such as "tranquilizers" and "stimulants."

To explore potential data quality issues further, specific types of item nonresponse in the 2005 through 2008 NSDUH data were examined for the measures shown in Tables 6.7 and 6.8. If respondents aged 12 to 17 had missing data for the substance use measure in Table 6.7, the most common pattern by far was for them to answer a question as DK. Of the 503 respondents aged 12 to 17 in 2005 through 2008 who answered the question about lifetime nonmedical use of Darvocet®, Darvon®, or Tylenol® with codeine as DK or REF, for example, 469 answered the question as DK and only 34 refused. Except in situations where small numbers of respondents at a specific age answered a question as DK or REF, this pattern of youths with missing data tending to answer the questions as DK rather than REF also held within single years of age.

Responses of DK also tended to occur more commonly in the noncore measures in Table 6.8, but not to the extent as was observed for the substance use measures in Table 6.7. For example, 788 youths aged 12 to 17 answered the first adolescent depression question as DK or REF. Of these, 531 answered the question as DK and 257 refused. These findings, along with the results provided in Tables 6.7 and 6.8, suggest that younger respondents may have more difficulty understanding the meaning of complex questions or words, leading to item nonresponse.

A number of studies have examined children's cognitive capacity and how this relates to collecting accurate data from children. Based on this literature, it seems possible to conduct interviews with children under age 12 that elicit valid responses, as long as the instrument is designed to be age appropriate. This stipulation implies the need to make significant changes to the NSDUH instrument to make it appropriate for younger children, which will likely involve considerable effort.

6.2.5 Questionnaire Mode

The National Survey of Child and Adolescent Well-Being (NSCAW),[22] sponsored by the Administration for Children & Families (ACF) within the U.S. Department of Health and Human Services (HHS), examines child and family well-being outcomes in detail and seeks to relate those outcomes to their experience with the child welfare system. NSCAW is conducted by RTI International. NSCAW's sample includes a cohort of 6,000 children and adolescents who have come into contact with the child welfare system. Data are collected in four annual waves from the children, their biological mother, primary caregiver (if different), caseworker, teacher, and agency administrative records. Because the responses of children are such a large part of the NSCAW dataset, their experiences can be drawn from when designing survey instruments and modes for young children.

[22] Data and documentation can be found at http://www.acf.hhs.gov/programs/opre/abuse_neglect/nscaw/.

Evidence from NSCAW's cognitive testing of youth's understanding of informed consent procedures indicates that younger children are capable of participating in an audio computer-assisted self-interviewing (ACASI) interview. In NSCAW, children between the ages of 11 and 17 participate in an ACASI interview. This interview is administered in much the same way as the NSDUH interview, including the use of a tutorial at the front end. The tutorial is quite similar to the one used in NSDUH. For children who are aged 5 through 10, an "assisted ACASI" is used. Respondents listen to the question and point to their response on a showcard. The interviewer then enters the response.

Use of ACASI with children younger than 12 is rare. One study evaluated the validity of reports of the use of asthma-controller medication from children aged 8 to 18 and their parents (Bender et al., 2007). Researchers compared ACASI, face-to-face interviewing, and PAPI methods. They found that discrepancy between self-reported use and objective measures taken from electronic devices on the inhalers was highest in the ACASI mode and lowest in the face-to-face interview. Black and Ponirakis (2000) showed the advantages of using a computer-administered questionnaire to ask children sensitive questions and listed developmental considerations when preparing the instrument to maximize valid answers.

Based on these findings, it appears that most children 11 years of age or older have the cognitive ability to respond appropriately to the ACASI instrument. Careful consideration should be given to administering the instrument via ACASI to children younger than 11 without assistance. Usability testing and cognitive interviewing can further inform this issue.

6.2.6 Questionnaire Complexity

A reading assessment has been performed on the current NSDUH questionnaire, identifying needed edits in order to maintain a 6th grade reading level. In order to ensure that the instrument is age-appropriate, it is recommended that a similar assessment be performed if the target population were to be expanded to include children younger than 12 years of age. Items identified as containing potentially age-inappropriate items or constructs will have to be pretested with members of the younger age group to ensure that younger children are able to comprehend the questionnaire items. Because of the complex drug names and multiple reference dates that require recall (i.e., past 30 days, past 12 months, and lifetime), it might not be possible to reduce the NSDUH questionnaire level further to a 4th or 5th grade reading level.

The NSCAW questionnaire contains a wealth of questions that cover sensitive topics. These questions are asked of all respondents who are 11 or older. Topics include pregnancy and sexually transmitted diseases (STDs). The inclusion of such items suggests that NSDUH can question 11 year olds about the topics that are currently contained within the NSDUH instrument, but as discussed previously in Section 6.2.4, the validity of young children's self-reports may be questionable. This will likely lead to the alteration of some NSDUH questionnaire items in order to maintain reliability and validity among younger children.

6.2.7 Questionnaire Items

6.2.7.1 Current NSDUH Mental Health Measures for Youths Aged 12 to 17

For youths aged 12 to 17, NSDUH currently only collects data on depression using the major depression module (the MDE measure) from the National Comorbidity Survey's (NCS's) Composite International Diagnostic Interview (CIDI). There are no other measures within NSDUH to assess adolescent mental health.

The adult mental health measures in NSDUH (the Kessler-6 or K6, the Structured Clinical Interview for DSM-IV [SCID], and the World Health Organization Disability Assessment Schedule [WHODAS]) are not immediately "transferable" to either children or adolescents. So, any expansion to examine mental health in children will have to reflect measures for both adolescents and children because the existing measurement in adolescents is limited to depression.

The current NSDUH's serious psychological distress (SPD) and MDE measures (SCID interview, K6 drawn from the CIDI, and WHODAS) have been used with children as young as 13, but not younger than that. Ronald Kessler and his colleagues conducted a field-based household study with adolescents where they used the WHO-CIDI with youths aged 13 to 17 (Kessler et al., 2009). It has, though, been shown to have limited validity with children younger than 13 years (Merikangas, Avenevoli, Costello, Koretz, & Kessler, 2009).

6.2.7.2 Mental Health Measures for Youths Younger Than Age 12

Some child-oriented measures parallel those used in NSDUH to assess adult mental health *status*. A screening instrument can be used, which will be similar to how the K6 functions for adults within NSDUH, to identify children with serious emotional disturbance (to parallel the adult indicator of serious mental illness [SMI]), such as the Strengths and Difficulties Questions (SDQ) (Goodman, 1999). The SDQ and an impact supplement have been used for several years in another large household-based survey, the National Health Interview Schedule (NHIS). Then a "gold-standard" interview can be used to conduct a validation study to determine the predictive validity of the SDQ within NSDUH for children, as the NSDUH has done with the SCID for adults.

For youths under age 13, a full or partial parent report is generally recommended, which will be a significant deviation from the current NSDUH data collection protocol and instrument. There are structured interviews designed for administration down to 8 years of age, such as the Kiddie-Schedule for Affective Disorders (K-SADs) (Kaufman et al., 1997) and the Diagnostic Interview Schedule for Children (DISC) (Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000). These are the two leading structured clinical interviews (or gold-standard interviews) for children and are very similar in structure to the SCID for adults. Short forms of each are available, and one can "pull out" certain modules (just as NSDUH has done to assess depression in adolescents). These are designed to be administered to both children and their parents. The SDQ, on the other hand, is intended to be parent-completed only.

For mental health *service use*, the Service Assessment for Children and Adolescents, or SACA (Horwitz et al., 2001), which is currently used in NSDUH for adolescents, also can be used for children under 12 years of age. However, again, parents will need to be the reporters.

In summary, if younger age groups were added to NSDUH, and mental health were to be measured among this group, new mental health status scales will be required. The same mental health service utilization scale can be used, but it will be recommended that both types of scales be administered to adult proxies, rather than directly to children. Introducing parent reporting of sensitive behaviors in NSDUH might decrease the accuracy of the data and will require significant modifications to the data collection protocols, field interviewer (FI) manual, FI training programs, and the computer-assisted interviewing (CAI) instrument. For more discussion on proxy reporting, see Section 6.2.12.

6.2.8 Questionnaire Timing

A review of the timing data from the 2007 and 2008 NSDUHs revealed that 12 year olds had the longest interview lengths of interviewed minors. There was a general negative relationship between age and length of the interview. As the respondent's age increased, the length of the interview decreased (see Table 6.10). This pattern began to taper off with 15-year-old respondents, perhaps because of the increased reporting of behavior. This relationship held true for timing figures of the core modules, the noncore modules, the overall ACASI, and the total interview time. The completion times of respondents who were 15, 16, or 17 years old at the time of the interview were quite similar to each other, while the 12-year-olds' timing was the highest.

Table 6.10 Adolescent Total Timing Data, by Age: 2007 and 2008 NSDUHs

	2007/Age (in Years)						2008/Age (in Years)					
Timing Data	**12**	**13**	**14**	**15**	**16**	**17**	**12**	**13**	**14**	**15**	**16**	**17**
Sample Used in Analysis	3,531	3,759	3,702	3,816	3,857	3,706	3,475	3,665	3,777	3,808	3,933	3,814
Mean	65.5	63.5	61.2	63.0	63.2	63.3	64.5	61.5	60.1	60.7	61.3	62.2
Variance	246.9	280.1	256.3	280.4	271.6	287.5	257.4	243.0	257.3	245.3	261.6	280.0
Standard Deviation	15.7	16.7	16.0	16.7	16.5	17.0	16.0	15.6	16.0	15.7	16.2	16.7

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007-2008.

Across all age groups, timing data for the tobacco, alcohol, and marijuana modules were quite similar. Younger respondents took more time to answer the risk availability and youth mental health service utilization modules when compared with older adolescent respondents. Additional modules that take longer to administer to younger adolescents when compared with older adolescents include the inhalants module, the psychotherapeutic modules, and the youth experiences module. There are also modules within the questionnaire that take longer to administer to older adolescents when compared with younger adolescents. These differences could be due to increased reporting of behavior, as opposed to the increased cognitive burden of answering the questions at a younger age. The modules that saw a positive relationship between age and timing data include adolescent depression and substance dependence and abuse.

Because the mean timing data for the overall interview was higher for 12 year olds than for older adolescent respondents, it can be expected that increasing NSDUH's target population to include children younger than age 12 will increase the duration of the overall instrument administration. Depending on the age of the child respondent, a shorter questionnaire may be needed. Younger children may not be able to concentrate long enough to complete an instrument that averages 60 to 65 minutes in length. However, it is possible that introducing an abbreviated questionnaire for younger respondents can lead to context effects.

6.2.9 Costs

The more visits required by an FI to complete an interview, the higher the field costs associated with that effort. To investigate if interviewing younger respondents will have a negative impact on costs, the number of field visits needed to successfully interview respondents of different ages were compared. Tables 6.11 and 6.12 compare the records of calls (ROCs) needed to complete interviews (code 70s) among different age categories.

Table 6.11 shows that the 50 or older age group had the lowest number of ROCs for all age categories in both 2007 and 2008 (i.e., 4.52 and 4.44, respectively). Given the relatively low response rate among this age group, the low number of ROCs for code 70s might be due to these respondents being more likely to respond in the extreme, adamantly refusing or eagerly agreeing, than other respondents when asked to complete an interview. At 4.59 (2007) and 4.47 (2008) ROCs per successful interview, the youngest group (i.e., the 12 to 17 year olds) had the second lowest number of contacts for the remaining age groups (i.e., 18 to 25, 26 to 34, and 35 to 49) in both survey years.

Table 6.12 separates the minor age category into specific ages for a comparison with the 18 or older respondents' group. In 2007 and 2008, there were fewer ROCs recorded for successful interview completion with minors than for those respondents aged 18 years or older. In fact, except for 14 year olds in 2007, the data show an exponential increase in ROCs as the age of the respondent increased.

Based on the number of field visits needed to capture an interview, it appears to cost slightly less to interview younger respondents already in the NSDUH sample than to interview most older respondents. However, as discussed in Section 6.2.1, if increased sample sizes are required to expand the sample to include children under 12, field costs will likely increase due to the cost of needing to contact additional SDUs.

Table 6.11 Average Number of ROCs for Successful Interview Completion, by Age Category: 2007 and 2008 NSDUHs

	Age of Respondents				
Year	**12 to 17**	**18 to 25**	**26 to 34**	**35 to 49**	**50 or Older**
2007	4.59	4.94	4.98	4.91	4.52
2008	4.47	4.82	4.81	4.79	4.44

ROC = record of call.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007-2008.

Table 6.12 Average Number of ROCs for Successful Interview Completion, Minors and Those 18 or Older: 2007 and 2008 NSDUHs

Year	Age of Respondents						
	12	13	14	15	16	17	18 or Older
2007	4.46	4.53	4.45	4.57	4.68	4.85	4.88
2008	4.29	4.32	4.41	4.45	4.65	4.69	4.76

ROC = record of call.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH, 2007-2008.

6.2.10 Incentives

The positive effects of incentives on survey participation of adult respondents have been well documented in the research literature. In a meta-analysis of studies that experimented with the effects of incentives on in-person surveys, Singer, Gebler, Raghunathan, Van Hoewyk, and McGonagle (1999) found that cash incentives were more effective in raising survey response than were gifts and prepaid incentives. Simmons and Wilmot's (2004) literature review of the effects of incentive payments on social surveys corroborated these findings. Further, Davies and Rupp (2005-2006) found that debit card incentives—which are akin to offering cash—were associated with higher response rates than were checks and prepaid telephone cards (see their Appendix A: An Experiment with Response Incentives).

Beyond the research on the impact of alternative incentive payment methods for adults, little research has assessed the effects of different types of incentives on children's participation in surveys. Currently, all NSDUH respondents are given $30.00 cash as a token of appreciation for completing the NSDUH interview. Participation data from a previous round of NSDUH showed that cash incentives effectively lower refusal rates among survey participants aged 12 to 17 (Eyerman, Bowman, Butler, & Wright, 2005), suggesting that similar results might be observed if younger children were included in the study. Perhaps the most relevant study regarding the use of incentives in children's survey research was conducted by Bagley, Reynolds, and Nelson (2007). Their interview-based study included children aged 4 to 16 who were presented hypothetical situations regarding their participation in research, which included depictions of incentives that differed by type and dollar value. The researchers found that cash incentives were most likely to be effective with children who are at least 9 years of age, for these children could appreciate the value of money in general and as compensation for their research participation. Children younger than 9 years of age, however, did not seem to have the same grasp on the value of money and were instead drawn more to the influence of nonmonetary gifts. This study's results suggest that a different form of incentive, such as a gift card to a store well known to children, may be better received by many children under age 9, but a monetary incentive might be effective for 9 to 12 year olds.

Another issue to consider is how incentives paid to parents might influence this process. Using a token cash incentive to help obtain parental consent was found to increase early response among young adults and, thus, lessen the need for intensive nonrespondent follow-up (Mann, Lynn, & Peterson, 2008). However, the impact of parental incentives on the overall final response rates among young adults was found to be insignificant. This study's results suggest that paying an incentive to parents may have at best a modest positive impact on the completion rate of interviews with their children, including those under 12 years of age.

6.2.11 Field Interviewer Training

Prior to the development of any FI training content for a redesigned NSDUH with an expansion of its target population, all data collection instruments, protocols, and procedures must first be in final form. Only after all details of survey operations are in place, including the final programming of computer equipment to disseminate the survey instrument(s), can any interviewer manuals, materials, and training programs be constructed.

