

# 2002 NSAF Variance Estimation

Report No. 4

Prepared by:

J. Michael Brick

Teresa Strickler

David Ferraro

Westat



Assessing  
the New  
Federalism

An Urban Institute  
Program to Assess  
Changing Social Policies

Methodology Reports

## PREFACE

*2002 NSAF Variance Estimation* is the fourth report in a series describing the methodology of the 2002 National Survey of America's Families (NSAF). The NSAF is part of the *Assessing the New Federalism* project at the Urban Institute, conducted in partnership with Child Trends. Data collection for the NSAF was conducted by Westat.

The NSAF is a major household survey focusing on the economic, health, and social characteristics of children, adults under the age of 65, and their families. During the third round of the survey in 2002, interviews were conducted with over 40,000 families, yielding information on over 100,000 people. The NSAF sample is representative of the nation as a whole and of 13 states, and therefore has an unprecedented ability to measure differences between states.

### About the Methodology Series

This series of reports has been developed to provide readers with a detailed description of the methods employed to conduct the 2002 NSAF. The 2002 series of reports include:

- No. 1: An overview of the NSAF sample design, data collection techniques, and estimation methods
- No. 2: A detailed description of the NSAF sample design for both telephone and in-person interviews
- No. 3: Methods employed to produce estimation weights and the procedures used to make state and national estimates for *Snapshots of America's Families*
- No. 4: Methods used to compute and results of computing sampling errors
- No. 5: Processes used to complete the in-person component of the NSAF
- No. 6: Collection of NSAF papers
- No. 7: Studies conducted to understand the reasons for nonresponse and the impact of missing data
- No. 8: Response rates obtained (taking the estimation weights into account) and methods used to compute these rates
- No. 9: Methods employed to complete the telephone component of the NSAF
- No. 10: Data editing procedures and imputation techniques for missing variables
- No. 11: User's guide for public use microdata
- No. 12: 2002 NSAF questionnaire

## **About This Report**

Report No. 4 describes the methods and estimated sampling errors for statistics from the 2002 survey. The report covers an overview of the sample design, summarizes the precision of the survey estimates, and discusses computing sampling error estimates for differences estimated from NSAF data. Report No. 4 also provides a general review of the two main methods of computing sampling errors or variances of estimates from surveys with complex survey designs like the NSAF, including a detailed discussion of the replication method of variance estimation and an overview of available software available for computing sampling errors.

## **For More Information**

For more information about the National Survey of America's Families, contact:

*Assessing the New Federalism*  
Urban Institute  
2100 M Street, NW  
Washington, DC 20037

E-mail: [nsaf@ui.urban.org](mailto:nsaf@ui.urban.org)  
Web site: <http://anf.urban.org/nsaf>

*Adam Safir and Tim Triplett*

## CONTENTS

<u>Chapter</u>		<u>Page</u>
1	INTRODUCTION .....	1-1
	1.1 Overview of the Survey .....	1-1
	1.2 Design Effects .....	1-3
	1.3 Estimates of Change .....	1-14
2	METHODS FOR VARIANCE ESTIMATION .....	2-1
	2.1 Replication Methods .....	2-1
	2.2 Taylor Series Method.....	2-2
	2.3 Differences between National and Study Area Variance Estimation .....	2-2
	2.4 Estimating Variances of Change Over Time .....	2-3
3	REPLICATION FOR THE NSAF .....	3-1
	3.1 Choice of Method .....	3-1
	3.2 Design of Replicates .....	3-2
4	SOFTWARE FOR COMPUTING VARIANCES .....	4-1
	4.1 WesVar .....	4-1
	4.2 SUDAAN.....	4-2
	4.3 Review of Software Output .....	4-4
	REFERENCES .....	R-1
Appendixes		
<u>Appendix</u>		<u>Page</u>
A	Round 1 Design Effect Tables .....	A-1
B	Round 2 Design Effect Tables .....	B-1
C	Design Effect Statistics .....	C-1

## CONTENTS (CONTINUED)

### Tables

<u>Table</u>		<u>Page</u>
1-1	Average DEFF and DEFT for Estimates from the 2002 NSAF National Sample, by Estimate Type .....	1-5
1-2	Average DEFF and DEFT for Estimates from the 2002 NSAF Study Area Child Sample, by Study Area.....	1-8
1-3	Average DEFF and DEFT for Estimates from the 2002 NSAF Study Area Child Sample, Hispanic Children, by Study Area.....	1-8
1-4	Average DEFF and DEFT for Estimates from the 2002 NSAF Study Area Child Sample, Black Children, by Study Area .....	1-9
1-5	Average DEFF and DEFT for Estimates from the 2002 NSAF Study Area Adult Pair File Sample, by Study Area.....	1-10
1-6	Average DEFF and DEFT for Estimates from the 2002 NSAF Study Area Adult Pair Sample, Hispanic Adults, by Study Area .....	1-11
1-7	Average DEFF and DEFT for Estimates from the 2002 NSAF Study Area Adult Pair Sample, Black Adults, by Study Area .....	1-11
1-8	Average DEFF and DEFT for Estimates from the 2002 NSAF Study Area Adult Pair Sample, “Option B” Adults, by Study Area .....	1-12
1-9	Average DEFF and DEFT for Estimates from the 2002 NSAF Study Area Adult Pair Sample, Adults in Households with Children, by Study Area .....	1-13
1-10	Average DEFF and DEFT for Estimates from the 2002 NSAF Study Area Adult Pair Sample, Adults in Households with No Children, by Study Area .....	1-13
4-1	Estimates of Children Computed Using WesVar and SUDAAN.....	4-7

## CONTENTS (CONTINUED)

### Figures

<u>Figure</u>		<u>Page</u>
1-1	Study Areas.....	1-1

### Exhibits

<u>Exhibit</u>		<u>Page</u>
4-1	Computing Sample Estimates and Their Corresponding Standard Errors Using WesVar.....	4-3
4-2	Computing Sample Estimates and Their Corresponding Standard Errors Using SUDAAN.....	4-5

# 1. INTRODUCTION

This report describes the methods and estimated sampling errors for statistics from the 2002 National Survey of America's Families (NSAF). The first chapter gives an overview of the sample design and summarizes the precision of the survey estimates for both children and adults. This chapter also discusses computing sampling error estimates for differences estimated from NSAF data collected in 2002 (Round 3) and the earlier two rounds (Round 1 in 1997 and Round 2 in 1999).

The rest of the report describes the methodology used to create these types of estimates of sampling variability. Chapter 2 reviews the two main methods of computing sampling errors or variances of estimates from surveys with complex survey designs like the NSAF. Chapter 3 discusses the replication method of variance estimation for the NSAF in more detail. Chapter 4 discusses software available for computing sampling errors.

## 1.1 Overview of the Survey

The NSAF collected information on the economic, health, and social dimensions of the well-being of children, adults under the age of 65, and their families in 13 states and the balance of the nation. The Urban Institute selected these study areas (see figure 1-1) before the first survey in 1996 because they represent a broad range of fiscal capacity, child well-being, and approaches to government programs. Data were also collected in the balance of the nation to permit estimates for the United States as a whole. The sample design is briefly outlined here. The complete details are provided in the *2002 NSAF Sample Design*, Report No. 2.

**Figure 1-1.  
Study Areas**

Alabama	Massachusetts	New Jersey	Wisconsin Balance of nation
California	Michigan	New York	
Colorado	Minnesota	Texas	
Florida	Mississippi	Washington	

The primary goal of the survey in all three rounds was to obtain social and economic information about children in low-income families (with incomes below 200 percent of the poverty threshold) since the impact of new federalism on these children was likely to be greatest. Secondary goals included obtaining similar data on children in higher income families, plus adults under age 65 (with and without children).

One major change in the design for the 2002 survey is that Milwaukee County in Wisconsin, which had been a separate study area with its own sample in 1997 and 1999, was no longer a separate study area. The Milwaukee County study area was included with the rest of Wisconsin as a single study area.