The materials development schedule is further complicated by the fact that the various manuals and materials cannot all be drafted on a concurrent schedule, but instead must be constructed in the proper order along with input and review from Substance Abuse and Mental Health Services Administration (SAMHSA) staff at each progressive step. At such a time when the survey instruments are finalized, as well as all materials to be used in conjunction with those instruments (lead letters, brochures, etc.), can these details be incorporated into an FI manual. Once the FI manual content is finalized, with text in place outlining all procedures and protocols, which will most likely take 6 to 8 months at minimum, then the development of a project training guide can begin, along with all associated exercises, videos, PowerPoint slides, and so on. Approximately 4 to 6 months of additional lead time will be required to finalize this FI training program.

Although NSDUH training and materials development staff can draw on experience from previous versions of NSDUH for the initial drafts of the manuals and materials, special attention will be required on content for issues specific to interviewing younger respondents and other redesign changes, such as needing to demonstrate a greater sensitivity to confusion about confidentiality assurances among younger respondents, as well as additional refusal conversion training for parents and special techniques for engaging younger respondents.

It will be necessary to develop, test, and produce all materials to be used with this youth-specific questionnaire content, in addition to a series of materials for FIs to present to concerned parents when asking permission to interview their young children, and possibly even a handout for the youth respondents themselves, detailing what will be expected of them. The reading level of any respondent materials must match the reading level of the youth-specific questionnaire content.

6.2.12 Use of Parents as Proxy Respondents for Younger Age Groups

A number of studies have investigated the validity of proxy reporting for young children. However, the value of proxy reporting has not been viewed as consistent across surveys. This finding is in line with the comments of an RTI children's mental health expert, who indicated that there were many survey instruments for parents to report upon the emotional and behavioral problems of their children. However, these proxy reports are not considered valid for delinquent behavior and other "externalizing" behavior problems. Thus, parent proxy reporting may not be adequate for surveys like NSDUH that ask questions about delinquent behavior.

When the survey topic is not sensitive and measures are related to facts that can be easily observed by others, proxy reporting seems to be more accurate. For example, Varni, Limbers, and Burwinkle (2007) evaluated the quality of parent reports for children aged 2 to 16 and found the parent proxy-report scales exceeded the minimum internal consistency reliability and construct validity. Russell, Hudson, Long, and Phipps (2006) compared parents' and child self-reporting on health-related quality of life and found high correlations; they concluded that parent reporting can be viewed as a reliable substitute when the child is either too young or too ill to provide a self-report. Finkelhor, Hamby, Ormrod, and Turner (2005) assessed the caregiver proxy report for juvenile victimization for children 2 to 9 years old; in a test of construct validity, endorsement of juvenile victimization items correlated well with measures of traumatic symptoms. They also found adequate test-retest reliability in a 3- to 4-week survey readministration.

Discrepancies between proxy reporting and children's self-reporting also are often identified in the literature. On measures of reports of depressive symptoms in one study, children reported more symptoms than their parents did (Angold et al., 1987). Jensen et al. (1999) found that parents and children rarely agreed on the presence of diagnostic conditions (such as mental disorders and psychiatric symptoms) regardless of diagnostic type, when children ranging from age 6 to 19 and their parents were independently interviewed. It also has been reported that parents' proxy reporting is not very accurate for assessing children's behavioral and emotional problems, such as delinquent activity (Achenbach, McConaughy, & Howell, 1987; Augustyn, Frank, Posner, & Zuckerman, 2002; Barry, Frick, & Grafeman, 2008).

A few studies showed the potential value of proxy reporting, but also acknowledged that there was variance in the quality of proxy reporting depending on the topics of measurement. Rosenbaum, Elliott, Kanouse, and Schuster (2007) compared child and parent reporting on their relationship and communication about sexual behavior; they found that survey response discordance was not solely attributable to response bias. Items about concrete events yielded better agreement than hypothetical items. By assessing agreement between parent proxy and child respondents on an epidemiological study, Whiteman and Green (1997) asserted that agreement between responses from children and parent proxies depends largely on the types of information sought, rather than the characteristics of the respondents. In summary, the literature suggests that the nature of items asked in the NSDUH interview, such as key measures of substance use, may significantly limit the ability of parents to accurately respond for children under age 12.

6.3 Conclusions

In this chapter, several issues have been documented for consideration when determining whether or not to expand the NSDUH target population to include children under age 12. The main benefit of expanding the sample is the ability to provide accurate estimates for this age group. Under the current design, the undercoverage of past year initiates aged 11 or younger also affects the mean age at first use estimate. The sample can be expanded to include children under age 12, but there are several issues to consider:

- **Sampling.** Expanding the sample to include children under the age of 12 will require modifying the screening software to roster all eligible persons, including the newly eligible age group. Because the number of selected households is typically based on the number of persons in the youngest age group, expanding the sample to persons under age 12 will increase the number of household screenings required. Applying the current sampling rates for 12 to 17 year olds to persons aged 10 and persons aged 11 also will add 7,500 interviews. Although increased sample sizes are operationally feasible, they will come with an increase in costs. The larger sample can be accommodated by either increasing the average cluster size or by allowing up to three persons to be selected per household. The latter may have a negative impact on interview response rates. From an operational sampling perspective, the sample can be extended below age 10, but the issues described in this chapter will demand extensive changes to the survey to make it practical.

- **Refusal Rates.** The parental refusal rate currently increases as the age of the youth respondent decreases. Because parents are more likely to refuse participation of their younger children, this trend might be exacerbated if children under 12 were included in the sample. This increase in parental refusals may have a negative impact on the response rate of the younger age group as a whole.
- **Data Quality.** An item nonresponse analysis of the 2005 through 2008 NSDUHs showed that the percentage of youths at age 12 who had missing data for select NSDUH modules tended to be higher than corresponding percentages for older adolescents. The majority of missing data was answered as "DK" as opposed to "REF." This implies that younger respondents have difficulty understanding some of the complex questions or wording in the NSDUH questionnaire, such as "hallucinogens." Furthermore, the research literature suggests that the validity of young children's self-reports may only be acceptable when using an age-appropriate instrument. This stipulation implies the need to make significant changes to the NSDUH instrument to make it appropriate for younger children. Based on the NSCAW experience, it appears that NSDUH can assume that children as young as 11 will have the cognitive ability to respond to the ACASI instrument, but children younger than 11 will need assistance.
- **Questionnaire Issues.** A reading assessment has been performed on the current NSDUH questionnaire, identifying needed edits in order to maintain a 6th grade reading level. A similar assessment will need to be performed again if the target population were expanded to include children under age 12. New mental health status scales also will be required. These changes will have to be pretested heavily with members of the younger age group to ensure that younger children will be able to comprehend the questionnaire items. Lastly, as the mean completion data for the overall interview is higher for younger respondents, one can assume that increasing the NSDUH target population to include children under age 12 will increase the duration of the overall instrument administration and it may be too burdensome or long for younger children.
- **IRB Concerns.** IRB representatives raised potential concerns related to question comprehension and cognitive ability, including the ability for young children to understand respondent rights, such as privacy, confidentiality, and the voluntary nature of the study. There are also incentive issues, including the type and amount of

incentive, that should be considered if expanding the survey population to include children under age 12.

In summary, a number of issues for consideration have been raised in this chapter for evaluating the expansion of the NSDUH population to include children under age 12. Extensive usability testing and cognitive interviewing will need to be conducted to inform this decision and to ensure that children under age 12 have the cognitive ability to comprehend confidentiality procedures and questionnaire items.

7. Flexible Design

7.1 Introduction

The current National Survey on Drug Use and Health (NSDUH) design structure and scope have been in place since 1999. A 5-year design was developed for the 1999 to 2003 survey years with reserve panels that allowed extending its use into 2004. A fresh 5-year sample was drawn for the 2005 to 2009 survey years using population projections based on the 2000 decennial census; reserve panels drawn at that time will allow continued use of this sample through 2014 if necessary. Reserve sample panels and the availability of unscreened dwelling units in phased-out segments have permitted the conduct of nationally representative special studies based on probability subsamples under both of these designs. In the event that the Substance Abuse and Mental Health Services Administration (SAMHSA) is required to reduce or supplement the sample to address a particular need, change from a 50-State design to a national design, or make other significant changes, it is important to know how flexible the current design is and, if not, how design adjustments can be incorporated to allow for more flexibility in the future.

This chapter briefly addresses several flexible design options for NSDUH. Section 7.2 discusses the past and current 5-year designs and their use or planned use over the period from 1999 through 2013. Section 7.3 discusses how the current design has been and can be adapted to fit other purposes, including a national design. Section 7.4 discusses features that could be incorporated to increase flexibility in the design. Section 7.5 summarizes the discussion.

7.2 Current Design

The following material has been taken primarily from the 2009 sample design plan (Morton, Chromy, & Hirsch, 2008a) and is included here for convenient reference.

7.2.1 Stratification

The first level of stratification in NSDUH is the State. The sample selection procedures began by geographically partitioning each State into roughly equal-sized State sampling (SS) regions; that is, regions were formed so that each area would yield, in expectation, roughly the same number of interviews during each data collection period. This partition divided the United States into 900 SS regions. Within each of these SS regions, a sample of census tracts was selected. Because census tracts usually greatly exceed the dwelling unit size requirement for a sampling cluster (150 dwelling units in urban areas or 100 in rural areas [more details to follow]), one smaller geographic region, or segment, was selected within each sampled census tract. In general, segments consisted of adjacent census blocks. In summary, the first-stage stratification for the 2005 to 2009 studies were States and then SS regions within States, the first-stage sampling units were census tracts, and the second-stage sampling units were smaller area segments.

Additional implicit stratification was achieved by sorting the first-stage sampling units (usually census tracts) by a core-based statistical area/socioeconomic status (CBSA/SES) indicator[23] and by the percent whites who are not Hispanic or Latino. The first-stage sample units were selected from this well-ordered sample frame using sequential probability proportionate to size (PPS) sampling with minimum replacement.

7.2.2 First- and Second-Stage Sample Selections

The design of 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. When necessary, census tracts were aggregated until each primary sampling unit (PSU, a census tract or collection of tracts) had the required minimum number of dwelling units for a cluster. After PSUs were formed, a sample of them was selected within each SS region as described above using a composite size measure (and sequential PPS sampling).

For the second stage of sampling for the 2005 to 2009 survey years, each selected PSU was partitioned into compact clusters of dwelling units by aggregating adjacent census blocks. Consistent with the terminology used in previous studies, these geographic clusters of blocks are referred to as "segments."

One segment was selected within each sampled census tract with PPS. Each segment was designed to contain a sufficient number of dwelling units to support one field test and two annual NSDUH samples. This allowed half of the segments used in any given year's main sample to be used again in the following year as a means of improving the precision of measures of annual change. This also allowed for any special supplemental sample or field test that SAMHSA may have wished to conduct within the same segments. This approach reflected an urban-area target of at least 150 dwelling units in a cluster (segment). Because it was more difficult to meet that target in rural areas, the target was lowered to at least 100 dwelling units in those areas, which still was sufficient to support one field test and two main study samples. Segments were constructed using 2000 decennial census data supplemented with revised population counts obtained from Nielsen Claritas.

Although the entire segment was a compact cluster, the final sample of dwelling units was a noncompact cluster. A sample dwelling unit in NSDUH refers to either a housing unit or a group quarters listing unit, such as a dormitory room or a shelter bed.

Noncompact clusters differ from compact clusters in that only a random sample of dwellings is included in the sample. Although compact cluster designs are less costly and more stable, 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).

Table 7.1 shows the allocation of the annual area segment sample for large sample and small sample States.

---

[23] 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, block group-level median rents and property values were given a rank (1…5) based on State and CBSA quintiles. The rent and value ranks then were averaged, weighting by the percent renter- and owner-occupied dwelling units, 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."

Table 7.1 Annual National Sample of Area Segments

Design Parameters	Small States	Large States	Total
Total Sample			
SS Regions	516	384	900
Segments	4,128	3,072	7,200
Total per State			
SS Regions	12	48	
Segments	96	384	
Total per SS Region			
Segments per Quarter	2	2	
Segments over Four Quarters	8	8	
Total States	43	8	

SS = State sampling.

To coordinate the sample selection for the 2005 to 2009 survey years, 48 census tracts and then segments were selected within each SS region. An equal probability subsample of eight segments was used for each NSDUH survey year. These eight segments were randomly assigned to quarters and to two panels within each quarter. The panels used in the 2009 NSDUH were designated as panels 5 and 6. Panel 5 segments were used for the 2008 survey and were used for the second time in 2009. Panel 6 segments were used for the 2009 survey only. Dwelling units that were not selected for the 2008 study were eligible for selection in the panel 5 segments in 2009.

With the 50 percent segment overlap from one analysis year to the next, the coordinated 5-year sample required a sample of 24 segments. An additional 24 segments per SS region were selected at the same time to provide additional flexibility in case the design requirements changed.

Table 7.2 displays how a sample of 24 segments selected in an SS region is used to provide 8 segments per year for 5 years. Note that the first and last samples are used only once, but all other samples are used in 2 successive years.

Table 7.2 Sample Rotation within an SS Region

Subsample		Year				
No.	**Segments**	**2005**	**2006**	**2007**	**2008**	**2009**
1	4	X				
2	4	X	X			
3	4		X	X		
4	4			X	X	
5	4				X	X
6	4					X

SS = State sampling.

As noted previously, enough reserve samples were selected to continue use of the sample through 2014 if desired. In fact, additional panels of four segments per SS region were selected from the reserve sample to field each of the 2010 through 2013 surveys. Samples 6 and 7 were used for the 2010 NSDUH, and the 50 percent overlap in sample segments will continue through 2013.

In addition to the reserve sample, "retired" segments or segments that have already been used for 2 survey years have been used to supplement the sample for special field studies. For example, the 2007 Mental Health Surveillance Feasibility Study was fielded in 10 NSDUH segments from 2006 that had been retired from use in the main study (Morton, Martin, Hirsch, & Chromy, 2008b). The advantage of using retired segments for field tests is that the sample segments have already been field enumerated and there are no additional costs associated with preparing the list frames. However, because these segments were field enumerated 2 or more years earlier, special procedures are required to handle changes associated with using old listings (e.g., the half-open interval procedure for adding missing dwelling units).

7.2.3 Third and Subsequent Stages of Sample Selection

The selection of dwelling units (third-stage units) and persons (fourth-stage units) is specific to each annual survey or supplement.

For the 2009 NSDUH, the sampling frame for the third stage of sample selection was the lines of listed dwelling units. After accounting for eligibility, nonresponse, and the fourth-stage sample selection procedures, it was estimated that roughly 190,800 lines needed to be selected to obtain a sample of 67,500 responding persons distributed by State and age group as shown in Table 7.3.

Table 7.3 Summary of the 2009 NSDUH Design

Statistic	Small States	Large States	Total
Total Sample			
SS Regions	516	384	900
Segments	4,128	3,072	7,200
Selected Lines	109,392	81,408	190,800
Expected Eligible Dwelling Units	90,214	67,136	157,349
Expected Completed Screening Interviews	80,290	59,751	140,041
Expected Selected Persons	48,977	36,448	85,425
Expected Completed Interviews	38,700	28,800	67,500
Total per State			
SS Regions	12	48	
Segments	96	384	
Expected Selected Lines	2,544	10,176	
Expected Completed Interviews	900	3,600	
Expected Interviews per Segment	9.38	9.38	
Total per SS Region and Segment, by Quarter			
Segments per SS Region	2	2	
Expected Interviews per SS Region	18.75	18.75	
Expected Interviews per Segment	9.38	9.38	
Total States	43	8	51

Note: Table 7.3 was updated using 2007 NSDUH experience.

SS = State sampling.