Another major change in the sample design was to eliminate the nontelephone sample in the study areas representing specific states. The NSAF uses two frames—a random digit dialing (RDD) sample of households with telephones and an area sample conducted in person for those households without telephones. In 1997 and 1999, this dual-frame approach (see Waksberg et al. 1997) was used for each study area and for the nation. In 2002, only a national area sample was selected, and the same dual-frame method was used to produce national estimates. No separate sample of households without telephones was selected in the study areas, so estimates for the study areas are based only on the RDD sample selected. The weights of the RDD sample in the study areas were adjusted to reduce the potential bias associated with not sampling households without telephones in the study areas.

The national area sample consisted of a subsample of primary sampling units (PSUs) used in the 1999 survey. The 1999 PSUs were subsampled in the study areas, while all the PSUs from the balance of the United States were retained. Since block groups with very high telephone coverage rates as of the 1990 Census were excluded in the Round 1 and Round 2 sampling, this same restriction was carried forward to the 2002 survey of nontelephone households.

Another new feature introduced in the 2002 survey was the subsampling of refusal cases. A random sample of telephone numbers was selected, and only those identified in the subsample were followed to completion if they refused the screening interview. The numbers that were not subsampled were classified separately to make it easier to compute response rates.

In the RDD sample, a screener-based subsampling of households was used to sample low-income households at a higher rate than other households. A very short income question was asked during the RDD screening interview, and those that reported an absence of children or incomes above 200 percent of the poverty threshold were subsampled.

Within both the RDD and area samples, household members were subsampled to reduce the respondent burden. If there were multiple children under age 6, one was randomly selected. The same was done for children 6 to 17 years old. Data were collected from the most knowledgeable adult (MKA) in the household for the sampled child. During the MKA interview, data were also collected about the MKA and about his/her spouse or partner. Most questions asked about the MKA were repeated in reference to the spouse or partner; however, some questions on health insurance and health care usage were asked in reference to only one of the two. The target of these questions was randomly assigned to either the MKA or his/her spouse or partner. Questions asked only about the MKA were those related to feelings, religious activities, and opinions.

Other adults in households with children were subsampled, as were adults in adult-only households. Adults were eligible only if they would not have been the MKA for other children in the household if those children had been selected. Self-response was required for sampled adults—with proxy data collected about his/her spouse or partner (if living in the same household). Data were not collected directly from the spouse of a sampled adult. As in the MKA interview, there were also some questions related to feelings, religious activities, and opinions that were asked only about the sampled adult.

## 1.2 Design Effects

The precision of the sample estimates can be derived from a survey by computing sampling errors from the sample itself. Estimates of sampling errors are used to make inferences about the differences between two population parameters or differences between the population parameter in 2002 and the same parameter in 1997. For instance, the difference between the proportions of persons in poverty in two regions of the country can be estimated by computing the difference of the estimated proportions and the estimated sampling error of this difference. Estimated sampling errors can appear in several forms, including the estimated standard error of a survey estimate, the ratio of the estimated standard error to the survey estimate (called the coefficient of variation, or CV of the survey estimate), and a confidence interval about the survey estimate.

Another way of describing the variability of an estimate from a survey is by using the *design effect*. The term design effect is used to describe the variance of a sample estimate under a particular sample design relative to the variance of the estimate under a simple random sample with the same sample size. Design effects are used to evaluate the efficiency of the sampling design and estimation procedure utilized to develop the estimates.

The concept of design effect was popularized by Kish (1965) to deal with complex sample designs involving stratification and clustering. Stratification can increase efficiency over simple random sampling, but clustering usually reduces efficiency of the estimate owing to positive intracluster correlation among the subunits in the clusters. To determine the effect of a complex sample design on the sampling variance of an estimate, this variance is often compared with the variance expected under simple random sampling. This is accomplished by calculating a ratio of variances associated with an estimate, namely

$$\text{DEFF} = \frac{\text{sampling variance of a complex sample}}{\text{sampling variance of a simple random sample}}.$$

This ratio is called the design effect (DEFF) of the estimate for the sample design. At the analysis stage, the DEFF is sometimes used because standard statistical analysis software (such as the usual procedures in SAS and SPSS) assume the data are from a simple random sample when computing sampling errors of estimates. The DEFF can, in some circumstances, indicate how appropriate this is and can be used to adjust these simple estimates to produce ones closer to the actual sampling errors of the estimates (Skinner, Holt, and Smith 1989).

The design effect for a proportion is estimated as

$$DEFF_{PROP} = \frac{v(p)_{COMPLEX}}{v(p)_{SRS}},$$

where  $p$  denotes the estimate of the proportion,  $v(p)_{SRS}$  is the estimated simple random variance  $v(p)_{SRS} = \frac{p(1-p)}{n}$ , and  $v(p)_{COMPLEX}$  is the variance of the complex sample calculated appropriately.

In most cases, design effects for complex samples are larger than 1. In the NSAF, design effects are greater than 1 because of differential sampling fractions and the intracluster correlation of units within clusters (Kish 1992). As will be seen shortly, some design effects from the survey are considerably greater than 1.

In the NSAF and most large-scale surveys, a large number of data items or variables are collected from respondents. Each variable has its own design effect. One way to represent all of these is to compute design effects for a number of similar variables and then average them. This average is used to represent the design effects for the group of variables. This practice is described in Wolter (1985).

Table 1-1 gives the average design effects for estimates of proportions using the national weights. The table has the average design effects for estimates from the child file and the adult pair file. In the reports from previous rounds, the same tables contained both the average design effects for both the national estimates and the study area estimates. In Round 3, the sample for the national estimates include the nontelephone cases, while the samples for the study areas do not. Thus, the estimates, variances, and design effects for the national and area study are different. The average design effects for the study areas using the appropriate weights are given in subsequent tables.

Table 1-1 presents the national estimates for subgroups from the child and adult pair weight files (for the child file the subgroups are black children and Hispanic children; for the adult pair file the subgroups are black adults, Hispanic adults, Option “B” adults, adults in households with children, and adults in households without children). “Option B” adults are adults who are not MKAs or the spouse of an MKA. The table further subdivides each subgroup by giving the estimates for all persons and for low-income persons (less than or equal to 200 percent of the poverty level).

The estimates of design effects in the table are averages. For the child file entries, the averages were computed from 28 statistics. Similarly, the estimates in the tables from the adult pair file were based on 20 statistics for each group. The statistics used for both sets of estimates are shown in appendix C. If an estimate had fewer than 10 observations (this is the unweighted number of sampled persons in the denominator of the estimated percentage), then the statistic was dropped from the calculations for that subgroup.

**Table 1-1.**  
**Average DEFF and DEFT for Estimates from the 2002 NSAF National Sample, by Estimate Type**

Type of estimate	All						Low-Income					
	Average	Maximum	Minimum	DEFT	Nominal sample size	Effective sample size	Average	Maximum	Minimum	DEFT	Nominal sample size	Effective sample size
<b>Child</b>												
All races	3.17	5.89	1.54	1.77	34,332	10,828	3.23	5.86	1.58	1.78	13,030	4,031
Black	3.03	4.38	1.77	1.73	4,382	1,447	3.08	4.27	1.76	1.74	2,611	848
Hispanic	2.86	5.07	1.80	1.67	6,146	2,150	2.82	4.42	1.51	1.66	3,846	1,362
<b>Adult</b>												
All race	4.62	7.74	2.71	2.13	69,421	15,015	4.64	7.70	2.84	2.14	21,888	4,722
Black	3.55	5.22	1.52	1.86	6,839	1,925	3.46	5.05	1.74	1.84	3,431	993
Hispanic	3.79	6.22	2.05	1.93	9,267	2,447	3.88	6.72	2.15	1.94	5,441	1,403
Option B adults	3.89	7.11	1.90	1.94	49,111	12,633	3.84	7.41	2.19	1.93	15,590	4,059
Households with children	4.12	9.42	1.32	1.99	52,294	12,694	4.00	8.68	1.33	1.96	16,612	4,158
Adult-only households	2.96	4.92	1.09	1.70	17,127	5,781	3.44	5.89	1.58	1.83	5,276	1,534

Each subgroup has four entries: the average DEFF, the maximum DEFF, the minimum DEFF, and the DEFT. The DEFT is the square root of the design effect, so it is like the DEFF but on the scale of the standard deviation of the estimate rather than the variance. The figures labeled DEFT in the tables are the averages of the DEFTs.

The DEFT is a convenient measure because it can be used directly in computing confidence intervals for the estimates, whereas the square root of the DEFF must be computed before it can be used (Verma, Scott, and O’Muircheartaigh 1980). One reason for presenting the DEFTs is to dampen some of noise associated with the DEFFs. The maximum and minimum values of the DEFFs in the tables show that there is considerable variability in these quantities. By taking the square root of the DEFF and averaging these values, the variability is somewhat reduced.

Table 1-1 also shows the effective sample sizes for each subgroup. Effective sample sizes were calculated by dividing the nominal sample sizes by the average design effects. The effective sample size is the size of a simple random sample that has the same precision as that of the complex design.

Before discussing this and other tables of design effects, the most important factors that result in design effects larger than one in the NSAF should be reviewed. These factors are the following:

1. Oversampling by study area. The need for both study area and national estimates required oversampling to produce stable separate estimates for the study areas. This oversampling increased the design effect for national estimates.
2. Household screening. Additional variability comes from the subsampling of households without children and those above 200 percent of the poverty level. The misclassification of households as “above 200 percent poverty” when they actually were not also increases the variance of estimates restricted to persons at or below 200 percent of the poverty level. See Flores-Cervantes et al. (1998) for a discussion of this topic.
3. Within-household subsampling. Differential sampling rates at the person level also contribute to increases in design effects. Children and adults were subsampled within households for both the RDD and area sample components.<sup>1</sup>
4. Differential sampling rates associated with subsampling and attempting to complete only a portion of the households that initially refused the screener interview. The weights for the subsampled cases were increased slightly to account for the initial refusal cases that were not subsampled.
5. Clustering of households in the area sample. This only affects the national weights because the area sample cases are not included in the study area weights. Households were clustered within segments and segments were clustered within PSUs. Design effects

---

<sup>1</sup> Households without children and households above 200 percent of the poverty level were not subsampled in the area component of the study, so design effects associated with such subsampling were not incurred in the area sample but were for the RDD sample.

increase to the extent that respondents in the same cluster are similar in their responses to survey items. Larger design effects due to area sample clustering are more likely to affect low-income households because a larger percentage of low-income persons are in nontelephone households.

It is important to remember that the average design effects are themselves estimates, and are based on a relatively small set of estimated variances from the survey. The estimates included in the tables also affect the average design effects. The estimates used were specifically selected, and many of them are related to lower income status. Other choices of estimates would give different averages. Thus, while the averages in the tables provide some guidance, computing the actual standard error rather than using these averages is generally more reliable. The methods described in the rest of this report show how the appropriate standard errors can be computed from the survey data.

The average DEFTs for all children and for low-income children in table 1-1 are approximately 1.8. The average DEFT does not vary much by the race of the child, ranging from 1.7 to 1.8. Thus, the sampling error for the national estimates is about 70 to 80 percent above what would be expected in a simple random sample of the same size. All the reasons noted here contribute to inflating the national DEFT above unity, but a main contributor is the differential sampling of households by study area.

Table 1-1 also shows the average DEFTs for national estimates of adults fall between about 1.7 and 2.1. The average DEFTs for low-income adults are approximately equal to the DEFTs for all adults. There is some variation in the DEFTs by presence of children in the household, with estimates for adults in households with children having larger DEFTs (2.0) than estimates for adults in households without children (1.7).

The average DEFTs for national estimates from Round 3 are similar to the corresponding DEFTs for estimates from Round 2 (Round 2 DEFTs are lower than those from Round 1 as discussed in *1999 Variance Estimation*, Report No. 4). To aid in this type of comparison, the average design effects tables for national estimates from Round 1 and Round 2 are given in appendices A and B, respectively, in the same format as table 1-1. These DEFTs differ slightly from those presented in earlier reports because the weights for these rounds were recomputed using the 2000 Census (see *2002 NSAF Sample Estimation Survey Weights*, Report No. 3) and because the estimates used in the computations differ somewhat from those used in previous rounds. The estimates used in the computations were changed so the design effects for all three rounds were computed using the same set of statistics. These revisions in the computations for the design effects from previous rounds were made to make them more comparable to the Round 3 estimates.

The average DEFTs for the national child estimates from Round 3 are about 10 to 20 percent smaller than the corresponding DEFTS from Round 2. The average DEFTs for the national adult estimates from Round 3 are nearly all within 5 to 7 percent of the corresponding Round 2 estimates. Some major factors that could cause differences in the design effects between Round 2 and Round 3 are the elimination of Milwaukee as a separate study area (should decrease the

Round 3 design effects), the reduction of the area sample size, and the introduction of refusal subsampling in Round 3 (both should increase the Round 3 design effects).

Tables 1-2, 1-3, and 1-4 present the average design effects for study area estimates from the child file. Table 1-2 gives the average DEFTs for children by study area. The average DEFTs are relatively consistent for the child estimates across study areas, ranging from 1.1 to 1.3. There is also little variability in the estimates for low-income children compared with all children. The lower average DEFT for the study areas than for the national estimates is largely because the national estimate DEFTs are inflated by sampling the smaller study areas at higher rates in order to make precise estimates at the study area level. Within study areas, the differential sampling was only because of other factors such as income.

Table 1-3 gives the average DEFTs for Hispanic children. The averages for some of the study areas are based on a very small number of statistics because any estimate with a sample size of less than 10 children was excluded from the computations. For example, in Alabama only 15 estimates of Hispanic children (9 estimates for low-income Hispanic children) have sample sizes (the numerator of the estimated percentage) of 10 or more, so the estimated average DEFT for Hispanic children in Alabama is unstable. The small sample size is a consequence of there being so few Hispanics in Alabama. The other study area where this occurred is Mississippi where only 6 estimates of Hispanic children were used to compute the average DEFTs. The average DEFTs for Hispanics are about the same as the average DEFT for all children shown in table 1-1, but it should be remembered that these averages may be unstable.

**Table 1-2.**  
**Average DEFF and DEFT for Estimates from the 2002 NSAF Study Area Child Sample, by Study Area**

Study area	All Children				Low-Income Children			
	Average	Maximum	Minimum	DEFT	Average	Maximum	Minimum	DEFT
Alabama	1.43	2.49	0.69	1.18	1.54	2.49	0.66	1.23
California	1.71	2.37	0.86	1.30	1.88	2.53	0.91	1.36
Colorado	1.59	2.57	1.04	1.25	1.71	2.66	0.99	1.30
Florida	1.39	2.15	0.80	1.17	1.47	2.06	0.83	1.21
Massachusetts	1.30	2.33	0.81	1.13	1.61	2.27	1.08	1.27
Michigan	1.60	2.65	1.02	1.25	1.80	2.93	1.04	1.33
Minnesota	1.45	2.29	0.89	1.19	1.82	3.37	0.95	1.33
Mississippi	1.55	2.15	0.67	1.23	1.69	2.40	0.70	1.29
New Jersey	1.55	2.21	0.85	1.24	1.70	2.56	0.74	1.30
New York	1.49	2.42	0.87	1.21	1.70	2.87	0.99	1.29
Texas	1.63	2.99	0.93	1.27	1.69	2.90	0.99	1.29
Washington	1.81	3.21	1.01	1.33	1.74	2.91	1.14	1.31
Wisconsin	1.66	2.82	0.81	1.27	1.60	2.76	0.57	1.26
Balance of nation	1.48	2.04	0.91	1.21	1.60	2.44	0.70	1.25

**Table 1-3.**  
**Average DEFF and DEFT for Estimates from the 2002 NSAF Study Area Child Sample,**  
**Hispanic Children, by Study Area**

Study area	All Children				Low-Income Children			
	Average	Maximum	Minimum	DEFT	Average	Maximum	Minimum	DEFT
Alabama	1.51	1.89	0.78	1.27	1.63	1.99	1.31	1.27
California	1.76	3.52	0.82	1.31	1.80	3.65	0.97	1.33
Colorado	1.61	2.70	0.68	1.25	1.70	2.58	0.71	1.29
Florida	1.48	2.57	0.84	1.21	1.48	2.71	0.67	1.20
Massachusetts	1.55	2.15	0.75	1.24	1.52	2.04	0.82	1.23
Michigan	1.46	2.33	0.76	1.20	1.61	2.29	0.93	1.26
Minnesota	1.41	1.80	1.09	1.19	1.40	1.60	0.73	1.18
Mississippi	1.52	1.83	1.34	1.23	1.89	2.44	1.63	1.37
New Jersey	1.58	2.73	0.79	1.24	1.65	3.32	0.62	1.27
New York	1.62	2.55	0.75	1.26	1.69	2.85	0.86	1.28
Texas	1.69	3.27	1.05	1.29	1.71	3.01	1.07	1.30
Washington	1.73	3.09	1.09	1.31	1.83	3.18	1.05	1.34
Wisconsin	1.50	2.29	0.78	1.21	1.57	2.79	0.78	1.24
Balance of nation	1.64	2.39	0.87	1.27	1.69	2.44	0.86	1.29

Table 1-4 is the corresponding table of average DEFTs for black children. Only two study areas have averages based on fewer than 15 statistics: Washington and Colorado. The average DEFTs for all black children and low-income black children are about the same as the average DEFTs for all children given in table 1-1. Given the instability of all these average DEFTs, there does not seem to be a great deal of variability across all, Hispanic, and black children.