7.2.4 Comparisons with the First 5-Year Design for 1999-2003

The 1999 to 2003 design formed first-stage units directly as compact area segments. The additional stage of sampling in the 2005 to 2009 design provided better control over forming area segments within census tracts. It also provides some additional flexibility in sampling within census tracts, such as for adjusting the minimum size of segments in future panels.

Larger sizes of segments were required in the 1999 to 2003 design, with a minimum of 175 dwelling units for both urban and rural segments (Bowman, Penne, Chromy, & Odom, 2001, p. 2). This larger size permitted the simultaneous fielding of the 1999 paper-and-pencil interviewing (PAPI) survey in a probability subsample of area strata. Unused lines in retired panels also were used to field a special Validity Study based on the collection and testing of hair and urine samples (Harrison, Martin, Enev, & Harrington, 2007).

Additional sample panels also were selected for the 1999 to 2003 design, and a subsample of the unused panels was used in the 2004 survey.

7.3 Flexibility of Current Design

This section addresses the current design's ability to adjust to possible changes without completely redesigning a 5-year survey plan, such as the one employed in the 2005 to 2009 NSDUHs. This discussion relates to Task 1 in the work plan: "Document the features in the current sample design that allow for flexibility, such as drawing additional PSUs to allow for supplemental studies."

It is important to note that although there is some flexibility in the current NSDUH design, there may be issues such as loss of data comparability (i.e., trends), decreased precision, and changes in operational procedures that may have cost implications. Thus, the term "flexible" refers to the ability to change from one design to another from a sample design perspective. The discussion of such designs does not include all potential consequences associated with these changes.

7.3.1 National Design

The design goals and major features of a national design will differ from those of the current design, as illustrated in Table 7.4.

The quickest adaptation of the current design to a national design will involve designating the current SS regions as PSUs. This approach was applied to the PAPI subsample in the 1999 survey in a combined survey context, but it could be applied equally well to a stand-alone national survey. It also was applied to a study of alternative incentive plans during the first two quarters of the 2001 survey. Preserving the current allocation of eight segments per SS region to the new PSUs allows continuation of the control of the sample distribution across quarterly surveys and across replacement panels. Designs for a national sample of 40,000 with PSUs drawn from the current frame of 900 SS regions are presented in Table 3.18 of Chapter 3.

Table 7.4 Comparison of Design Goals and Major Features for the Current 50-State Design and a National Design

Design Goal/Feature	Current 50-State Design	National Design
Sample Size	Sample size is set by precision at State level, resulting in large national sample. Sample size for major demographic groups is expected to be satisfied by State precision requirements. Current sample size is 67,500 persons aged 12 or older.	Sample size is set by precision at national level for major demographic groups. It is reasonable to expect that a much smaller sample size will satisfy precision requirements.
Use of Large Geographic Primary Sampling Units (PSUs)	Direct sampling of census tracts or geographic segments is done because State sample size requirements are large enough to require travel to most parts of the State.	Travel cost efficiencies are possible through use of larger areas (usually metropolitan areas or groups of counties) as PSUs.
Geographic Sample Distribution	Minimum sample sizes are allocated to each State.	Sample is distributed in proportion to population. Largest States will get largest share of the sample. Some of the smallest States might not have any sample in a particular year.

7.3.2 Hybrid Designs

Designs that preserve minimum State sample sizes at some level by supplementing a national sample also have been considered. One approach when selecting a new sample will require redefining the SS regions in most States. Table 3.19 in Chapter 3 shows some of the possible results of defining new SS regions summarized to the national level under a specific requirement for an effective sample size of at least 400 for past month marijuana use estimates among persons aged 12 or older. The presumption made in those designs is that the State estimate requirement will make it optimal to sample all SS regions (i.e., treat them as strata rather than PSUs).

Another approach will be to use the current annual sample as an initial phase for the sampling of segments, followed by further stratification and selection of a final subsample from each replacement panel that still meets the quarterly allocation requirements. This approach will allow switching to a hybrid design while continuing to utilize a multiyear sample selected under the current design objectives.

7.3.3 Changing Cluster Size

The current design can be adapted fairly easily to cluster size changes that involve doubling or tripling the cluster size. Some examples of alternative designs that also triple the cluster size were discussed in Chapter 3. A tripling of cluster size under the current design will be accompanied by collapsing sets of three neighboring SS regions and selecting subsamples of the current samples of area segments. A similar procedure will work equally well for increasing the cluster size by some other integer multiple. Intermediate adjustments in cluster size will be difficult to implement under the current design with 900 SS regions, 8 segment samples per SS region, and holding the total sample size constant. If the cluster size is increased by a noninteger multiple, tailoring the current design to this change will require subdividing and reconfiguring the SS regions. For example, increasing the cluster size by a factor of 1.75 will require forming approximately 514 SS regions from the current 900 regions.

7.3.4 Address-Based Sampling

One of the advantages of address-based sampling (ABS) is that larger geographic areas can be used to define segments without increasing the total cost per unit sampled. The savings associated with frame development will most likely exceed the increased cost of data collection in larger geographic areas. The economy of ABS enables a segment to be defined as a census block group or even an entire census tract. The introduction of census tracts as intermediate stages of sampling provides immediate flexibility in taking advantage of this feature.

In addition, ABS obviates the need to use retired segments whose listings are at least 2 years old for special studies. Instead, current address listings can be purchased shortly before data collection. In the event that additional sample panels are needed (e.g., for a new survey year), ABS can be used to develop a frame in a matter of weeks instead of the months needed for conventional listing. In cases where the coverage of city-style mailing addresses is not adequate, the land area of a census tract can still be subsampled for conventional listing purposes. Based on prior research, it is expected that approximately 85 percent of NSDUH segments have adequate ABS coverage. Conventional listing will be retained in the remaining, primarily rural segments.

Finally, the ABS cost savings eliminate the need for retaining half samples of segments. If an overlapping sample design is preserved, the ABS lists will be retained (but updated) for 2 years, similar to the traditional lists.[24] Dwelling units that were selected one survey year will be ineligible for selection the following year.

7.3.5 Continuous Survey Options

Many national surveys field a single sample and allow the work to extend for a period of months. If response rates are low, the survey period may sometimes be extended to allow additional follow-up of nonrespondents. Adaptive procedures (e.g., subsampling nonrespondents for more intensive efforts) may be applied as the survey work proceeds in attempts to gain cooperation from hard-to-reach or generally uncooperative sample members.

Telephone and log-based surveys may use very short data collection periods targeting as little as a single week. The Current Population Survey (CPS) selects monthly samples and defines 1 week in the month as a reference period (U.S. Census Bureau, 2006, pp. 3-14 and 5-2).

In contrast to the one-time surveys, the current NSDUH design is set up specifically for fielding four nonoverlapping samples to be worked completely in each of four quarterly surveys.

[24] To remain current, the ABS lists can be updated annually, and dwelling units that were previously selected can be flagged as ineligible.

No work is carried over to the next quarter for additional follow-up. The third month in each quarter often is devoted to "clean up" operations or for assisting interviewers who have fallen behind schedule in their work. Although the distribution of completed interviews by quarter is controlled by the sample allocation and work assignment procedures, there is little control over the distribution of completed interviews by month or week within a quarter. In practice, interviewers are encouraged to advance their work as quickly as feasible to permit time for nonresponse follow-up in the third month if necessary. If the cumulative yield for a State is lagging, quarterly allocations may be increased for subsequent quarters and additional sample in the same segments may be released during the quarter for work in the latter part of the quarter.

Before implementing a more continuous data collection plan, decisions will need to be made on the following design issues:

1. *Choice of period*: It is assumed that more continuous data collection implies shorter survey periods, such as months or weeks versus the current quarterly surveys.
2. *Scheduling of cleanup operations*: Currently, interviewers have 3 months to complete all calls within an assigned sample segment. This could be shortened to correspond to the survey period, or some cleanup could be allowed while the next sample release is in progress.
3. *Release of the sample*: The sample could be released by segment, by addresses within segment, or by some combination of the two.
4. *Residence rules*: Current residence rules are based on being in residence at an address during half or more of the calendar quarter. Some alternative rules should be considered with more continuous data collection.

These topics (and others) are discussed in Chapter 4. A few comments are included in this chapter about the flexibility of the current design to transition toward more continuous data collection.

Assuming a release of monthly survey samples, State samples can be partitioned into twelfths, but some SS regions will be omitted from each twelfth. In other words, in order to adapt the current quarterly design to a monthly design, a third of the SS regions will be fielded each month such that sample is selected from each region 4 times a year. Weekly releases under the current design will require some partitioning of the sample within segments similar to the practice of releasing additional lines currently employed to control sample yield after the quarter has begun. None of these partitions will be particularly appropriate for designing analysis strata and replicates without making some convenient assumptions about temporal and spatial effects.

7.3.6 Biennial Surveys

The current sample design can be transitioned easily to a biennial survey design as long as there is no concern about the aging of dwelling unit addresses. However, if conducted every other year, it is important to note that NSDUH no longer will have the flexibility to quickly respond to substantive issues or to provide data to conduct special analyses on the impact of national events.

As discussed in the Chapter 5, a major disadvantage to the biennial survey design is that it will be difficult to retain experienced and talented professional staff and interviewers who are reassigned to work on other projects or who seek new employment in an "off year." The inability to retain field staff can contribute to an inconsistency in interviewer experience levels, which can affect response rates and respondent reporting of sensitive items. In addition, estimates produced using combined years of data, such as those for smaller geographic regions (e.g., States and substate areas) or for rare subpopulations (e.g., pregnant women), no longer will be timely and in some cases will become impractical.

7.3.7 Person Sampling

Sampling allocations by age group are currently reviewed before fielding each annual year of quarterly surveys. Allocations by other characteristics or by a different age group will require reprogramming the screening software, but it will not require any major adjustment to the basic design.

7.3.8 Target Population Definitions

Target population definitions currently being considered relate to including persons younger than 12 years of age. The main benefit of expanding the sample to children younger than 12 years old is the ability to provide accurate estimates for this age group. As discussed in Chapter 6, such transitions are feasible within the current overall design, but there are several issues to consider, such as the impact on refusal rates, data quality, costs, and questionnaire comprehension.

Expanding the sample to include children under the age of 12 will require modifying the screening software to roster all eligible persons (including the newly eligible age group). Because the number of selected households is typically based on the number of persons in the youngest age group, expanding the sample to persons under the age of 12 will also increase the number of household screenings required. The larger sample could be accommodated by either increasing the average cluster size or by allowing up to three persons to be selected per household. The latter may have a negative impact on interview response rates.

If the target population were expanded in some way that required additional geographic coverage or the inclusion of persons not residing in the currently defined NSDUH dwelling units, some design adjustments will be required.

7.4 Design Adjustments for More Flexibility

This section addresses ways to increase the flexibility of the next design for NSDUH, as outlined in Task 2 of the work plan: "Investigate options for transitioning to different designs. Two options to look into are increasing the number of State sampling [SS] regions to allow for supplementing the sample in specific areas and the ability to transition to a national design from a 50-State design."

Since 1999, two coordinated multiyear plans have been implemented. Prior to that time, a mix of single year plans with new samples and continuing surveys that used the same sample, in whole or at least in part, were employed with less forethought to long-term transition.

Although there are options for transitioning to a different design, it is important to consider that there may be issues, such as loss of data comparability (i.e., trends). A new design is free to make a complete break with the past; however, some smoothing of the transition can be beneficial for maintaining comparable trend measures and for retaining experienced field staff. The shift to a 50-State design and the large accompanying increase in sample size in 1999 had both expected and unexpected consequences. One of the unexpected consequences was the resulting change in the distribution of interviewers by prior relevant experience and the effects this change had on both response rates and the reporting of sensitive behaviors in self-administered questionnaires. This section addresses ways to increase flexibility on the next design from a sample design perspective. The discussion of such designs does not include all potential consequences associated with these changes.

7.4.1 National Design

The approach outlined in Section 7.3.1 is very convenient for switching to a national design because it would also allow each transition of the sample back to a 50-State design. A less flexible national design might take better advantage of defining PSUs based on type of community and other summary demographic measures. These types of designs often lead to the use of counties or contiguous sets of counties as PSUs.

7.4.2 More State Sampling Regions

The current design is based on annual samples of eight segments within each stratum (SS region). Designs could be developed that select two segments per SS region, allowing a quadrupling of the number of strata and the development of a coordinated randomization scheme to assign the sample to quarters or to months if more continuous data collection is considered. For annual estimates, variance estimation strata and replicates will be defined as SS regions and segments, respectively. For analyses involving shorter periods (months, quarters, or half years), SS regions would be collapsed into a smaller number of strata so that each strata is represented in the analytic period.

One purpose of stratification is to allow a different sampling strategy in different strata. Adding strata that focus on certain populations could be built into a new, more flexible design. For example, census tracts with high college student populations could be identified in which special approaches might be followed to sample dormitories and other student housing.

7.4.3 Hybrid Designs

As discussed in Section 7.3.2, the current design can be adapted to hybrid designs.

7.4.4 Changing Cluster Size

Some initial decision about the general magnitude of cluster size in advance will help to make the design flexible for implementing other features given the cluster size. Increasing cluster size under any scheme will be accompanied by decreasing the number of clusters (segments) and the number of strata (SS regions).

7.4.5 Address-Based Sampling

Using ABS will make it feasible to redraw and fine-tune the sample for each successive year. Some inefficiencies and increased costs for areas where traditional listing is required will result and will need to be considered in any cost tradeoffs among design alternatives.

7.4.6 Continuous Survey Options

For monthly surveys, an ideal approach will be to select 12 segments per stratum and allocate 1 to each month. This approach runs contrary to increasing the number of strata based on the current design, so some compromise approach will most likely be required.

7.4.7 Biennial Surveys

The current design has adequate flexibility with regard to biennial surveys.

7.4.8 Person Sampling

The current design has adequate flexibility with regard to person sampling.

7.4.9 Target Population Definitions

Depending on the type of target population definitions considered, a new design could be made more flexible. As an example, military personnel are currently excluded from NSDUH. Separate strata and special procedures will be required to include this population in the survey. Changes within the current dwelling unit population can be easily implemented from a strictly sampling viewpoint.

7.5 Conclusions

In the event that SAMHSA is required to reduce or supplement the NSDUH sample to address a particular need, change from a 50-State design to national design, or make other significant changes, there are flexible design options from a sample design perspective. Depending on the objective, flexible design options for the current and next NSDUH design include hybrid designs, changing cluster sizes, ABS, continuous or biennial data collection, and expanding the target population. However, although they provide flexibility, the impact of these options on other areas of the survey design should be considered, such as loss of data comparability (i.e., trends), decreased precision, and changes in operational procedures that may have cost implications.

References

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*, 213-232.

Angold, A., Weissman, M. M., John, K., Merikangas, K. R., Prusoff, B. A., Wickramaratne, P., Gammon, G. D., & Warner, V. (1987). Parent and child reports of depressive symptoms in children at low and high risk of depression. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 28*, 901-915.

Augustyn, M., Frank, D. A., Posner, M., & Zuckerman, B. (2002). Children who witness violence, and parent report of children's behavior. *Archives of Pediatrics & Adolescent Medicine, 156*, 800-803.

Bagley, S. J., Reynolds, W. W., & Nelson, R. M. (2007). Is a "wage-payment" model for research participation appropriate for children? *Pediatrics, 119*(1), 46-51.

Barry, C. T., Frick, P. J., & Grafeman, S. J. (2008). Child versus parent reports of parenting practices: Implications for the conceptualization of child behavioral and emotional problems. *Assessment, 15*, 294-303.

Bender, B. G., Bartlett, S. J., Rand, C. S., Turner, C., Wamboldt, F. S., & Zhang, L. (2007). Impact of interview mode on accuracy of child and parent report of adherence with asthma-controller medication. *Pediatrics, 120*, e471-e477.

Black, M. M., & Ponirakis, A. (2000). Computer-administered interviews with children about maltreatment: Methodological, developmental, and ethical issues. *Journal of Interpersonal Violence, 15*, 682-695. doi: 10.1177/088626000015007002.

Borgers, N., deLeeuw, E., & Hox, J. (2000). Children as respondents in survey research: Cognitive development and response quality. *Bulletin de Méthodologie Sociologique, 66*, 60-75.

Borgers, N., & Hox, J. (2001). Item nonresponse in questionnaire research with children. *Journal of Official Statistics, 17*, 321-335.

Bowman, K. R., Penne, M. A., Chromy, J. R., & Odom, D. M. (2001). 1999 National Household Survey on Drug Abuse: Sample design report. In *1999 National Household Survey on Drug Abuse: Methodological resource book* (Section 2, prepared for the Substance Abuse and Mental Health Services Administration under Contract No. 283-98-9008, Deliverable No. 10, RTI/7190.130.400). Research Triangle Park, NC: RTI International.

Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin, 113*, 403-439.

Chromy, J. R. (1981). Variance estimators for a sequential sample selection procedure. In D. Krewski, R. Platek, & J. N. K. Rao (Eds.), *Current topics in survey sampling: Proceedings of the symposium held May 7-9, 1980, sponsored by the Ottawa Chapter and the Survey Research Methods Section of the American Statistical Association* (pp. 329-347). New York: Academic Press.

Chromy, J. R. (1987). Design optimization with multiple objectives. In *Proceedings of the 1987 American Statistical Association, Survey Research Methods Section, San Francisco, CA* (pp. 194-199). Alexandria, VA: American Statistical Association. [Available as a PDF at http://www.amstat.org/sections/SRMS/proceedings/]

Chromy, J. R., Eyerman, J., Odom, D., McNeeley, M. E., & Hughes, A. (2005). Association between interviewer experience and substance use prevalence rates in NSDUH. In J. Kennet & J. Gfroerer (Eds.), *Evaluating and improving methods used in the National Survey on Drug Use and Health* (HHS Publication No. SMA 05-4044, Methodology Series M-5, pp. 59-88). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

Chromy, J. R., & Penne, M. A. (2002). Pair sampling in household surveys. In *Proceedings of the 2002 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, New York, NY* (pp. 552-554). Alexandria, VA: American Statistical Association. [Available at http://www.amstat.org/sections/SRMS/Proceedings/]

Cochran, W. G. (1963). *Sampling techniques* (2nd ed.). New York: 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.

Davies, P. S., & Rupp, K. (2005-2006). An overview of the National Survey of SSI Children and Families and related products. *Social Security Bulletin, 66*(2), 7-20.

Davis, J. A., Smith, T. W., & Marsden, P. V. (2009). *General Social Surveys, 1972-2008: Cumulative codebook*. Chicago, IL: National Opinion Research Center, National Data Program for the Social Sciences. [Available at http://www.norc.org/GSS+Website/Documentation/]

Entwisle, B. (2007). Hard choices: Reflections on the design of the General Social Survey. In *The General Social Survey: The next decade and beyond (Workshop on Planning for the Future of the GSS, Arlington, Virginia, Room 515, Stafford II, May 2-3, 2007)* (pp. 36-39). Arlington, VA: National Science Foundation. [Available as a PDF at http://www.nsf.gov/pubs/2007/nsf0748/nsf0748.jsp]

Eyerman, J., Bowman, K., Butler, D., & Wright, D. (2005). The differential impact of incentives on cooperation and data collection costs: Results from the 2001 National Household Survey on Drug Abuse incentive experiment. *Journal of Economic and Social Measurement, 30*(2-3), 157-169.

Finkelhor, D., Hamby, S. L., Ormrod, R., & Turner, H. (2005). The Juvenile Victimization Questionnaire: Reliability, validity, and national norms. *Child Abuse & Neglect, 29*, 383-412.

Gfroerer, J. C., Wu, L.-T., & Penne, M. A. (2002). *Initiation of marijuana use: Trends, patterns, and implications* (HHS Publication No. SMA 02-3711, Analytic Series A-17). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

Goodman, R. (1999). The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 40*, 791-799.

Groves, R. M., Mosher, W. D., Lepkowski, J., & Kirgis, N. G. (2009). Planning and development of the continuous National Survey of Family Growth. *Vital and Health Statistics, Series 1*(48), 1-64.

Hamby, S. L., & Finkelhor, D. (2000). The victimization of children: Recommendations for assessment and instrument development. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 829-840.

Harrison, L. D., Martin, S. S., Enev, T., & Harrington, D. (2007). *Comparing drug testing and self-report of drug use among youths and young adults in the general population* (HHS Publication No. SMA 07-4249, Methodology Series M-7). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

Horwitz, S. M., Hoagwood, K., Stiffman, A. R., Summerfeld, T., Weisz, J. R., Costello, E. J., Rost, K., Bean, D. L., Cottler, L., Leaf, P. J., Roper, M., & Norquist, G. (2001). Reliability of the services assessment for children and adolescents. *Psychiatric Services, 52*, 1088-1094.

Huang, L. X., Schildhaus, S., & Wright, D. (1999, September). *Seasonality of substance use: National Household Survey on Drug Abuse 1992-96* (OAS working paper). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

Jensen, P. S., Rubio-Stipec, M., Canino, G., Bird, H. R., Dulcan, M. K., Schwab-Stone, M. E., & Lahey, B. B. (1999). Parent and child contributions to diagnosis of mental disorder: Are both informants always necessary? *Journal of the American Academy of Child and Adolescent Psychiatry, 38*, 1569-1579.

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., Williamson, D., & Ryan, N. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): Initial reliability and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 980-988.

Kennet, J., Gfroerer, J., Barker, P., Piper, L., Chromy, J. R., Hirsch, E. L., & Granger, R. K. (2010, May). Reliability and data quality in the National Survey on Drug Use and Health. In L. A. Aday & M. Cynamon (Eds.), *Ninth Conference on Health Survey Research Methods: Conference proceedings* (pp. 107-123). Hyattsville, MD: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Center for Health Statistics. [Available as a PDF at http://www.cdc.gov/nchs/data/misc/proceedings_hsrm2010.pdf]

Kessler, R. C., Avenevoli, S., Costello, E. J., Green, J. G., Gruber, M. J., Heeringa, S., Merikangas, K. R., Pennell, B. E., Sampson, N. A. , & Zaslavsky, A. M. (2009). National Comorbidity Survey Replication Adolescent supplement (NCS-A): II. Overview and design. *Journal of the American Academy of Child and Adolescent Psychiatry, 48*, 380-385.

Kirgis, N., & Lepkowski, J. (2010, June). *A management model for continuous data collection: Reflections from the National Survey of Family Growth, 2006-2010* (NSFG Survey Methodology Working Papers, Report 10-011). Ann Arbor, MI: University of Michigan, Institute for Social Research. [Available as a PDF at http://www.psc.isr.umich.edu/pubs/pdf/ng10-011.pdf]

Kish, L. (1965). *Survey sampling*. New York: John Wiley & Sons.

Knight, E. D., Runyan, D. K., Dubowitz, H., Brandford, C., Kotch, J., Litrownik, A., & Hunter, W. (2000). Methodological and ethical challenges associated with child self-report of maltreatment: Solutions implemented by the LongSCAN Consortium. *Journal of Interpersonal Violence, 15*, 760-775. doi: 10.1177/088626000015007006.

Lepkowski, J. M., Mosher, W. D., Davis, K. E., Groves, R. M., & Van Hoewyk, J. (2010). The 2006-2010 National Survey of Family Growth: Sample design and analysis of a continuous survey. *Vital and Health Statistics, Series 2*(150), 1-36.

Mann, S. L., Lynn, D. J., & Peterson. Jr., A. V. (2008). The "downstream" effect of token prepaid cash incentives to parents on their young adult children's survey participation. *Public Opinion Quarterly, 72*, 487-501. doi: 10.1093/poq/nfn027.

Merikangas, K., Avenevoli, S., Costello, J., Koretz, D., & Kessler, R. C. (2009). National Comorbidity Survey Replication Adolescent supplement (NCS-A): I. Background and measures. *Journal of the American Academy of Child and Adolescent Psychiatry, 48*, 367-369.

Morton, K. B., Chromy, J. R., & Hirsch, E. L. (2008a). *2009 National Survey on Drug Use and Health: Sample design plan* (prepared for the Substance Abuse and Mental Health Services Administration under Contract No. 283-2004-00022, RTI/0209009). Research Triangle Park, NC: RTI International.

Morton, K. B., Martin, P. C., Hirsch, E. L., & Chromy, J. R. (2008b, January). Sample design report. In *2007 National Survey on Drug Use and Health: Methodological resource book* (Section 2, prepared for the Substance Abuse and Mental Health Services Administration under Contract No. 283-2004-00022, Phase III, Deliverable No. 8, RTI/0209009.330.004). Research Triangle Park, NC: RTI International.

Nelson, K. P., & McGough, D. T. (1980). Redesigning the Annual Housing Survey: Policy considerations. In *Proceedings of the 1980 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Houston, TX* (pp. 109-114). Alexandria, VA: American Statistical Association. [Available as a PDF at http://www.amstat.org/sections/SRMS/proceedings/]

Office of Applied Studies. (2008). *Results from the 2007 National Survey on Drug Use and Health: National findings* (HHS Publication No. SMA 08-4343, NSDUH Series H-34). Rockville, MD: Substance Abuse and Mental Health Services Administration.

Office of Applied Studies. (2010). *Results from the 2009 National Survey on Drug Use and Health: Detailed tables*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Peytcheva, E., & Currivan, D. (2009). *2008 National Survey on Drug Use and Health: Summary report of the 2008 Field Supervisor Survey* (prepared for the Substance Abuse and Mental Health Services Administration under Contract No. 283-2004-00022). Research Triangle Park, NC: RTI International.

Peytcheva, E., Currivan, D., & Cunningham, D. (2010). *Analysis of NSDUH field interviewer turnover 2005-2008* (prepared for the Substance Abuse and Mental Health Services Administration under Contract No. 283-2004-00022). Research Triangle Park, NC: RTI International.

Peytcheva, E., Murphy, J., Cunningham, D., & Currivan, D. (2009). *2008 National Survey on Drug Use and Health: Summary report of the 2008 Field Interviewer Survey* (prepared for the Substance Abuse and Mental Health Services Administration under Contract No. 283-2004-00022). Research Triangle Park, NC: RTI International.

Rosenbaum, J., Elliott, M., Kanouse, D., & Schuster, M. (2007). Measuring parent-child agreement, or lack thereof. *Journal of Adolescent Health, 40*(2, Suppl. 1), S30.

RTI International. (2004). *SUDAAN®, Release 9.0.1*. Research Triangle Park, NC: Author.

Russell, K. M., Hudson, M., Long, A., & Phipps, S. (2006). Assessment of health-related quality of life in children with cancer: Consistency and agreement between parent and child reports. *Cancer, 106*, 2267-2274.

Schwab-Stone, M., Fallon, T., Briggs, M., & Crowther, B. (1994). Reliability of diagnostic reporting for children aged 6-11 years: A test-retest study of the Diagnostic Interview Schedule for Children-Revised. *American Journal of Psychiatry, 151*, 1048-1054.

Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M. K., & Schwab-Stone, M. E. (2000). NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 28-38.

Simmons, E., & Wilmot, A. (2004). Incentive payments on social surveys: A literature review. *Social Survey Methodology Bulletin, 53*, 1-11.

Singer, E., Gebler, N., Raghunathan, T., Van Hoewyk, J., & McGonagle, K. (1999). The effect of incentives on response rates in interviewer-mediated surveys. *Journal of Official Statistics, 15*, 217-230.

Stueve, A., & O'Donnell, L. (2000). Inconsistencies over time in young adolescents' self-reports of substance use and sexual intercourse. *Substance Use & Misuse, 35*, 1015-1034.

U.S. Census Bureau. (2006). *Current Population Survey: Design and methodology* (Technical Paper [TP] 66). Washington, DC: Author. [Available as a PDF at http://www.census.gov/prod/2006pubs/tp-66.pdf]

U.S. Census Bureau. (2009). *American Community Survey: Design and methodology* (ACS-DM1). Washington, DC: Author. [Available as a PDF at http://www.census.gov/acs/www/methodology/methodology_main/]

U.S. Census Bureau. (2012, January 31). *State & County QuickFacts: Kalawao County, Hawaii*. Retrieved February 9, 2012, from http://quickfacts.census.gov/qfd/states/15/15005.html

Varni, J. W., Limbers, C. A., & Burwinkle, T. M. (2007). Parent proxy-report of their children's health-related quality of life: An analysis of 13,878 parents' reliability and validity across age subgroups using the PedsQL™ 4.0 Generic Core Scales. *Child: Care, Health and Development, 33*, 649-650. doi:10.1186/1477-7525-5-2.

Weber, L. R., Miracle, A., & Skehan, T. (1994). Interviewing early adolescents: Some methodological considerations. *Human Organization, 53*(1), 42-47.

Whiteman, D., & Green, A. (1997). Wherein lies the truth? Assessment of agreement between parent proxy and child respondents. *International Journal of Epidemiology, 26*, 855-859.

Appendix A: Explaining the Variance Models

Parametric variance models allow a sampling statistician to represent the variance of key estimates as a function of sample design parameters. The 2007 National Survey on Drug Use and Health (NSDUH) data were used to estimate some of the key population parameters needed for these models. The required parameters include the following:

- variance components;
- unequal weighting effects (UWEs), national and by State;
- averages and coefficients of variation for respondents per segment;
- averages and coefficients of variation for respondents per dwelling unit; and
- prevalence estimates by age group.

A.1 Variance Components

The notation for the variance components associated with State, primary sampling unit (PSUs, regional stratum), segment, dwelling unit, and person is shown in Table A1.

Table A1 Variance Component Notation

Level	Symbol
State	σ^2_{state}
Primary Sampling Unit/Regional Stratum*	σ^2_{PSU}
Segment	$\sigma^2_{segment}$
Dwelling Unit	σ^2_{DU}
Person	σ^2_{person}
Total	$\sigma^2 = \sigma^2_{state} + \sigma^2_{PSU} + \sigma^2_{segment} + \sigma^2_{DU} + \sigma^2_{person}$

* The regional strata in the current design are the PSUs in a national design.

Variance components were estimated using the method of moments in SAS PROC NESTED with equal weights. Variance components were computed for 11 substance use and treatment variables of interest, controlling for age group. Variance components for the 11 selected measures are shown in Table A2.

Table A2 Variance Components of Residuals as a Percentage of Total Variance for Selected Measures, Based on 2007 NSDUH

Variable	Mean	σ^2_{state}	σ^2_{PSU}	$\sigma^2_{segment}$	σ^2_{DU}	σ^2_{person}
SMI Proxy, 18+	0.1092	0.1792	0.0000	0.8217	5.3469	93.6522
Past Year MDE, 18+	0.0746	0.1503	0.0000	0.7603	9.7051	89.3842
Past Month Alcohol Use, 12+	0.5114	1.2160	1.3526	2.8701	20.9771	73.5842
Past Month Binge Alcohol Use, 12+	0.2331	0.5051	0.7855	1.6434	19.0616	78.0044
Past Month Marijuana Use, 12+	0.0583	0.6064	0.6175	1.0939	17.7466	79.9357
Past Month Cigarette Use, 12+	0.2424	0.6066	0.8255	2.2496	25.6162	70.7021
Past Month Pain Reliever Use, 12+	0.0209	0.1067	0.1332	0.6198	4.2437	94.8966
Past Year Alcohol Use Disorder, 12+	0.0752	0.1970	0.4292	0.5913	12.6936	86.0890
Past Year Illicit Drug Use Disorder, 12+	0.0276	0.0888	0.2135	0.3413	9.1019	90.2544
Substance Use Disorder, 50+	0.0371	0.0378	1.2822	9.5271	0.0000	89.1530
Past Year Specialty Substance Use Treatment, 12+	0.0097	0.1020	0.0000	0.3457	15.0642	84.4882

MDE = major depressive episode; SMI = serious mental illness.