Tables 1-5 through 1-10 present the similar tables of average design effects for study area estimates from the adult pair file. Table 1-5 gives the average DEFTs for adults by study area. The differences between the average DEFTs for the study areas are not very large but do range from about 1.4 to 1.7. The DEFTS for low-income and all adults for a study area are approximately equal. The average DEFTs for the estimates from the adult pair file are somewhat larger than those from the child file.

Tables 1-6 and 1-7 give the estimates for adult Hispanics and blacks, respectively. The average DEFTs for Hispanics is very unstable for many study areas because the average is computed from only a few statistics. For example, in Alabama only 9 adult statistics had sample sizes of 10 or more needed for inclusion in the average DEFT, and in Mississippi only 6 adult statistics had these sample sizes. Other study areas where the average DEFT was based on fewer than 15 statistics for either all adult Hispanics or low-income adult Hispanics are Michigan and Minnesota. For blacks, only Washington and Colorado have average DEFTs based on fewer than 15 statistics for all adults or low-income adults. The average DEFTs by study area are relatively consistent for Hispanic and blacks, with some variation by study area.

**Table 1-4.**  
**Average DEFF and DEFT for Estimates from the 2002 NSAF Study Area Child Sample,**  
**Black Children, by Study Area**

Study area	All Children				Low-Income Children			
	Average	Maximum	Minimum	DEFT	Average	Maximum	Minimum	DEFT
Alabama	1.66	2.28	0.70	1.28	1.57	2.14	0.84	1.24
California	1.45	1.87	1.06	1.20	1.53	1.92	1.03	1.24
Colorado	1.89	2.50	1.28	1.37	2.01	2.71	1.22	1.41
Florida	1.47	2.02	0.85	1.20	1.45	1.89	0.97	1.20
Massachusetts	1.47	2.07	0.93	1.21	1.56	2.45	1.05	1.24
Michigan	1.57	2.20	0.73	1.25	1.55	2.31	0.80	1.24
Minnesota	1.73	2.96	0.94	1.30	1.81	3.24	0.89	1.32
Mississippi	1.64	2.47	0.75	1.27	1.65	2.60	0.69	1.27
New Jersey	1.65	2.48	0.78	1.27	1.87	2.92	0.79	1.35
New York	1.47	2.07	0.86	1.20	1.48	2.06	0.96	1.21
Texas	1.57	2.37	0.87	1.24	1.60	2.74	0.90	1.25
Washington	1.59	2.12	1.09	1.25	1.62	2.11	1.19	1.27
Wisconsin	1.76	2.67	0.47	1.31	1.91	2.97	0.72	1.36
Balance of nation	1.62	2.29	0.92	1.26	1.59	2.41	0.66	1.25

**Table 1-5.**  
**Average DEFF and DEFT for Estimates from the 2002 NSAF Study Area Adult Pair File**  
**Sample, by Study Area**

Study area	All Adults				Low-Income Adults			
	Average	Maximum	Minimum	DEFT	Average	Maximum	Minimum	DEFT
Alabama	2.66	4.38	1.28	1.61	2.49	4.11	1.52	1.56
California	2.21	3.93	1.29	1.47	2.58	4.67	1.45	1.59
Colorado	2.95	4.51	1.54	1.70	2.73	4.39	1.31	1.63
Florida	2.51	4.09	1.34	1.57	2.73	3.92	1.87	1.64
Massachusetts	2.37	3.80	1.14	1.53	2.48	3.68	0.84	1.55
Michigan	2.04	3.57	1.46	1.42	2.15	3.30	1.28	1.45
Minnesota	1.93	3.38	0.93	1.37	1.84	2.85	1.02	1.35
Mississippi	2.02	2.73	1.37	1.41	2.20	3.09	1.47	1.48
New Jersey	2.92	4.83	1.42	1.70	2.76	3.92	1.65	1.65
New York	2.34	4.42	1.45	1.51	2.57	3.84	1.32	1.59
Texas	2.61	4.47	1.93	1.60	2.66	3.90	1.54	1.62
Washington	2.53	3.94	1.17	1.57	2.29	3.23	1.29	1.50
Wisconsin	2.15	3.62	1.20	1.45	2.09	3.43	1.27	1.43
Balance of nation	2.04	2.88	1.06	1.42	2.25	3.20	0.69	1.48

**Table 1-6.**  
**Average DEFF and DEFT for Estimates from the 2002 NSAF Study Area Adult Pair**  
**Sample, Hispanic Adults, by Study Area**

<b>Study area</b>	<b>All Adults</b>				<b>Low-Income Adults</b>			
	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>
Alabama	2.52	4.50	0.99	1.53	3.21	5.93	1.33	1.74
California	2.68	7.08	1.29	1.60	2.59	6.76	1.23	1.56
Colorado	2.88	5.15	1.51	1.67	2.48	6.09	1.08	1.53
Florida	2.26	3.72	1.14	1.49	2.20	3.17	1.15	1.47
Massachusetts	2.09	4.41	0.96	1.42	1.90	3.50	0.42	1.35
Michigan	2.61	5.31	1.09	1.57	3.04	5.77	1.15	1.69
Minnesota	1.50	3.23	0.67	1.20	1.53	3.10	0.70	1.21
Mississippi	1.96	3.01	0.74	1.36	2.13	3.79	0.69	1.40
New Jersey	3.34	7.20	1.58	1.78	2.97	5.39	1.49	1.69
New York	2.58	4.08	0.59	1.58	2.47	4.43	0.78	1.54
Texas	2.59	4.85	1.29	1.59	2.30	4.43	1.39	1.50
Washington	2.11	3.73	0.99	1.43	2.08	3.71	1.31	1.42
Wisconsin	1.70	2.90	1.02	1.29	1.71	3.07	0.85	1.28
Balance of nation	2.12	3.50	1.04	1.44	2.30	3.55	1.10	1.50

**Table 1-7.**  
**Average DEFF and DEFT for Estimates from the 2002 NSAF Study Area Adult Pair**  
**Sample, Black Adults, by Study Area**

<b>Study area</b>	<b>All Adults</b>				<b>Low-Income Adults</b>			
	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>
Alabama	2.18	4.24	0.87	1.44	2.20	5.12	0.79	1.44
California	1.71	2.86	0.71	1.28	1.56	3.09	0.97	1.24
Colorado	2.99	14.47	0.93	1.61	2.37	3.13	0.96	1.52
Florida	2.81	4.72	1.26	1.66	2.63	4.90	1.36	1.60
Massachusetts	2.34	3.64	0.92	1.50	1.87	4.11	1.00	1.34
Michigan	1.85	3.23	0.63	1.34	1.83	3.91	1.10	1.33
Minnesota	1.42	2.76	0.74	1.18	1.48	2.11	0.85	1.21
Mississippi	2.16	3.65	1.39	1.46	2.20	3.48	1.57	1.47
New Jersey	2.24	3.67	0.69	1.47	1.78	2.74	1.04	1.33
New York	1.93	3.21	0.77	1.37	1.96	3.89	1.12	1.39
Texas	2.56	5.88	1.00	1.56	2.14	3.23	1.18	1.45
Washington	1.99	3.54	0.91	1.38	1.93	3.50	0.77	1.36
Wisconsin	1.54	2.64	0.56	1.22	1.44	2.83	0.58	1.17
Balance of nation	1.84	3.08	0.75	1.34	1.84	2.76	1.31	1.35

Table 1-8 gives the same statistics for “Option B” adults. The average DEFTs are smaller for this subset of adults than for all adults, and they are more consistent across study areas. The two main reasons for the smaller DEFTs are (1) the rate of sampling for “Option B” adults within study area was more consistent, whereas all adults include MKAs and “Option B” adults who were selected at different rates; and (2) “Option B” adults have smaller cluster sizes than all adults, reducing the variance of the estimate. Most averages are based on 19 statistics.