A.2 Unequal Weighting Effects

The UWE for an age group a, is

$$\text{UWE}_a = \frac{n_a \sum_{i \in a} w_i^2}{\left(\sum_{i \in a} w_i\right)^2}, \tag{A1}$$

where w_i is the analysis weight of respondent i, and a is the set of respondents in age group a. These were calculated for various age groups both nationally and by State. For a given age group a, the weighted averages of the State-specific UWE_a (weighted by the population size of the State) is denoted UWE_a^A.

The national UWE_a, denoted UWE_a^{US}, is used in the variance model for the U.S.-level estimates under the current design. The analysis weights vary across States because of the State-specific sample targets. With a national design, that variability is removed, and the UWE_a^A is used instead.

Table A3 shows the UWEs for the 2007 national estimates by age group.

Table A3 Unequal Weighting Effects for the National (U.S.-Level) Estimates, Based on 2007 NSDUH

Age Group	UWE^{US}	UWE^{A}
12 or Older	3.586	2.671
12 to 17	1.738	1.271
18 to 25	1.757	1.293
26 to 34	1.755	1.305
35 to 49	1.628	1.212
50 or Older	1.667	1.192

A.3 Cluster Sizes and Cluster Size Variation

The parametric models also require the average cluster size and the coefficient of variation of the cluster size at the segment level and at the dwelling unit level, as shown in Table A4 for the domains of interest to this study.

Table A4 Average Cluster Sizes and Coefficients of Variation, Based on 2007 NSDUH

Age Group Domain (δ)	Population Size ($\sum_{i\in\delta} w_i$)	N_δ	Respondents/DU[1]		Respondents/Segment	
			$\bar{m}_{\delta,DU}$	$CV_{m_{\delta,DU}}$	$\bar{m}_{\delta,seg}$	$CV_{m_{\delta,seg}}$
12 or Older	247,845,208	67,870	0.4797	1.5180	9.4264	0.5804
12 to 17	25,241,088	22,433	0.1586	2.7205	3.1157	0.8168
12 to 20	38,475,786	31,175	0.2203	2.3319	4.3299	0.7636
18 to 25	32,730,854	22,187	0.1568	2.7844	3.0815	1.0617
26 to 34	35,300,404	6,908	0.0488	4.8620	0.9594	1.3283
50 or Older	89,729,008	6,366	0.0450	5.0413	0.8842	1.2866
18 or Older	222,604,119	45,437	0.3212	1.8350	6.3107	0.6835
12 or Older API	12,060,590	3,079	0.02176	8.0462	0.4276	2.9876
12 or Older AIAN	2,925,989	3,163	0.02236	7.6888	0.4393	2.7723
12 or Older Pregnant	2,624,322	1,123	0.0079	11.2598	0.1560	2.6798

AIAN = American Indian or Alaska Native; API = Asian or Pacific Islander; DU = dwelling unit.

[1] The total number of dwelling units in the sample is actually more than the number of respondents. This does not matter because the key statistic has the form $\bar{m}_{\delta,DU}(1+CV^2_{\delta,DU})$ (see Equation A2 in Section A.4), which is insensitive to the number of dwelling units in the sample.

A.4 Variance Models

A.4.1 A National Design

Let δ be a domain of interest (e.g., 12 to 20, 12 or older Asian or Pacific Islander [API], 18 or older). The parametric variance model of an estimated proportion for the domain under a proposed national design is as follows:

$$Var_{nat}(\hat{p} \mid n = 40,000) = \qquad (A2)$$

$$\sum_{relevant\ a} W_{a\delta}^2 \frac{p_a(1-p_a)}{f_a c_{a\delta} 40,000} \mathrm{UWE}_a{}^A \left\{ (8\sigma_{PSU}^2 + \sigma_{seg}^2)\, \bar{m}_{\delta,seg} (1 + CV_{m_{\delta,seg}}^2) + \sigma_{DU}^2 \bar{m}_{\delta,DU} (1 + CV_{\delta,DU}^2) + \sigma_{person}^2 \right\},$$

where

a = an age group (e.g., *relevant a* means that the 26 or older age group is excluded when the domain is 18 to 20 year olds);

p_a = the estimated (domain) proportion of interest within age group a;

$W_{a\delta}$ = the (estimated) population share of age group a within domain δ, that is, $\sum_{i \in a \cap \delta} w_i \big/ \sum_{i \in \delta} w_i$;

f_a = the proposed sampling fraction for age group a; and

$c_{a\delta}$ = usually the fraction of the population in age group a that is also in domain δ, that is, $\sum_{i \in a \cap \delta} w_i \big/ \sum_{i \in a} w_i$ as explained below.

The value of $c_{a\delta}$ is 1 except when the domain is API, American Indian or Alaska Native (AIAN), pregnant women, or 12 to 20 year olds. For the 12 to 20 domain, .375 is used as the fraction of 18 to 20 year olds within the 18 to 25 age group.

Table A5b displays the realized fractions from the 2007 NSDUH sample analogous to the population fractions in Table A5a. The two sets of numbers are generally close, with some exceptions. The sample appears to have more than 3 times the relative number of API persons than the population. This is likely due to larger-than-their-size-would-dictate NSDUH samples in States like Oklahoma, Arizona, Alaska, and New Mexico.

Table A5a 2007 Population Fraction of an Age Group Originating in Selected Domains: $\sum_{i \in a \cap \delta} w_i / \sum_{i \in a} w_i$

Age Group	Domain		
	Pregnant Women	**AIAN**	**API**
12 to 17	0.0040	0.0157	0.0488
18 to 25	0.0299	0.0136	0.0536
26 to 34	0.0341	0.0134	0.0654
35 to 49	0.0052	0.0114	0.0538
50 or Older	0.0000	0.0097	0.0365

AIAN = American Indian or Alaska Native; API = Asian or Pacific Islander.

Table A5b 2007 Sample Fraction of an Age Group Originating in Selected Domains: $\sum_{i \in a \cap \delta} 1 / \sum_{i \in a} 1$

Age Group	Domain		
	Pregnant Women	**AIAN**	**API**
12 to 17	0.0041	0.0528	0.0444
18 to 25	0.0325	0.0492	0.0498
26 to 34	0.0363	0.0537	0.0553
35 to 49	0.0058	0.0378	0.0408
50 or Older	0.0000	0.0219	0.0300

AIAN = American Indian or Alaska Native; API = Asian or Pacific Islander.

The variance model in Equation (A2) assumes a probability proportional to size (PPS) sequential sample of PSUs with eight sampled segments per PSU and 9.375 respondents per segment. Because every State except Wyoming is certain to be in the sample, the State variance component (σ^2_{state}) has not been included in Equation (A2). It has been assumed that the underestimation of the variance that this may cause is compensated for by the equation not capturing the benefits at the PSU/region level from the implicit stratification of systematic sampling.[25]

For a national design with 3 times as many targeted respondents per segments, the $\overline{m}_{\delta,seg}$ in Equation (A2) can be multiplied by 3. This has the effect of reducing the number of selected PSUs to a third of the original number.

A.4.2. 50-State Design

The parametric variance model for a U.S.-level estimate under the current 50-State design, with differing age group targets, is as follows:

[25] In fact, the inclusion of the $8\sigma^2_{PSU}$ term in Equation (A2) will likely make the variance prediction too high. When it comes to actually estimating the variance from a national sample, variance strata can be created to capture most of the variance reduction because of systematic sampling.

$$Var_{st}(\hat{p} \mid n = 67{,}500) = \qquad \text{(A3)}$$

$$\sum_{relevant\ a} W_{a\delta}^2 \frac{p_a(1-p_a)}{f_a \tilde{c}_{a\delta} 67500} \text{UWE}_a^{US} \left\{\sigma_{seg}^2\, \bar{m}_{\delta,seg}(1 + CV_{m_{\delta,seg}}^2) + \sigma_{DU}^2 \bar{m}_{\delta,DU}(1 + CV_{\delta,DU}^2) + \sigma_{person}^2\right\},$$

where

$\tilde{c}_{a\delta}$ = usually 1 except in selected domains (in most of these exceptions, it has the values displayed in Table A5b).

On the one hand, there is a missing variance component in Equation (A3) (σ_{PSU}^2), which reduces the variance. On the other, the UWEs (the UWE_a^{US}) are higher.

A.4.3 Hybrid Design

A hybrid design combines the national design shown above with the current design in States with fewer than *Q* expected respondents under the national design. The size of *Q* is to be chosen so that the effective sample size for estimating the fraction of past month marijuana users among persons aged 12 or older is 400. The effective sample size in a State is the actual sample size divided by the (estimated) design effect. For a generic State, the design effect is as follows:

$$deff_s = \frac{\sum_{relevant\ a} \left(\frac{1}{f_a}\right) W_{a12+}^2\, p_a(1-p_a) \text{UWE}_a^A \left\{\sigma_{seg}^2 \bar{m}_{12+,seg}(1 + CV_{m_{12+,seg}}^2) + \sigma_{DU}^2 \bar{m}_{12+,DU}(1 + CV_{m_{12+,seg}}^2) + \sigma_{person}^2\right\}}{\left(\sum_{relevant\ a} W_{a12+}\right)^2 p_{12+}(1-p_{12+})}. \qquad \text{(A4)}$$

Equation (A4) assumes that 3.375 respondents are targeted per segment as in the current design. For a design with 3 times as many respondents per segment, the $\bar{m}_{12+,seg}$ in Equation (A4) is multiplied by 3. Table 3.15 in Section 3.6 of Chapter 3 shows the design effects that were computed using this analysis.

The estimated variance under a hybrid design is as follows:

$$Var_{hyb}(\hat{p}) =$$

$$\left(\frac{N_1}{N}\right)^2 \sum_{relevant\ a} W_{a\delta}^2 \frac{p_a(1-p_a)}{f_a c_{a\delta} n_1} \text{UWE}_a^A \left\{(8\sigma_{PSU}^2 + \sigma_{seg}^2)\, \bar{m}_{\delta,seg}(1 + CV_{m_{\delta,seg}}^2) + \sigma_{DU}^2 \bar{m}_{\delta,DU}(1 + CV_{\delta,DU}^2) + \sigma_{person}^2\right\} + \qquad \text{(A5)}$$

$$\left(\frac{N_2}{N}\right)^2 \sum_{relevant\ a} W_{a\delta}^2 \frac{p_a(1-p_a)}{f_a \tilde{c}_{a\delta} n_2} \text{UWE}_a^{US} \left\{\sigma_{seg}^2\, \bar{m}_{\delta,seg}(1 + CV_{m_{\delta,seg}}^2) + \sigma_{DU}^2 \bar{m}_{\delta,DU}(1 + CV_{\delta,DU}^2) + \sigma_{person}^2\right\},$$

where

$\bar{m}_{\delta,seg}$ = the average cluster size and is multiplied by 3 when there are 3 times as many targeted respondents in a segment as in the current design;

$n_1(N_1)$ = the size of the sample (population) covering the national component of the sample; and

$n_2(N_2)$ = the size of the sample (population) covering the State component.

An empirical investigation revealed that the population fractions of age groups within selected domains among the larger States (e.g., the largest 14) were roughly equal to those for the Nation, whereas the sampling fractions among the remaining States were roughly equal to the sampling fractions for the Nation (neither shown). This explains the use of $c_{a\delta}$ and $\tilde{c}_{a\delta}$ in Equation (A5).

Appendix B: Investigating Several New Options

At the request of the Substance Abuse and Mental Health Services Administration (SAMHSA), additional analyses were conducted for the National Survey on Drug Use and Health (NSDUH) on several new 50-State design options with State allocation different from the current design. These were grouped under the general headings of Options A, B, and C and had the parameters specified in Table B1. Within the tables presented in this appendix, *n* represents the targeted respondent sample size and *m* the average number of respondents per cluster.

Table B1 Summary of Design Parameters for Design Options A (*n* = 58,400), B (*n* = 67,500), and C (*n* = 67,521)

Design Parameter	Design Option A			Design Option B				Design Option C
	***m* = 9.375**	***m* = 18.75**	***m* = 28.125**	***m* = 9.375**	***m* = 18.75**	***m* = 28.125**	***m* = variable**[1]	***m* = 10.420**
Age Group Sample Size								
12 to 17	14,600	14,600	14,600	16,875	16,875	16,875	16,875	16,880
18 to 25	14,600	14,600	14,600	16,875	16,875	16,875	16,875	16,880
26 to 34	8,760	8,760	8,760	10,125	10,125	10,125	10,125	10,128
35 to 49	11,680	11,680	11,680	13,500	13,500	13,500	13,500	13,505
50 or Older	8,760	8,760	8,760	10,125	10,125	10,125	10,125	10,128
Total	58,400	58,400	58,400	67,500	67,500	67,500	67,500	67,521
State Allocation								
CA	4,800	4,800	4,800	4,600	4,600	4,600	4,600	4,501
TX, NY, FL	3,200	3,200	3,200	3,300	3,300	3,300	3,300	3,334
IL, PA, OH, MI	1,600	1,600	1,600	2,000	2,000	2,000	2,000	2,001
GA, NC, NJ, VA	1,600	1,600	1,600	1,500	1,500	1,500	1,500	1,500
Remainder	800	800	800	1,000	1,000	1,000	1,000	1,000
Other								
State Sampling (SS) Regions	779	389	260	900	450	300	652	810
Segments	6,229	3,115	2,076	7,200	3,600	2,400	5,216	6,480
Expected Completed Screening Interviews	90,871	90,871	90,871	105,031	105,031	105,031	105,031	105,043
Expected Completed Interviews	58,400	58,400	58,400	67,500	67,500	67,500	67,500	67,521

Note: Allocation for each new option: 12 to 17: 25 percent; 18 to 25: 25 percent; 26 to 34: 15 percent; 35 to 49: 20 percent; and 50 or older: 15 percent.

[1] *m* = 20.536 in California; 20.625 in Texas, New York, and Florida; 20.833 in Illinois, Pennsylvania, Ohio, and Michigan; 15.625 in Georgia, North Carolina, New Jersey, and Virginia; and 10.417 in the remaining 38 States and the District of Columbia.

SAMHSA also requested that two frame sources be evaluated:

- 100 percent field enumeration (FE) (50 percent overlap of segments annually, as is done currently), and
- address-based sampling (ABS) with a 50 percent coverage threshold (updated annually with no overlap).

Option B with variable cluster sizes and Option C were evaluated in terms of precision, cost, field management, and response rates. The remaining design options were evaluated in terms of the following:

- precision,
- cost,
- field management,
- small area estimation (SAE) and substate estimates,
- response rates, and
- interviewer effects on prevalence.

B.1 Precision

For all options except Option B with variable cluster sizes, variance estimates for the same measures as in the text were computed using the following modification of Equation (A3) (see Appendix A):

$$Var_{st}(\hat{p} \mid n_{option}) = \sum_{relevant\ a} W_{a\delta}^2 \frac{p_a(1-p_a)}{f_a \tilde{c}_{a\delta} n_{option}} \mathrm{UWE}_a^{option} \left\{ \sigma_{seg}^2 \bar{m}_{\delta,seg}(1 + CV_{m_{\delta,seg}}^2) + \sigma_{DU}^2 \bar{m}_{\delta,DU}(1 + CV_{\delta,DU}^2) + \sigma_{person}^2 \right\}, \tag{B1}$$

where

n_{option} = 58,400 for Option A, 67,500 for Option B, and 67,521 for Option C.

$$\mathrm{UWE}_a^{option} = n_{option} \sum_{all\ states} \frac{1}{n_{option}^{state}} \left(\frac{\sum_{k \in a, state} w_k}{\sum_{k \in a, all\ states} w_k} \right)^2 \mathrm{UWE}_a^{state}, \tag{B2}$$

where

n_{option}^{state} = allocation to a particular State under a particular option, and

UWE_a^{state} = current unequal weighting effect (UWE) in a specific State for age group a.

For Option B with variable cluster sizes, the following equation was used:

$$Var_{st}(\hat{p} \mid n = 67{,}500) = \sum_{g=1}^{5} \left[\sum_{relevant\ a} W_{a\delta(g)}^2 \frac{p_a(1-p_a)}{f_a \tilde{c}_{a\delta} n_{B-\text{var}}^{(g)}} \text{UWE}_a^{(g)} \left\{ \sigma_{seg}^2 \, \overline{m}_{\delta,seg}^{(g)} (1 + CV_{m_{\delta,seg}}^2) + \sigma_{DU}^2 \overline{m}_{\delta,DU} (1 + CV_{\delta,DU}^2) + \sigma_{person}^2 \right\} \right], \quad \text{(B3)}$$

where

g = one of five State groups (1: California; 2: New York, Texas, and Florida; 3: Illinois, Ohio, Michigan, and Pennsylvania; 4: Georgia, North Carolina, New Jersey, and Virginia; and 5: remaining States and District of Columbia) such that each State in a group has the same allocation and the same expected respondent sample size per segment;

$n_{B-\text{var}}^{(g)}$ = target respondent sample size across primary sampling units (PSUs) in State group g ($n_{B-\text{var}}^{(1)} = 4{,}600$; $n_{B-\text{var}}^{(2)} = 9{,}900$; $n_{B-\text{var}}^{(3)} = 8{,}000$; $n_{B-\text{var}}^{(4)} = 6{,}000$; $n_{B-\text{var}}^{(6)} = 39{,}000$);

$\overline{m}_{\delta,seg}^{(g)} = \frac{z_g}{9.375} \overline{m}_{\delta,seg}$, z_g = target respondent sample size per segment in State group g ($z_1 = 20.536$; $z_2 = 20.625$; $z_3 = 20.833$; $z_4 = 15.625$; $z_5 = 10.417$), while $\overline{m}_{\delta,seg}$ is from Table A3;

$W_{a\delta(g)}^2$ = (estimated) population share of age group a and State group g within domain δ (i.e., $\sum_{i \in a \cap \delta \cap g} w_i \big/ \sum_{i \in \delta} w_i$); and

$$\text{UWE}_a^{(g)} = \frac{\sum_{k \in a \cap State\ group\ g} w_k^2}{\left(\sum_{k \in a \cap State\ group\ g} w_k\right)^2} \times$$ 2007 number of respondents in State group g and age group a.

Precision results are displayed in Tables B2 and B3 for Option A and variants of Option B with cluster sizes of 9.375, 18.75, and 28.125. They are the same whether 100 percent FE or partial ABS is used, as are the results, which are not displayed.

Table B2 Estimated Prevalence Estimates and Relative Standard Errors for Option A

Measure	Domain	Prevalence	Relative Standard Error (RSE)		
			Cluster Size		
			9.375	18.750	28.125
Past Year Alcohol Use Disorder	12+	0.0752	0.0199	0.0205	0.0211
Substance Use Disorder	50+	0.0371	0.0686	0.0751	0.0812
Past Year Illicit Drug Use Disorder	12+	0.0276	0.0296	0.0302	0.0307
Past Month Alcohol Use	12-20	0.2789	0.0107	0.0116	0.0123
Past Month Alcohol Use	50+	0.4692	0.0139	0.0143	0.0148
Past Month Alcohol Use	AIAN, 12+	0.4181	0.0276	0.0289	0.0302
Past Month Alcohol Use	12+	0.5114	0.0070	0.0078	0.0085
Past Month Alcohol Use	API,12+	0.3670	0.0286	0.0301	0.0314
Past Month Alcohol Use	Pregnant, 12+	0.1197	0.0830	0.0846	0.0861
Past Month Pain Reliever Use	18-25	0.0464	0.0480	0.0489	0.0498
Past Month Pain Reliever Use	12+	0.0209	0.0373	0.0386	0.0398
Past Month Binge Alcohol Use	18-25	0.4178	0.0131	0.0137	0.0143
Past Month Binge Alcohol Use	12+	0.2331	0.0115	0.0124	0.0132
Past Month Cigarette Use	12-17	0.0983	0.0337	0.0353	0.0369
Past Month Cigarette Use	12+	0.2424	0.0121	0.0133	0.0144
Past Year MDE (18+)	18+	0.0788	0.0171	0.0176	0.0181
Past Month Marijuana Use	12-17	0.0668	0.0402	0.0413	0.0423
Past Month Marijuana Use	18-25	0.1643	0.0246	0.0254	0.0261
Past Month Marijuana Use	50+	0.0154	0.1038	0.1051	0.1063
Past Month Marijuana Use	AIAN, 12+	0.0729	0.0608	0.0619	0.0631
Past Month Marijuana Use	12+	0.0583	0.0218	0.0230	0.0241
Past Month Marijuana Use	API,12+	0.0279	0.0979	0.0999	0.1019
Past Month Marijuana Use	Pregnant, 12+	0.0404	0.1336	0.1345	0.1354
SMI Proxy (18+)	18+	0.1192	0.0130	0.0134	0.0138
Past Year Specialty Substance Use Treatment	12+	0.0097	0.0603	0.0615	0.0623

AIAN = American Indian or Alaska Native; API = Asian or Other Pacific Islander; MDE = major depressive episode; SMI = serious mental illness.

Table B3 Estimated Prevalence Estimates and Relative Standard Errors for Option B

Measure	Domain	Prevalence	Relative Standard Error (RSE)		
			Cluster Size		
			9.375	18.750	28.125
Past Year Alcohol Use Disorder	12+	0.0752	0.0188	0.0195	0.0200
Substance Use Disorder	50+	0.0371	0.0648	0.0710	0.0767
Past Year Illicit Drug Use Disorder	12+	0.0276	0.0281	0.0286	0.0292
Past Month Alcohol Use	12-20	0.2789	0.0102	0.0110	0.0117
Past Month Alcohol Use	50+	0.4692	0.0131	0.0136	0.0140
Past Month Alcohol Use	AIAN, 12+	0.4181	0.0261	0.0274	0.0286
Past Month Alcohol Use	12+	0.5114	0.0066	0.0074	0.0081
Past Month Alcohol Use	API,12+	0.3670	0.0271	0.0285	0.0298
Past Month Alcohol Use	Pregnant, 12+	0.1197	0.0789	0.0804	0.0819
Past Month Pain Reliever Use	18-25	0.0464	0.0456	0.0465	0.0473
Past Month Pain Reliever Use	12+	0.0209	0.0354	0.0366	0.0378
Past Month Binge Alcohol Use	18-25	0.4178	0.0124	0.0130	0.0135
Past Month Binge Alcohol Use	12+	0.2331	0.0109	0.0117	0.0125
Past Month Cigarette Use	12-17	0.0983	0.0320	0.0335	0.0350
Past Month Cigarette Use	12+	0.2424	0.0115	0.0126	0.0136
Past Year MDE (18+)	18+	0.0788	0.0162	0.0167	0.0172
Past Month Marijuana Use	12-17	0.0668	0.0382	0.0392	0.0401
Past Month Marijuana Use	18-25	0.1643	0.0234	0.0241	0.0248
Past Month Marijuana Use	50+	0.0154	0.0981	0.0933	0.1005
Past Month Marijuana Use	AIAN, 12+	0.0729	0.0577	0.0588	0.0599
Past Month Marijuana Use	12+	0.0583	0.0207	0.0218	0.0229
Past Month Marijuana Use	API,12+	0.0279	0.0929	0.0948	0.0967
Past Month Marijuana Use	Pregnant, 12+	0.0404	0.1269	0.1278	0.1287
SMI Proxy (18+)	18+	0.1192	0.0123	0.0127	0.0132
Past Year Specialty Substance Use Treatment	12+	0.0097	0.0572	0.0583	0.0594

AIAN = American Indian or Alaska Native; API = Asian or Other Pacific Islander; MDE = major depressive episode; SMI = serious mental illness.

Table B4 displays the average percentage increase in relative standard error (RSE) of all the new options compared with the current design computed on the log scale so that a 1 percent increase in RSE translates into a 2 percent increase in relative variance or a 2 percent decrease in effective sample size. Averages were taken across the eight variables for the 12 or older domain, the two variables for the 18 or older domain, the two variables for the American Indian or Alaska Native (AIAN) domain, the two variables for the Asian or Other Pacific Islander (API) domain, the three variables for the 50 or older domain (denoted "OLD"), the six variables for the younger than age 26 domain (denoted "YOUNG"), and the two variables for the pregnant women domain, respectively.

For the RSE of a State within one of the five State groups, averages were taken across the eight variables for the 12 or older domain. For the calculation, Equation (B1) was used with UWE_a^A from Table A3 in place of UWE_a^{option} and n_{option} set appropriately.

Table B4 Average Relative Percentage Change in RSE Compared with the Current Design: New Options

Group	Option A (*n* = 58,400)			Option B (*n* = 67,500)				Option C (*n* = 67,521)
	Cluster Size			Cluster Size				Cluster Size
	9.375	18.750	28.125	9.375	18.750	28.125	Variable[1]	10.420
12 or Older	-15	-10	-5	-21	-15	-10	-16	-20
18 or Older	-51	-48	-45	-57	-53	-50	-54	-56
AIAN	-49	-46	-43	-55	-51	-48	-48	-54
API	-48	-45	-41	-53	-50	-47	-46	-53
OLD	-23	-19	-14	-29	-24	-20	-25	-28
YOUNG	10	14	18	5	9	12	13	5
PREGNANT	-44	-42	-41	-49	-47	-46	-47	-48
12 or Older: CA	-14	-9	-4	-12	-7	-2	-6	-11
12 or Older: TX, NY, FL	6	11	16	4	10	15	11	4
12 or Older: IL, PA, OH, MI	41	46	51	29	35	40	36	36
12 or Older: GA, NC, NJ, VA	-29	-23	-19	-26	-20	-15	-22	-25
12 or Older: Remainder	6	11	16	-5	0	5	-5	-5

AIAN = American Indian or Alaska Native; API = Asian or Other Pacific Islander; OLD = 50 or older; YOUNG = younger than 26.

[1] 20.536 for California; 20.625 for Texas, New York, and Florida; 20.833 for Illinois, Pennsylvania, Ohio, and Michigan; 15.625 for Georgia, North Carolina, New Jersey, and Virginia; and 10.417 for the remaining 38 States and District of Columbia.

B.2 Cost Comparisons

Key design parameters and total costs for design Options A, B, and C are compared in Table B5. Cost savings for Year 1 are relative to the current design with 7,200 segments field enumerated with an average cluster size of 9.375. Cost savings for All Other Years are relative to the current design with 3,600 segments field enumerated and 3,600 segments reused with an average cluster size of 9.375. The table shows that the greatest savings will occur by lowering the sample size and increasing the cluster size.

B.3 Field Management Issues

Larger clusters will mean fewer travel points. More field supervisors (FSs) will handle multiple State assignments as the sample size is reduced. An attempt has been made to incorporate the shift to more travel within and less travel between segments in developing the cost model parameters.

Some realignment of field interviewer (FI) staff will be required. States that have an increase in sample size will require recruiting and training new FIs. States that have decreases in sample size will face termination of FIs who have no local work and do not have the ability and performance potential to travel to States where the sample is increased. Once the FI staff stabilizes, it is expected that there will be about the same level of turnover as is experienced under the current design.

Based on all of the above, Table B6 provides the approximate number of FIs, FSs, regional supervisors (RSs), and regional directors (RDs) under design Options A, B, and C.

B.4 Small Area Estimates and Substate Estimates

The new design options with cluster sizes of 9.375 will support State and substate SAE on the same schedule as the current design. The root posterior variance of small area estimates for the two younger age groups will be increased by 10 to 15 percent, as reflected in Table B4. When data are combined over 2 and 3 years to produce small area estimates, the overlapping sample design causes the cluster size to double in one third and one half of the segments, respectively. A hybrid frame consisting of ABS with some field enumeration would maintain the 50 percent overlapping sample design. Thus, the loss of precision caused by pooling data over multiple years and increasing the average cluster size would be roughly the same for the hybrid frame as was previously projected for the field-enumerated frame. If the overlapping sample design were to be eliminated, the m = 18.75 and the 28.125 cluster sizes in design Options A and B would be more efficient for SAE with the exception of the two younger age groups.

B.5 Response Rates

As with the national and hybrid designs, the weighted interview response rates are not expected to be affected by any of the new design options. However, there will be implications for the unweighted response rates as the distribution of sample changes by State and age group.

Table B5 Cost Comparisons for Design Options A, B, and C

	Design Option A			Design Option B				Design Option C
	m = 9.375	*m* = 18.750	*m* = 28.125	*m* = 9.375	*m* = 18.750	*m* = 28.125	Variable[1]	
Design Parameter								
PSUs (SS Regions)	779	389	260	900	450	300	652	810
Segments	6,229	3,115	2,076	7,200	3,600	2,400	5,216	6,480
Completed Screenings	90,871	90,871	90,871	105,031	105,031	105,031	105,031	105,043
Completed Interviews	58,400	58,400	58,400	67,500	67,500	67,500	67,500	67,522
100 Percent FE - Year 1								
Variable Costs ($ in Thousands)	2,566	1,925	1,497	2,966	2,225	1,730	2,149	2,670
Cost Savings[2] ($ in Thousands)	5,821	10,391	12,128	1,486	6,768	8,776	4,806	2,690
100 Percent FE - All Other Years								
Variable Costs ($ in Thousands)	1,283	962	749	1,483	1,112	865	1,074	1,335
Cost Savings[3] ($ in Thousands)	5,621	9,870	11,393	1,486	6,397	8,158	4,397	2,542
ABS 50 Percent Threshold – Year 1								
Variable Costs ($ in Thousands)	293	219	171	338	254	197	245	304
Cost Savings[2] ($ in Thousands)	8,095	12,096	13,454	4,114	8,739	10,309	6,710	5,056
ABS 50 Percent Threshold – All Other Years								
Variable Costs ($ in Thousands)	146	110	85	169	127	99	122	152
Cost Savings[3] ($ in Thousands)	6,758	10,723	12,056	2,800	7,383	8,925	5,349	3,725

ABS = address-based sampling; FE = field enumeration; PSU = primary sampling unit; SS = State sampling.

[1] 20.536 for California; 20.625 for Texas, New York, and Florida; 20.833 for Illinois, Pennsylvania, Ohio, and Michigan; 15.625 for Georgia, North Carolina, New Jersey, and Virginia; and 10.417 for the remaining 38 States and District of Columbia.