The last two tables of adults are for those who live in households with and without children (tables 1-9 and 1-10). The average DEFT is slightly smaller for adults in households without children. For example, in New York the average DEFT is 1.2 for all adults in households without children and 1.4 for adults in households with children.

The design effects for the study area estimates can also be compared across rounds. In the first two rounds, a sample of nontelephone households was selected in each study area and used in the estimates for the study areas. To support more precise estimates of change with Round 3, when the Round 1 and Round 2 weights were recomputed using the 2000 Census as the base, the study area weights were computed two ways. Study area weights were computed both with the nontelephone sample from the study area (similar to what was done in the original Round 1 and Round 2 weighting) and without the nontelephone sample from the study area (comparable to the

**Table 1-8.**  
**Average DEFF and DEFT for Estimates from the 2002 NSAF Study Area Adult Pair**  
**Sample, “Option B” Adults, by Study Area**

Study area	All Adults			Low-Income Adults				
	Average	Maximum	Minimum DEFT	Average	Maximum	Minimum DEFT		
Alabama	1.60	2.39	1.03	1.26	1.53	2.30	1.07	1.23
California	1.44	2.19	1.06	1.19	1.62	2.10	0.96	1.27
Colorado	1.88	2.87	0.85	1.35	1.76	2.73	0.99	1.32
Florida	1.87	3.08	1.04	1.35	1.83	2.63	1.18	1.34
Massachusetts	1.46	2.44	0.89	1.20	1.72	2.19	1.19	1.31
Michigan	1.78	2.42	1.03	1.32	1.74	2.60	0.98	1.31
Minnesota	1.80	2.92	1.00	1.33	1.87	3.00	0.91	1.35
Mississippi	1.56	2.64	1.00	1.24	1.59	2.46	0.98	1.25
New Jersey	1.81	3.31	0.88	1.33	1.56	2.45	0.81	1.24
New York	1.79	2.57	1.03	1.33	1.66	2.58	1.06	1.28
Texas	1.86	2.62	0.95	1.35	1.92	3.02	0.89	1.37
Washington	1.83	2.74	1.06	1.34	1.87	3.04	1.30	1.36
Wisconsin	1.86	2.67	1.15	1.35	1.47	2.14	0.99	1.21
Balance of nation	1.78	2.72	1.00	1.32	1.80	2.36	1.23	1.34

**Table 1-9.**  
**Average DEFF and DEFT for Estimates from the 2002 NSAF Study Area Adult Pair**  
**Sample, Adults in Households with Children, by Study Area**

<b>Study area</b>	<b>All Adults</b>				<b>Low-Income Adults</b>			
	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>
Alabama	1.90	4.02	1.19	1.36	1.91	3.32	1.02	1.37
California	1.74	2.53	1.07	1.31	1.78	2.64	1.22	1.33
Colorado	2.14	3.74	0.75	1.44	1.87	2.93	1.08	1.36
Florida	2.09	3.18	1.15	1.43	2.06	2.90	1.33	1.43
Massachusetts	1.62	2.70	0.91	1.26	1.89	2.64	1.04	1.37
Michigan	1.96	2.58	1.06	1.39	1.83	2.50	0.96	1.35
Minnesota	1.94	3.67	0.90	1.37	1.90	3.22	0.97	1.36
Mississippi	2.19	4.85	0.93	1.46	2.47	5.17	1.07	1.54
New Jersey	2.03	3.64	1.23	1.41	1.83	2.97	1.09	1.34
New York	1.98	3.14	1.18	1.40	1.88	3.42	1.17	1.35
Texas	2.17	3.44	1.09	1.46	2.24	3.52	1.15	1.48
Washington	2.02	3.83	1.19	1.41	2.06	3.03	1.16	1.42
Wisconsin	1.92	3.03	1.10	1.38	1.63	2.78	0.94	1.26
Balance of nation	1.95	2.97	1.10	1.38	1.97	3.11	1.22	1.39

**Table 1-10.**  
**Average DEFF and DEFT for Estimates from the 2002 NSAF Study Area Adult Pair**  
**Sample, Adults in Households with No Children, by Study Area**

<b>Study area</b>	<b>All Adults</b>				<b>Low-Income Adults</b>			
	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>
Alabama	1.92	3.42	1.03	1.37	2.06	3.54	0.95	1.42
California	1.90	3.83	1.01	1.36	2.47	4.70	0.85	1.54
Colorado	1.60	2.66	0.59	1.24	1.83	3.78	0.76	1.33
Florida	1.73	2.91	0.73	1.30	2.14	3.22	1.22	1.46
Massachusetts	1.69	2.73	0.78	1.29	1.95	3.82	0.91	1.37
Michigan	1.63	2.74	0.96	1.26	1.96	3.86	0.90	1.38
Minnesota	1.45	2.66	0.62	1.18	1.58	2.86	1.02	1.25
Mississippi	1.56	2.17	0.90	1.24	1.77	3.13	1.13	1.32
New Jersey	1.72	3.36	0.81	1.29	1.77	2.49	1.15	1.32
New York	1.46	2.41	0.82	1.19	2.00	3.67	1.09	1.39
Texas	1.67	2.54	1.12	1.28	1.84	2.61	0.90	1.34
Washington	1.50	2.51	0.39	1.20	1.53	2.43	0.58	1.22
Wisconsin	1.46	2.63	0.55	1.19	1.85	3.60	0.88	1.33
Balance of nation	1.57	2.61	0.67	1.24	1.98	3.29	1.01	1.39

Round 3 procedure). Before looking at the differences in the design effects for the study areas by round, we evaluate the differences in the design effects owing to the different sampling and weighting methods.

Appendixes A and B contain tables of average design effects for the study areas for the child and adult pair file estimates for Round 1 and Round 2 computed using both the national weights (with the nontelephone sample cases) and the study area weights (without the nontelephone sample but using weights adjusted for nontelephone households). These tables show that despite differences in the sample composition (including and excluding the nontelephone sample) and the weighting procedures, the average design effects computed from the national and study area weights are very similar. The DEFTs computed using the national weights are generally within 2 percent of the DEFTs of those computed from the study area weights, with only a small number of average DEFTs that differ by 5 to 8 percent. This result holds for both Round 1 and for Round 2. These findings indicate that the weighting method, including the differences in including the nontelephone sample, has a relatively minor effect on the average design effects.

The Round 3 average study area DEFTs in tables 1-2 (for the child file estimates) and 1-5 (for the adult pair file estimates) are very similar to the corresponding average study area DEFTs from Round 2 in appendix B. The only consistent and substantial difference in the DEFTs is for Wisconsin. The average DEFTs for the Round 3 Wisconsin estimates are consistently lower than the Round 2 average DEFTs by 5 to 10 percent. The Round 3 average is lower because no separate Milwaukee sample was drawn in Round 3, thus reducing the variance inflation due to the differential within Wisconsin sampling that was needed to support Milwaukee and balance of Wisconsin estimates in the earlier rounds.

### **1.3 Estimates of Change**

Estimates of change between 1997 and 2002 and between 1999 and 2002 are important statistics that can be calculated from three rounds of the NSAF. Estimates of change, like estimates of level discussed in the previous section, are subject to sampling error. The sampling error for estimates of change depends on the variances of the estimates for each round and on the correlation between the samples. The Round 1 and Round 2 samples were correlated because part of the Round 1 sample was retained for Round 2. However, the Round 3 telephone sample was independent of the two earlier rounds. Thus, estimates of change between 1997 and 2002 and between 1999 and 2002 are not correlated. Details on appropriate ways of computing sampling errors for estimates of change are presented in chapter 2. This section provides some guidance to analysts about the precision of estimates of change.