[2] Cost savings for Year 1 are relative to the current design with 7,200 segments field enumerated with an average cluster size of 9.375.

[3] Cost savings for All Other Years are relative to the current design with 3,600 segments field enumerated and 3,600 segments reused with an average cluster size of 9.375.

Table B6 Field Staff Size for Design Options A, B, and C

	Current Design	Design Option A			Design Option B				Design Option C
		$m = 9.375$	$m = 18.750$	$m = 28.125$	$m = 9.375$	$m = 18.750$	$m = 28.125$	Variable	
Design Parameter									
PSUs (SS Regions)	900	779	389	260	900	450	300	652	810
Segments	7,200	6,229	3,115	2,076	7,200	3,600	2,400	5,216	6,480
Completed Screenings	140,041	90,871	90,871	90,871	105,031	105,031	105,031	105,031	105,043
Completed Interviews	67,500	58,400	58,400	58,400	67,500	67,500	67,500	67,500	67,522
Field Interviewers (FIs)	680	563	470	439	636	529	493	577	614
Field Supervisors (FSs)	38	31	26	25	36	30	28	32	34
Regional Supervisors (RSs)	7	6	5	5	7	5	5	6	6
Regional Directors (RDs)	3	2	2	2	3	2	2	3	3

PSU = primary sampling unit; SS = State sampling.

B.6 Interview Experience Implications for Transition to Two Optional Designs (Option A and the Fixed Cluster Size Variants of Option B)

The major change in design that occurred from 1998 to 1999 had a significant impact on trend measures. Subsequent analysis of the data identified large changes in the distribution of interviewer experience as a possible explanation for any design impact on trend measures. Changes in design contemplated for future NSDUHs are much less severe than the changes implemented in 1999, but changes in the sample allocation may require a reallocation of interviewers geographically, which may lead to increased attrition in the initial year of implementation with a gradual return to a nearly stable distribution over succeeding years. Both increased attrition of current staff and large expansion in interviewing staff require the use of less experienced interviewers and can create a larger than usual shift in the distribution of completed interviews by interviewer experience at the time of the interview. Based on a study of interviewer experience, the distribution of interviews has been shifting gradually toward more experience, as shown in Table B7. Logistic regression coefficients for the effects of interviewer experience levels on the reported prevalence of past year marijuana are shown in the right column.

Table B7 Interviewer Experience Distribution, 2002 to 2007

Field Interviewer (FI) Experience at Completion of Interview	Interview Distribution, by FI Experience						Past Year Marijuana: Logistic Regression for 2007	
	2002	2003	2004	2005	2006	2007	Coefficients	Odds Ratios
1 to 19 (Base)	5.28	4.32	4.86	5.12	5.93	4.67	0.00000	1.00
20 to 39	5.14	4.30	3.72	4.61	4.41	4.32	-0.09312	0.91
40 to 59	5.18	3.48	3.52	3.74	3.62	3.84	-0.14685	0.86
60 to 99	9.32	7.06	5.70	6.34	6.02	6.09	-0.15538	0.86
100 to 249	32.33	25.12	19.22	16.87	17.27	14.58	-0.15970[1]	0.85
250 to 499	31.65	33.31	30.07	24.49	20.23	17.91	-0.10492	0.90
500 to 749	7.92	15.05	17.69	18.96	16.39	16.00	-0.17699[1]	0.84
750 to 999	1.97	4.25	9.40	9.31	11.40	12.51	-0.23447[1]	0.79
1,000+	1.21	3.11	5.81	10.56	14.73	20.08	-0.15785[1]	0.85

[1] These estimates are significantly different from zero at the .05 level.

Note: The estimated coefficient in the table is a sum; it does not convert into a log odds ratio. The right column shows the sum of the coefficients for categorical interviewer experience. It treats the category 1 to 19 interviews as the base and shows the interaction between that variable and the survey year. The other variables in the model are as follows: (a) average hours per week for the quarter; (b) average miles per week for the quarter; (c) FI gender; (d) race/ethnicity; (e) age group; (f) respondent race/ethnicity; (g) gender; (h) age group; (i) income; (j) interaction between FI and respondent gender; (k) interaction between FI and respondent age; (l) interaction between FI and respondent race/ethnicity; (m) interview nonresponse adjustment factor; (n) ever refused interview at the interview level; (o) percentage of population 65 or older in block group; (p) percentage of population 25 to 34 in block group; (q) percentage of population 35 to 44 in block group; (r) percentage of blacks in block group; (s) percentage of other race in block group; (t) percentage of whites in block group; (u) percentage of female head of household, no spouse, child younger than 18; (v) drug sale/manufacture arrest rate; (w) serious crime rate; (x) percentage of persons 16 to 64 with a work disability; (y) households with public assistance income; (z) percentage of housing units rented; (aa) median rents; (bb) metropolitan statistical area (MSA) status; (cc) survey year; (dd) State; and (ee) interaction between survey year and State.

The 2007 experience distribution was used as a starting point to study the impact of the proposed design options on the distribution of interviews by interviewer experience and its subsequent impact on the reported level of past year marijuana use. It should be noted that with continued application of the current design, the distribution will naturally shift toward more interviewing being conducted by highly experienced interviewers.

The current design sets the sample distribution by age group at one third to age groups 12 to 17, 18 to 25, and 26 or older. Both design options impose a change in the sample distribution to one fourth, one fourth, and one half to the three age groups. This allocation change alone leads to a relative decrease in screening interviews required per completed personal interview and allows some reduction in the size of the interviewing staff. The spreadsheet model assumed a negligible impact on the experience distribution for this design change alone.[26]

In order to compare the six design options involving two levels of change in sample size and three average cluster size options, average interviewer productivity of 81, 96, and 103 personal interviews completed per year along with sample size changes for specific State groupings were used to project new experience distributions. Additional assumptions made in applying the projection model were as follows:

- About one half of all new-to-project (NTP) interviewers remain on the project after 1 year (based on the 2002 to 2009 data).
- For modeling purposes, NTP interviewers were defined by the experience categories 1 to 59.
- Under normal conditions, 6.8 percent of experienced interviewers leave the project each year (based on the 2002 to 2009 data).
- Regression model parameters (see Table B7) were applied to the projected experience distribution to project the "weighted log odds" row (Tables B9 to B13).
- The log odds ratio to the base year was computed as the difference between the weighted log odds for the 2 years (where 2007 is assumed to be the base year and the year of the redesign is the other year). The odds ratio then was obtained by exponentiation.
- A starting value of 10 percent prevalence was used for 2007 for easy comparison and was about right for past year marijuana prevalence for persons aged 12 or older.
- Reference year odds were based on the specified prevalence of 10 percent and adjusted by the odds ratio for each design option.
- The shift in the annual distribution of interviews by experience was based on a transition matrix that aged each group based on increasing its experience by 90 interviews per year and moving part of the group to the next level.

A summary of the projected results is shown in Table B8. Details for each State grouping are in Tables B9 through B13. The impact on national results is obtained in Table B8 as a

[26] This aspect could be modeled more realistically in any updates to these models. For the current design, interviewers are estimated to complete 77 interviews per year.

population weighted average. The national effects are minimal; all modeled prevalence estimates round to 10.0 percent. Extra decimal places are shown to indicate the direction of the change even when the change is small.

A large increase in sample size is proposed for California under both Options A and B. The projected estimate of prevalence will round to 10.2 for the current cluster size of 9.375 when the sample size is increased to 4,800 and to 10.1 in two other cases. Tripling the cluster size to 28.125 persons reduces the impact because increased sample size is offset by increased interviewer productivity (more interviews completed).

For Florida, Texas, and New York, the impact of reducing the sample size is to shift the experience distribution toward more experience and reduce the projected prevalence estimates. All projected prevalence estimates round to 9.9, but the projected difference is less than 0.1 percentage points.

For Illinois, Michigan, Ohio, and Pennsylvania, the model projects no new hires, and the same projected experience distribution applies under all design options considered. This leads to a reduction of less than 0.1 percentage points, but the rounded projections are all at 9.9.

Table B8 Projected Reporting of Past Year Marijuana Use under Six Design Options When Assuming a Rate of 10.000 Percent under the Current Experience Distributions

	Sample Size		Cluster Size		
State/Option	**Current**	**New**	**Current**	**Double**	**Triple**
California					
A	3,600	4,800	10.162	10.072	10.030
B	3,600	4,600	10.142	10.047	10.004
Florida, Texas, New York					
A	3,600	3,200	9.920	9.913	9.914
B	3,600	3,300	9.942	9.913	9.913
Illinois, Michigan, Ohio, Pennsylvania					
A	3,600	1,600	9.916	9.916	9.916
B	3,600	2,000	9.916	9.916	9.916
Georgia, New Jersey, North Carolina, Virginia					
A	900	1,600	10.287	10.218	10.186
B	900	1,500	10.262	10.189	10.155
All Other States and District of Columbia					
A	900	800	9.920	9.916	9.916
B	900	1,000	10.065	9.957	9.916
All States					
A	67,500	58,400	9.989	9.967	9.959
B	67,500	67,500	10.047	9.977	9.952

Table B9 Projected Effect of Change in Interviewer Experience Based on Year in Which a Design Change Occurs: California

Experience at Time of Interview	2007 Distribution	No Change in Cluster Size	Option A			Option B		
			Cluster Size			Cluster Size		
		9.375	9.375	18.750	28.125	9.375	18.750	28.125
1 to 19	0.0425	0.040	0.113	0.073	0.054	0.104	0.061	0.041
20 to 39	0.0408	0.040	0.113	0.073	0.054	0.104	0.061	0.041
40 to 59	0.0367	0.040	0.113	0.073	0.054	0.104	0.061	0.041
60 to 99	0.0590	0.060	0.045	0.053	0.057	0.047	0.056	0.060
100 to 249	0.1398	0.107	0.081	0.095	0.102	0.084	0.099	0.107
250 to 499	0.1762	0.183	0.138	0.163	0.175	0.144	0.170	0.182
500 to 749	0.1593	0.154	0.116	0.137	0.147	0.121	0.143	0.153
750 to 999	0.1299	0.131	0.098	0.116	0.125	0.103	0.121	0.130
1,000+	0.2158	0.245	0.184	0.217	0.233	0.192	0.227	0.243
Weighted Log Odds	-0.152	-0.152	-0.134	-0.144	-0.149	-0.136	-0.147	-0.151
Log Odds Ratio		0.000	0.018	0.008	0.003	0.016	0.005	0.000
Odds Ratio		1.000	1.018	1.008	1.003	1.016	1.005	1.000
Odds	0.111	0.111	0.113	0.112	0.111	0.113	0.112	0.111
Prevalence (Past Year Marijuana Use)	10.000%	10.000%	10.162%	10.072%	10.030%	10.142%	10.047%	10.004%
Work by New Hires (1 to 59)	0.120	0.120	0.338	0.218	0.161	0.311	0.184	0.124
Attrition by Case-Level Experience								
1 to 59		0.500	0.500	0.500	0.500	0.500	0.500	0.500
60+		0.068	0.068	0.068	0.068	0.068	0.068	0.068
Interviewers Trained								
New to Project		10	36	20	14	32	16	10
Continuing		35	35	35	35	35	35	35
Total		45	71	55	49	67	51	45
Sample Size		3,600	4,800	4,800	4,800	4,600	4,600	4,600
Interviews Completed by								
Continuing FIs		90	90	107	114	90	107	114
New to Project FIs		45	45	53	57	45	53	57
Change in Effort		1.000	1.333	1.125	1.049	1.277	1.078	1.005

Table B10 Projected Effect of Change in Interviewer Experience Based on Year in Which a Design Change Occurs: Florida, Texas, and New York

Experience at Time of Interview	2007 Distribution	No Change in Cluster Size	Option A			Option B		
			Cluster Size			Cluster Size		
		9.375	9.375	18.750	28.125	9.375	18.750	28.125
1 to 19	0.0425	0.040	0.003	0.000	0.000	0.013	0.000	0.000
20 to 39	0.0408	0.040	0.003	0.000	0.000	0.013	0.000	0.000
40 to 59	0.0367	0.040	0.003	0.000	0.000	0.013	0.000	0.000
60 to 99	0.0590	0.060	0.067	0.080	0.086	0.065	0.078	0.083
100 to 249	0.1398	0.107	0.120	0.120	0.119	0.117	0.120	0.120
250 to 499	0.1762	0.183	0.206	0.206	0.204	0.200	0.206	0.205
500 to 749	0.1593	0.154	0.173	0.173	0.172	0.168	0.173	0.172
750 to 999	0.1299	0.131	0.147	0.147	0.146	0.143	0.147	0.146
1,000+	0.2158	0.245	0.275	0.274	0.273	0.267	0.275	0.274
Weighted Log Odds	-0.152	-0.152	-0.161	-0.162	-0.162	-0.158	-0.162	-0.162
Log Odds Ratio		0.000	-0.009	-0.010	-0.010	-0.006	-0.010	-0.010
Odds Ratio		1.000	0.991	0.990	0.990	0.994	0.990	0.990
Odds	0.111	0.111	0.110	0.110	0.110	0.110	0.110	0.110
Prevalence (Past Year Marijuana Use)	10.000%	10.000%	9.920%	9.913%	9.914%	9.942%	9.913%	9.913%
Work by New Hires (1 to 59)	0.120	0.120	0.010	0.000	0.000	0.040	0.000	0.000
Attrition by Case-Level Experience								
1 to 59		0.500	0.500	0.500	0.500	0.500	0.500	0.500
60+		0.068	0.068	0.216	0.274	0.068	0.190	0.249
Interviewers Trained								
New to Project		10	1	0	0	3	0	0
Continuing		35	35	30	28	35	31	29
Total		45	36	30	28	38	31	29
Sample Size		3,600	3,200	3,200	3,200	3,300	3,300	3,300
Interviews Completed by								
Continuing FIs		90	90	107	114	90	107	114
New to Project FIs		45	45	53	57	45	53	57
Change in Effort		1.000	0.889	0.750	0.699	0.917	0.773	0.721

Table B11 Projected Effect of Change in Interviewer Experience Based on Year in Which a Design Change Occurs: Illinois, Michigan, Ohio, and Pennsylvania

		No Change in Cluster Size	Option A			Option B		
	2007		Cluster Size			Cluster Size		
Experience at Time of Interview	Distribution	9.375	9.375	18.750	28.125	9.375	18.750	28.125
1 to 19	0.0425	0.040	0.000	0.000	0.000	0.000	0.000	0.000
20 to 39	0.0408	0.040	0.000	0.000	0.000	0.000	0.000	0.000
40 to 59	0.0367	0.040	0.000	0.000	0.000	0.000	0.000	0.000
60 to 99	0.0590	0.060	0.120	0.120	0.120	0.120	0.120	0.120
100 to 249	0.1398	0.107	0.115	0.115	0.115	0.115	0.115	0.115
250 to 499	0.1762	0.183	0.197	0.197	0.197	0.197	0.197	0.197
500 to 749	0.1593	0.154	0.165	0.165	0.165	0.165	0.165	0.165
750 to 999	0.1299	0.131	0.140	0.140	0.140	0.140	0.140	0.140
1,000+	0.2158	0.245	0.263	0.263	0.263	0.263	0.263	0.263
Weighted Log Odds	-0.152	-0.152	-0.161	-0.161	-0.161	-0.161	-0.161	-0.161
Log Odds Ratio		0.000	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009
Odds Ratio		1.000	0.991	0.991	0.991	0.991	0.991	0.991
Odds	0.111	0.111	0.110	0.110	0.110	0.110	0.110	0.110
Prevalence (Past Year Marijuana Use)	10.000%	10.000%	9.916%	9.916%	9.916%	9.916%	9.916%	9.916%
Work by New Hires (1 to 59)	0.120	0.120	0.000	0.000	0.000	0.000	0.000	0.000
Attrition by Case-Level Experience								
1 to 59		0.500	0.556	0.625	0.650	0.444	0.531	0.563
60+		0.068	0.556	0.625	0.650	0.444	0.531	0.563
Interviewers Trained								
New to Project		10	0	0	0	0	0	0
Continuing		35	18	15	14	22	19	18
Total		45	18	15	14	22	19	18
Sample Size		3,600	1,600	1,600	1,600	2,000	2,000	2,000
Interviews Completed by								
Continuing FIs		90	90	107	114	90	107	114
New to Project FIs		45	45	53	57	45	53	57
Change in Effort		1.000	0.444	0.375	0.350	0.556	0.469	0.437

Table B12 Projected Effect of Change in Interviewer Experience Based on Year in Which a Design Change Occurs: Georgia, New Jersey, North Carolina, and Virginia

Experience at Time of Interview	2007 Distribution	No Change in Cluster Size	Option A			Option B		
			Cluster Size			Cluster Size		
		9.375	**9.375**	**18.750**	**28.125**	**9.375**	**18.750**	**28.125**
1 to 19	0.0425	0.040	0.168	0.138	0.123	0.157	0.125	0.109
20 to 39	0.0408	0.040	0.168	0.138	0.123	0.157	0.125	0.109
40 to 59	0.0367	0.040	0.168	0.138	0.123	0.157	0.125	0.109
60 to 99	0.0590	0.060	0.034	0.040	0.043	0.036	0.043	0.046
100 to 249	0.1398	0.107	0.060	0.071	0.077	0.064	0.076	0.082
250 to 499	0.1762	0.183	0.103	0.122	0.131	0.110	0.130	0.140
500 to 749	0.1593	0.154	0.087	0.103	0.110	0.092	0.110	0.118
750 to 999	0.1299	0.131	0.074	0.087	0.094	0.079	0.093	0.100
1,000+	0.2158	0.245	0.138	0.163	0.175	0.147	0.174	0.187
Weighted Log Odds	-0.152	-0.152	-0.120	-0.128	-0.131	-0.123	-0.131	-0.135
Log Odds Ratio		0.000	0.031	0.024	0.020	0.029	0.021	0.017
Odds Ratio		1.000	1.032	1.024	1.021	1.029	1.021	1.017
Odds	0.111	0.111	0.115	0.114	0.113	0.114	0.113	0.113
Prevalence (Past Year Marijuana Use)	10.000%	10.000%	10.287%	10.218%	10.186%	10.262%	10.189%	10.155%
Work by New Hires (1 to 59)	0.120	0.120	0.505	0.413	0.370	0.472	0.374	0.328
Attrition by Case-Level Experience								
1 to 59		0.500	0.500	0.500	0.500	0.500	0.500	0.500
60+		0.068	0.068	0.068	0.068	0.068	0.068	0.068
Interviewers Trained								
New to Project		2	18	12	10	16	11	9
Continuing		9	9	9	9	9	9	9
Total		11	27	21	19	25	19	17
Sample Size		900	1,600	1,600	1,600	1,500	1,500	1,500
Interviews Completed by								
Continuing FIs		90	90	107	114	90	107	114
New to Project FIs		45	45	53	57	45	53	57
Change in Effort		1.000	1.777	1.500	1.398	1.667	1.406	1.311

Table B13 Projected Effect of Change in Interviewer Experience Based on Year in Which a Design Change Occurs: All Other States and District of Columbia

Experience at Time of Interview	2007 Distribution	No Change in Cluster Size	Option A			Option B		
			Cluster Size			Cluster Size		
		9.375	9.375	18.750	28.125	9.375	18.750	28.125
1 to 19	0.0425	0.040	0.003	0.000	0.000	0.069	0.020	0.000
20 to 39	0.0408	0.040	0.003	0.000	0.000	0.069	0.020	0.000
40 to 59	0.0367	0.040	0.003	0.000	0.000	0.069	0.020	0.000
60 to 99	0.0590	0.060	0.068	0.120	0.120	0.054	0.064	0.120
100 to 249	0.1398	0.107	0.121	0.115	0.115	0.096	0.114	0.115
250 to 499	0.1762	0.183	0.206	0.197	0.197	0.165	0.195	0.197
500 to 749	0.1593	0.154	0.173	0.165	0.165	0.139	0.164	0.165
750 to 999	0.1299	0.131	0.147	0.140	0.140	0.118	0.140	0.140
1,000+	0.2158	0.245	0.275	0.263	0.263	0.220	0.261	0.263
Weighted Log Odds	-0.152	-0.152	-0.161	-0.161	-0.161	-0.145	-0.157	-0.161
Log Odds Ratio		0.000	-0.009	-0.009	-0.009	0.007	-0.005	-0.009
Odds Ratio		1.000	0.991	0.991	0.991	1.007	0.995	0.991
Odds	0.111	0.111	0.110	0.110	0.110	0.112	0.111	0.110
Prevalence (Past Year Marijuana Use)	10.000%	10.000%	9.920%	9.916%	9.916%	10.065%	9.957%	9.916%
Work by New Hires (1 to 59)	0.120	0.120	0.010	0.000	0.000	0.208	0.061	0.000
Attrition by Case-Level Experience								
1 to 59		0.500	0.500	0.250	0.301	0.500	0.500	0.126
60+		0.068	0.068	0.250	0.301	0.068	0.068	0.126
Interviewers Trained								
New to Project		2	0	0	0	5	1	0
Continuing		9	9	7	7	9	9	9
Total		11	9	7	7	13	10	9
Sample Size		900	800	800	800	1,000	1,000	1,000
Interviews Completed by								
Continuing FIs		90	90	107	114	90	107	114
New to Project FIs		45	45	53	57	45	53	57
Change in Effort		1.000	0.889	0.750	0.699	1.111	0.938	0.874

The largest impact on State estimates occurs for Georgia, New Jersey, North Carolina, and Virginia. Increased sample sizes under both options shift the distribution of interviews toward administration by less experienced interviewers and results in increases of the projected prevalence estimates. Movement toward larger cluster sizes partially offsets the increases. Two estimates will round to 10.3 and the other four will round to 10.2 percentage points.

Options A and B shift the sample size in opposite directions for the remaining States and the District of Columbia. Under Option A, the cluster size options all lead to a reduction of less than 0.1 percentage points, but the projected estimates will round to 9.9. Under Option B, a small increase (less than 0.1 percentage points) is projected using the current average cluster size of 9.375, and a small decrease (less than 0.1) is projected using larger cluster sizes. The projected Option B prevalence estimates will round to 10.1, 10.0, and 9.9 depending on the cluster size option.

More details of the modeling process are shown in Tables B9 through B13.

In conclusion, although the impact of interviewer experience has been shown to be decreasing over time, some statistically significant effects can still be identified. These effects can be shown to increase or decrease reported prevalence measures when the experience distribution shifts in a major way, which may be the case with a design change. The potential for large increases is greatest when sample sizes increase. Reductions in sample size are more likely to decrease prevalence estimates, but the magnitude of the decrease is limited to the effect of using only experienced interviewers in the initial year.[27]

The largest impacts on reported prevalence are projected to occur with large increases in sample size, and the impact is reduced in the second and succeeding years. It may be possible to reduce this impact in any given year by phasing in sample increases gradually over several years.

As a caution, the modeling exercise used national averages and made many simplifying assumptions. As an example, the required number of interviewers shown in Tables B9 through B13 made rough assumptions based on 2008 and 2009 experience. Interviewers could complete an average of 81 interviews per year for the current design with the cluster size set at 9.375. With increases in cluster size, the judgment of field staff was that interviewers could complete 96 interviews for a cluster size of 18.750, or 103 interviews for a cluster size of 28.125. This was used to compute an efficiency factor for developing a projection of the number of interviews an interviewer could complete under a particular design. As an example, the efficiency factor for increasing the cluster size from 9.375 to 18.750 was approximated relative to the current design as 96/81 = 1.185.

For Tables B9 through B13, the current design interviewer production was set at 90 per year for continuing interviewers and 45 per year for newly hired interviewers; this is roughly consistent with the mix of continuing and newly hired interviewers projected under the "no change in cluster size" column. Adjustments in productivity then were computed for each design based on the efficiency factor for the design. For example, for Option B with a cluster size of 18.750, the productivity of interviewers was assumed to be 107 (90 × 1.185) for continuing

[27] Some new hiring may still be necessary during the course of the survey year if continuing interviewers leave the project unexpectedly, but this was not incorporated into the spreadsheet model.

interviewers and 53 (45 × 1.185) for new interviewers. The number of interviewers of each type then was projected based on the number of interviews expected to be completed by each type divided by the projected interviewer productivity. As an example, 16 (i.e., 3 × 0.061 × 4,600 / 53) new interviewers are projected for Option B with a cluster size of 18.750.

The "change in effort" row value for Option B with a cluster size of 18.750 then was computed as the ratio of the sample sizes divided by the efficiency factor, or (4,600 / 3,600) / 1.185 = 1.078. "Change in effort" is presented as summary measure of increased work required when accounting for both an increased sample size and a better cost efficiency when using larger cluster sizes. The process has to be viewed as conditional on the inputs and methods of the model. No attempt is made to incorporate uncertainty because of sampling or other errors in the model parameters. Because of unique local conditions, the results may not apply equally to all States or proportionally to substate areas. As discussed in Section 3.11.3 in Chapter 3, substate areas do not usually share boundaries with the State sampling regions used as strata in allocating the sample. This adds to the unpredictability of any proportional projections of sample size to substate areas. Other local factors can also override the impact of shifts in interviewer experience.

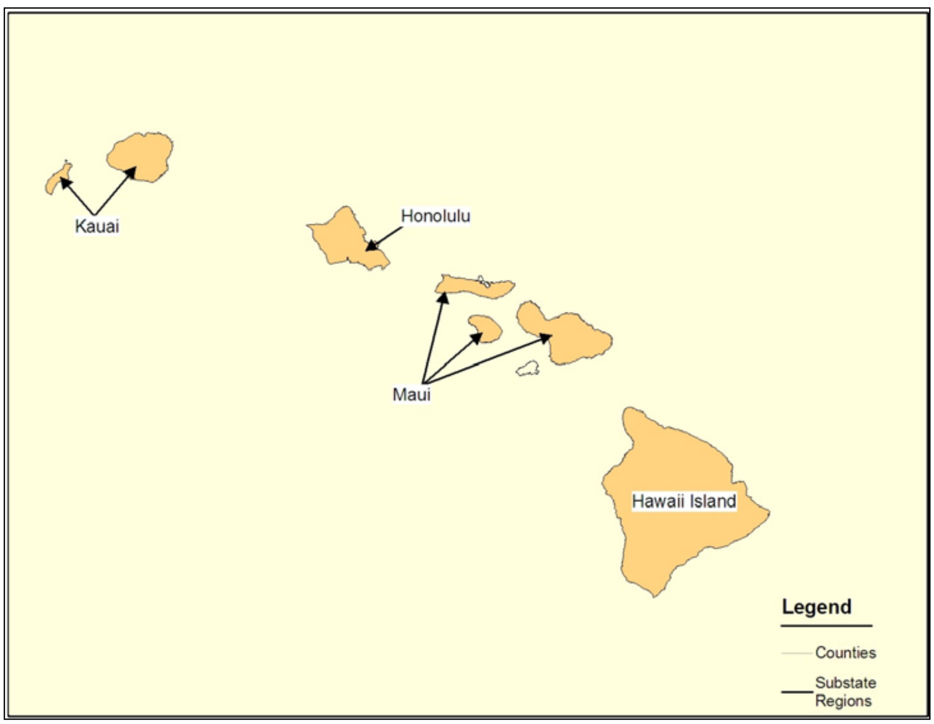
Finally, it should be noted that the projected impacts of the design options considered were between a 0.1 decrease and a 0.3 percentage point increase on an arbitrarily chosen baseline value of 10.0 when applied to State estimates. The seriousness of this range of error needs to be evaluated in terms of the usual precision of reported State estimates.

Appendix C: Investigating Design Options for Hawaii

The State of Hawaii has five counties. Four of these counties consist of one or more islands and are included in National Survey on Drug Use and Health (NSDUH) substate reports[28] (see Figure C1). Kalawao County, which shares Molokai island with one of the islands of Maui County, has a total population of 90 persons (U.S. Census Bureau, 2012). Commonly referred to as Kalaupapa Settlement, Kalawao County's very small population is primarily restricted to former patients with leprosy (i.e., Hansen's disease) who voluntarily choose to remain there as residents. Kalawao County is not used as a substate planning area by Hawaii officials because it is not a political subdivision of the State of Hawaii and does not have a county government.

The current sample design allocates the sample in proportion to a composite size measure based on the three NSDUH target age groups (12 to 17, 18 to 25, and 26 or older). Hawaii's State officials have expressed concern that NSDUH has not been able to produce a sufficiently large sample in the island of Kauai for previous substate reports. For these reports, data were combined for the islands of Maui and Kauai, which are hundreds of miles apart and may have different substance use and mental health characteristics. In order to address Hawaii's concern, options were considered to adjust the allocation to achieve a minimum 3-year sample of 200 persons aged 12 or older in the four more populous counties.

Figure C1 Hawaiian Islands and Counties Used in NSDUH's Substate Reports



[28] The island of Kahoolawe is uninhabited and was a former training ground and bombing range used by the U.S. armed forces. Access to the island is controlled by the Kahoolawe Island Reserve Commission and restricted to appropriate cultural, historical, archaeological, and educational activities (see http://kahoolawe.hawaii.gov/). Thus, this island is not identified on NSDUH's substate region map as being part of a treatment planning area for Hawaii.

Two options were considered: (1) increase the 3-year allocation to Kauai County to 200 persons and adjust the allocation to other counties to maintain a Hawaii 3-year sample of 2,700, and (2) increase the 3-year allocation to Kauai County to 200 and allow the other county allocations to remain at current levels. Table C1 summarizes the sample allocation to counties at current levels and for each of the two options. Note that the current and projected sample allocations are based on expected sample sizes. State sampling regions are constructed in terms of whole census tracts, general proximity, and approximately equal allocations; they do not necessarily correspond with counties. Rather than showing projected sampling errors, Table C1 shows the expected change in sampling errors under the two options. Both options improve the precision (reduce the modeled sampling error) of Kauai County estimates. The option to reallocate and maintain the current State sample size causes slight reductions in precision (increases in sampling error) for the other counties. Shifting more of the sample to Kauai County would also result in a small increase in cost. The option to simply supplement the sample in Kauai County provides the same improvement in precision for Kauai County and no change in the precision of other county estimates. The increase in expected sample size by 68 cases at the State level would result in a moderate increase in cost for the Hawaii NSDUH component.

Table C1 Current and Optional 3-Year Sample Allocations for Hawaii

	Hawaii County	Honolulu County	Kauai County	Maui County	State of Hawaii	Additional Ballpark Costs
County FIPS Code	**001**	**003**	**007**	**009**	**Total**	
3-Year Allocations (Expected Sample Sizes)						
Current	354	1,915	132	299	2,700	
Reallocate	344	1,864	200	291	2,700	$15,000
Supplement	354	1,915	200	299	2,768	$30,000
Percent Relative Change in Standard Errors						
Reallocate	1.35%	1.35%	-18.81%	1.35%	0.46%	
Supplement	0.00%	0.00%	-18.81%	0.00%	-0.84%	

FIPS = Federal information processing standards (codes).