For estimates of level, the design effect is useful for summarizing the precision of estimates of proportions and means. This representation is not as useful for estimates of change. To understand why, consider estimating a proportion in Round 3, say  $\hat{q}_3$ , and let  $v(\hat{q}_3)$  be its estimated variance. The estimated change in the proportion between Round 1 (or Round 2) and Round 3 can be written as  $\Delta = \hat{q}_1 - \hat{q}_3$ . Since the Round 3 sample is independent of the earlier rounds, the estimated variance of the difference is

$$v(\Delta) = v(\hat{q}_1) + v(\hat{q}_3) = D_1 \frac{\hat{q}_1(1-\hat{q}_1)}{n_1} + D_3 \frac{\hat{q}_3(1-\hat{q}_3)}{n_3},$$

where the subscripts indicate the round of data collection,  $D$  is the average design effect, and  $n$  is sample size. Unless the design effects and sample sizes are constant across the rounds, the expression for the design effect for an estimated change is a function of several factors. Since the sample sizes and design effects are not constant across the three rounds of NSAF, the design effect for estimates of change is not very useful.

Despite this, estimating the variance of estimates of change between 1997 and 2002 and between 1999 and 2002 is relatively simple. Chapter 2 describes the procedures used for estimating the variance for estimates of change. Computing variances for estimates of change between 1997 and 1999 is more complex because the samples in those two rounds are not independent. Report No. 4 in the 1999 series discusses this estimation problem.

## 2. METHODS FOR VARIANCE ESTIMATION

Variance estimation procedures have been developed to account for the sample design employed in a complex survey. Using these procedures, such factors as selecting sample clusters in multistage sampling and using differential sampling rates to oversample a targeted subpopulation can be appropriately reflected in estimates of sampling error. The two main methods for estimating variances from a complex survey are replication and Taylor series estimation. Wolter (1985) is a useful reference on the theory and applications of these methods. The next two sections briefly review these methods using essentially the same approaches as used in the first two rounds of NSAF. The third section discusses issues specific to the differences between the sampling and weighting methods used for the national and study area samples in Round 3. The last section of this chapter discusses estimating changes across the three rounds of the NSAF.

### 2.1 Replication Methods

The idea underlying replication is to draw subsamples from the sample, compute the estimate from each subsample, and estimate the variance from the variability of the subsample estimates. Specifically, subsamples of the original *full* sample are selected to calculate subsample estimates of a parameter for which a *full-sample* estimate of interest has been generated. The variability of these subsample estimates about the estimate for the full sample can then be computed. The subsamples are called *replicates* or replicate subsamples and the estimates from the subsamples are called *replicate estimates*. Balanced repeated replication (BRR) and jackknife replication are two approaches to forming subsamples. Rust and Rao (1996) discuss these and other replication methods, show how the units included in the subsample can be defined using variance strata and units, and describe how these methods can be implemented using weights.

Replicate weights facilitate the computation of replicate estimates. Each replicate weight is derived using the same estimation steps as the full sample weight, but including only the subsample of cases comprising each replicate. Once the replicate weights are developed, computing estimates of variance for sample estimates of interest is straightforward. Estimates of variance take the following form:

$$v(\hat{\mathbf{q}}) = c \sum_{k=1}^G (\hat{\mathbf{q}}(k) - \hat{\mathbf{q}})^2, \tag{2-1}$$

where

- $\hat{\mathbf{q}}$  is the estimate of  $\mathbf{q}$  based on the full sample.
- $\hat{\mathbf{q}}(k)$  is the  $k$ -th estimate of  $\mathbf{q}$  based on the observations included in the  $k$ -th replicate.
- $G$  is the total number of replicates formed.
- $c$  is a constant that depends on the replication method.
- $v(\hat{\mathbf{q}})$  is the estimated variance of  $\hat{\mathbf{q}}$ .

In the next chapter, the specific form of equation (2-1) used in the NSAF is presented. The use of WesVar (Westat, 2000) and SUDAAN (1998), two computer software programs that generate estimates of variance (standard errors, CVs, confidence intervals) with a specified set of replicate weights, are covered in chapter 4.

## **2.2 Taylor Series Method**

Another method for estimating variances in complex surveys is based on the Taylor series approximation. A Taylor series linearization of a statistic is formed and then substituted into the formula for calculating the variance of a linear estimate appropriate for the sample design. The Taylor series method relies on the simplicity associated with estimating the variance for a linear statistic even with a complex sample design. In most complex designs, the variance can be estimated by using the variance between PSUs and assuming the sample of PSUs is selected with replacement (Wolter 1985). For this approximation, the strata and PSUs must be defined, similar to the variance estimation strata and units discussed earlier.

SUDAAN (1998), Stata (1999), and SAS are software programs available to produce variance estimates for complex surveys using the Taylor series method. Examples using these programs for the NSAF were presented in *1999 NSAF Variance Estimation*. Those methods are directly applicable for Round 3 and are not otherwise discussed. The Taylor series method was not used to produce the variance estimates from the survey summarized in the first chapter.

## **2.3 Differences between National and Study Area Variance Estimation**

As discussed earlier in this report and in great detail in Report No. 3 in the 2002 series, the weights used to produce the national estimates differ from those used to produce estimates for the study area characteristics. Different weights were needed in the study area because samples of nontelephone households in the study areas, that had been included in previous rounds, were not retained in Round 3. Only a national sample of nontelephone households was included in Round 3. The study area weights were adjusted to account for the lack of a sample of nontelephone households. The national weight did not need this adjustment.

This change in the sample design and weighting procedures also affects variance estimation. Two sets of replicate weights were created—one for the national full sample weight and the second for the study area full sample weight. As described in Report No. 3, the study area weights were adjusted to account for households without a telephone by raking so the study area estimates summed across all study areas (including the balance of the nation) equaled the national area estimates for a set of variables. This form of raking is called sample-based raking (Brick and Kalton 1996) or Info-S raking (Lundström and Särndal 1999) because the control totals are actually estimates and subject to sampling errors.

Operationally, replicate variance estimates can be computed when the weights are sample-based raked using procedures similar to those used for raking to control known totals without error.

The only difference is the replicate sample estimates are adjusted so they equal the national full sample estimate on a replicate-by-replicate basis rather than to a fixed total (in other words, the estimate for replicate  $k$  from the aggregated study areas is adjusted to equal the  $k^{\text{th}}$  replicate estimate from the national sample). This procedure was followed for the study areas weights. Thus, the replicate weights for the study areas can be used in existing software packages that compute replicate variance estimates. The replicate weights capture the variability associated with sample-based control totals.

## 2.4 Estimating Variances of Change Over Time

Because the NSAF collected data in 1997, 1999, and 2002, estimating changes in the characteristics of the population—such as the change in the share of adults without insurance—is an important objective. In this report, the discussion focuses on estimating changes and the precision of those changes over the five years from 1997 to 2002 and over the more recent three years from 1999 to 2002. Estimating changes from 1997 to 1999 is covered in the 1999 weighting and variance estimation reports and is not repeated here.

Estimating the variance of estimates of change between 1997 and 2002 and between 1999 and 2002 is relatively simple, because the sample for 2002 is nearly independent of the sample from the previous rounds. The sample of telephone numbers for Round 3 is completely independent of the sample of telephone numbers in the earlier rounds. For the national sample, about half of the area sample PSUs from the earlier rounds was retained. Within the retained PSUs, this does not necessarily mean the same households or sampled persons within households were retained. Even though the 1997 and 1999 samples were intentionally overlapped and not independent, *1999 NSAF Variance Estimation* (Report No. 4) shows the average national correlation for low-income children was 0.07. With the independent telephone sample in Round 3 and at most half the overlap in the area sample, the correlation will be much less. Given this small overlap, the effect of this dependence on the estimates of totals is so small that it can safely be ignored for variance estimation. At worst, this will slightly overestimate the variance for estimates of change, which is a conservative approach. Even estimates for small subgroups are likely to have low correlations. The study area estimates for Round 3 use only the telephone sample and thus can be considered independent of the estimates from earlier rounds.

To understand the simplicity of estimating change over time from the NSAF assuming independence of the samples, consider estimating a characteristic or total at time  $t$ , say  $\hat{q}_t$ . Let  $v(\hat{q}_t)$  be its estimated variance (the square of the standard error). The estimated change between times  $t_1$  and  $t_2$  for this characteristic or total is  $\Delta = \hat{q}_{t_1} - \hat{q}_{t_2}$ . There is interest in estimating its variance,  $v(\Delta)$ .

Assuming the NSAF sample for Round 3 is independent of the samples for both Round 1 and Round 2, the variance of the difference is the sum of the variances for the two time periods,

$$v(\Delta) = v(\hat{q}_{t_1}) + v(\hat{q}_{t_2}). \quad (2.2)$$

The two variances on the right side of the equation can be computed separately. For example, consider estimating the change between 1997 ( $t_1$ ) and 2002 ( $t_2$ ). The first variance in equation (2.2) can be estimated using the variance estimation procedures described in *1997 NSAF Variance Estimation*, Report No. 4 for Round 1, while the second variance can be computed using the procedures described in this report. The estimated variance of the change can then be summed as indicated in equation (2.2). If the change is between 1999 and 2002, the same procedures can be used, but  $t_1$  now refers to the Round 2 estimate and its variance is computed using the variance estimation procedures described in *1999 NSAF Variance Estimation*, Report No. 4 for Round 2.

This approach is not appropriate for estimating changes between 1997 and 1999 because the samples in those two rounds are not independent. In fact, the sample design for Round 2 was established deliberately to have dependent samples. With dependent samples, the variance of the estimated change has an additional component to account for the correlation in the samples. The estimated variance is

$$v(\Delta) = v(\hat{\mathbf{q}}_{t_1}) + v(\hat{\mathbf{q}}_{t_2}) - 2 \cdot \mathbf{r} \cdot \sqrt{v(\hat{\mathbf{q}}_{t_1})v(\hat{\mathbf{q}}_{t_2})}, \quad (2.3)$$

where the last term accounts for the dependence of the two estimates. When the correlation,  $\mathbf{r}$ , is large and positive, then the variance of the estimated change may be much smaller than obtained from independent samples. With independent samples, the correlation is zero, and equation (2.3) reduces to equation (2.2). Report No. 4 in the 1999 series discusses this estimation problem, including the estimation of the correlation. However, as noted earlier, the 2002 sample is considered independent of the samples from the earlier rounds and equation (2.3) is not necessary for estimating changes between 2002 and the earlier rounds.

### **3. REPLICATION FOR THE NSAF**

The two sections in this chapter describe why replication was chosen as the preferred method for computing variance estimates from the NSAF and the details of the replication structure. The first section discusses the particular form of the jackknife replication method selected for the NSAF. The second section defines the variance strata and units used to construct the replicates.

#### **3.1 Choice of Method**

In Round 1, replication was selected as the primary method to estimate variances for the NSAF, although procedures for using Taylor series methods were also developed. All the key variance estimates included in the methodology series reports, including all the tables of design effects in this report, are computed using replication.

The two main reasons for choosing replication as the preferred method of variance estimation in Round 1 were operational convenience and the ability to reflect all components of the design and estimation in the variability estimates. These two considerations are still valid now. On operational convenience, once the replicate weights are constructed, computing estimates of sampling errors is simple. No special care is needed for subgroups of interest, and no knowledge of the sample design is required. If an estimator is needed that was not previously considered, replication methods can easily develop an appropriate estimate of variance. In such a case, variance estimates using a Taylor series approach would require additional work.

The ability to reflect all components of the design and estimation in the estimates of variability is still the most important reason for choosing replication. Variances are affected by both the nonresponse and poststratification types of adjustments made in developing the NSAF estimates. Replicate weights can be developed that reflect all such aspects of weighting. As noted in the previous chapter, the study area estimates for Round 3 include a new weighting adjustment called sample-based raking. Replication also accounts for this adjustment in the variances of the estimates. On the other hand, the currently available software for the Taylor series method of variance estimation cannot handle such a variety of adjustments. In fact, only some Taylor series software can even account for simple poststratification. Variance estimates using the Taylor series method cannot reflect nonresponse adjustments, raking, and sample-based raking.

Two different replication methods were considered in Round 1. The two methods are a jackknife approach sometimes referred to as a random group methodology (JK1), and a paired jackknife approach (JK2). For an RDD study, the JK1 method is often used because it is appropriate for a single-stage unclustered approach with no explicit strata. The JK2 approach is more frequently employed where first-stage sampling units are selected with two units selected per explicit stratum. Thus, the JK2 approach would be more appropriate for the area sample of PSUs.

In Round 1, the JK2 approach was selected, using the rationale described in Report No. 4 of the 1997 methodology series. The JK2 method was also used for Round 2 and Round 3. In addition

to the original reasons for choosing this method, it is very beneficial to users to retain the same method of replication across rounds.

The JK2 approach treats each pair of sampled telephone numbers as an implicit stratum, where each such stratum is defined by the sort order used in the sample selection of telephone numbers within a study area. It corresponds to the approach used for the area component of the study. The only difference is explicit strata were constructed within each site in the area sample before sample selection of PSUs, and segments are paired instead of telephone numbers. In the JK2 method, the constant,  $c$ , in equation (2-1) is equal to 1. The next sections describe the details on the way replicates were prepared using the JK2 approach.

### **3.2 Design of Replicates**

Two key issues for designing replicates are determining the number of replicates and deciding on the method for constructing the variance estimation strata and units. In the JK2 method, the number of replicates is equal to the number of strata, so the first step is to specify the number of variance estimation strata. In both Round 1 and Round 2, 60 variance estimation strata were used. The same number of replicates was used in Round 3.

The number of variance estimation strata was based on the desire to obtain adequate degrees of freedom to ensure stable estimates of variance, while not having so many replicates that the cost of computing variance estimates becomes unnecessarily high. Generally, at least 30 degrees of freedom are needed to obtain relatively stable variance estimates. A number greater than 30 was targeted because other factors reduce the contribution of a replicate to the total number of degrees of freedom, especially for estimates of subgroups. For example, estimating characteristics of the low-income population requires more strata because the low-income may be disproportionately distributed, with a relatively large share in households without telephones.

For the RDD sample, the variance strata and variance units were created using the same type of procedures as used in Round 1. Sampled telephone numbers were arranged in the same sort order used in sample selection within a study area. Adjacent sampled telephone numbers were paired to establish initial variance estimation strata (the first two sampled phone numbers were the first initial stratum, the third and fourth sampled telephone numbers were the second initial stratum, and so on). Each telephone number in the pair was randomly assigned to either the first or second variance unit within the variance stratum. Each pair was sequentially assigned to one of 60 final variance estimation strata (the first pair to variance estimation stratum 1, the second to stratum 2, ..., the 60th pair to stratum 60, the 61st pair to stratum 1, and so on). As a result, each variance stratum had approximately the same number of telephone numbers for each study area.

The procedures for the area sample only apply to the national weights, since the study area weights exist only for RDD cases. As described in Report No. 4 from Round 1, segments were used as the variance units or ultimate clusters because the sample of first-stage PSUs was small and led to unstable variance estimates. This procedure ignores the component of the variance due

to sampling the first-stage PSU but greatly improves the stability of the variance estimates. The same procedure was used for forming variance units in Round 3.

The Round 3 variance strata and variance units for the area sample were created by sorting the segment sample in the order of selection, then pairing adjacent segments within a PSU to form initial variance strata. For consistency with the previous area design, segments were not paired across study area. In study areas with an odd number of segments, three segments were included in the last variance strata (called a triplet) of the study area. Each initial variance stratum consisted of a pair of segments from the PSU. The initial strata were numbered from 1 to 60, beginning again with 1 after the 60th segment pair as described above for the RDD sample. Each segment in a pair was randomly assigned to variance unit 1 or 2. The initial variance strata were then combined to form 60 variance estimation strata, just as was done for the RDD sample.

For the study area weights, the 60 variance strata and the variance units from the RDD sample were used to compute replicates. For the national weights, the 60 area variance strata and variance units were then combined with the 60 variance strata from the RDD sample, resulting in a total of 60 variance estimation strata. Each variance stratum and variance unit for the national weights contains some RDD and area sample.

Once the variance strata were created, the replicate weights were created. The full replicate weights were constructed by modifying the full-sample base weights. Replicate base weight for replicate  $k$  for record  $i$  was

$$\begin{aligned} w_i^{(k)} &= 2 w_i \text{ if } i \text{ is in variance stratum } k \text{ and variance unit 1,} \\ &= 0 \text{ if } i \text{ is in variance stratum } k \text{ and variance unit 2,} \\ &= w_i \text{ if } i \text{ is not in variance stratum } k. \end{aligned}$$

In variance strata with 3 units, the replicate base weight for replicate  $k$  for record  $i$  is similar to the description above except the factor is  $1.5 w_i$  for two of the three variance units (units 1 and 2 or units 1 and 3) and 0 for the third variance unit.

The same sequence of weighting adjustments applied to the full sample weight was applied to the replicate base weight to create the final replicate weights. Thus, all components of the weighting process were fully reflected in the replicate weights, ranging from household adjustments (nonresponse, adjustment for household noncoverage, and adjustment to control totals) to person adjustments (nonresponse and raking). For the study area replicates, the person adjustments appropriately included the sample-based raking.

## 4. SOFTWARE FOR COMPUTING VARIANCES

Specialized software is needed to correctly compute sampling errors when a complex sample design or estimation scheme is employed. Most standard statistical software packages assume simple random sampling and weights all cases equally when computing estimates of variance. These estimates of sampling errors usually seriously understate the true variability of the survey estimates. Several specialized commercial software products have been developed to analyze data from complex surveys (Lepkowski and Bowles 1996; Cohen 1996; Broene and Rust 1998). Report No. 4 in the Round 2 series, *1999 NSAF Variance Estimation*, describes the procedures for computing sampling error estimates for the NSAF for the following software packages: WesVar, SUDAAN, Stata, SAS, and user-written Stata macros provided by the Urban Institute (Garrett 2000). The macros are given in the appendix of *1999 NSAF Public Use File Data Documentation*, Report No. 11. The Urban Institute also provides an online statistical analysis system that computes standard errors for means and proportions using the Taylor series approach on their web site, at <http://www.urban.org/content/Research/NewFederalism/NSAF/OnlineAnalysis/AnalysisPrelogin.htm>. These procedures are still appropriate and can be used to estimate sampling errors from the Round 3 data.

The following sections briefly describe WesVar, the package used to produce the estimates of design effects given earlier, and SUDAAN. SUDAAN was included in the 1999 report but the SUDAAN package at the time only supported Taylor series methods. In the interim, SUDAAN added replication capability, which is described below. Report No. 4 from the 1999 methodology series contains the pertinent details for all other software packages.

For software packages such as WesVar and SUDAAN, the default number of degrees of freedom is 60 for NSAF data. There are 60 variance estimation strata. Because of the NSAF design and especially for some subgroup estimates, the actual number of degrees of freedom will be smaller than 60 as is typical of complex samples. DiGaetano et al. (1998) discuss the efforts made to ensure that the actual number of degrees of freedom for the NSAF was maximized and thus close to 60. While there is no rule of thumb to easily determine the actual number of degrees of freedom for a complex sample, the practical implications generally will not be substantial because the critical values for  $t$  or  $F$  distributions with 30 and 60 degrees of freedom are very similar. For example, with a 95 percent confidence interval, the  $t$  value is  $t_{30}=1.697$  versus  $t_{60}=1.671$ . Similarly, the critical values for the  $F$  distribution are  $F_{1,30}=4.17$  and  $F_{1,60}=4.00$ . If users are concerned about the effects of underestimating the variance in certain subgroups, a smaller number of degrees of freedom could be used.

### 4.1 WesVar

WesVar is a package developed by Westat for the personal computer (PC). The latest version is 4.2, and this is the version used in the report. Older versions of WesVar could also be used to compute the same estimates described here. WesVar uses replication methods such as the jackknife and BRR to compute variance estimates. Through replicates, adjustments made during

weighting (nonresponse, raking) can be taken into account by making the same adjustments to each replicate separately. Replication is computer-intensive, but powerful PCs have largely eliminated this as an issue. However, it is still possible that for very large data sets the computations will exceed the capacity of the machine or take a long time. Although replication can be used for most estimates, replication techniques are not necessarily appropriate for all sample statistics of interest. Special care is needed when trying to estimate median, quartiles, or other quantiles. Direct estimates of quantiles using the jackknife method are not supported, but an alternative method is supported.

WesVar is an interactive program centered on sessions called *workbooks*. A workbook is a file linked to a specific WesVar data set. In a workbook, the user can request descriptive statistics and regression models, as well as analyze and create new statistics. The information about the design is incorporated into the replicate weights when the data file is created. Regression requests support both linear and logistic regression (both dichotomous and multinomial). Outputs include statistics of interest, such as the sum of weights, means, totals, percentages, ratios, regression coefficients, and log odds-ratios, along with their corresponding standard errors, CV, and confidence intervals. Chi-square tests of independence are performed on two-way tables, and goodness of fit statistics are produced for regression models. WesVar can also estimate linear combinations of parameter estimates and perform tests of hypotheses. It can output design effects for all the above statistics except computed statistics such as ratios.

Exhibit 4-1 demonstrates how to compute sample estimates and their corresponding standard errors using WesVar. In the example, proportions, totals, and their standard errors are computed for the variable POV200 (Focal child lives in a family below twice the federal poverty level) from the Round 3 national child data file. This variable is 1 if family income is 200 percent or less of the federal poverty level and is 2 otherwise.

The information about the design is already incorporated in the replicate weights. The variance estimation method (JK2) used is provided during the creation of the WesVar data file. The NSAF data files have one full sample final weight and 60 final replicate weights. The output of a request used to produce Round 3 estimates for the POV200 variable follows.

## **4.2 SUDAAN**

SUDAAN (Software for the Statistical Analysis of Correlated Data) is a package developed by Research Triangle Institute (RTI) to analyze data from complex sample surveys. The latest version is 8.0 and is used in this report. Like WesVar, SUDAAN computes standard errors of the estimates taking into account the survey design. SUDAAN and WesVar produce the same point estimates, but SUDAAN has one procedure that uses a first-order Taylor series approximation and another that uses replication. Because the Taylor series method is described in Report No. 4

**Exhibit 4-1.**  
**Computing Sample Estimates and Their Corresponding Standard Errors Using WesVar**

**WesVar Estimates of Proportions and Totals**

```

WESVAR VERSION NUMBER :          4.2
TIME THE JOB EXECUTED :          16:07:53 12/30/2003
INPUT DATASET NAME :             \UICYC3\R3 weighting\Sample-based raking
                                  \Ch\Wesvar\Data\nsaf_ch50_wesvar_sb1j.var
TIME THE INPUT DATASET CREATED : 10:14:03 11/21/2003
FULL SAMPLE WEIGHT :             CHW0
REPLICATE WEIGHTS :              CHW1...CHW60
VARIANCE ESTIMATION METHOD :     JK2

OPTION COMPLETE :                ON
OPTION FUNCTION LOG :            ON
OPTION VARIABLE LABEL :         OFF
OPTION VALUE LABEL :            OFF
OPTION OUTPUT REPLICATE ESTIMATES : OFF
FINITE POPULATION CORRECTION FACTOR : 1
VALUE OF ALPHA (CONFIDENCE LEVEL %) : 0.05000 (95.00000 %)
DEGREES OF FREEDOM :           60
t VALUE :                       2

ANALYSIS VARIABLES :            None Specified.
COMPUTED STATISTIC :            None Specified.
TABLE(S) :                      POV200

FACTOR(S) :                     1

NUMBER OF REPLICATES :          60
NUMBER OF OBSERVATIONS READ :   34332
WEIGHTED NUMBER OF OBSERVATIONS READ : 72645014
  
```

**WesVar Output**

POV200	STATISTIC	EST_TYPE	ESTIMATE	STDERROR	CELL_n	DENOM_n	DEFF
1	SUM_WTS	VALUE	27487924	313476.2	13030	N/A	N/A
2	SUM_WTS	VALUE	45157090	313476.2	21302	N/A	N/A
MARGINAL	SUM_WTS	VALUE	72645014	0	34332	N/A	N/A
1	SUM_WTS	PERCENT	37.84	0.432	13030	34332	2.718
2	SUM_WTS	PERCENT	62.16	0.432	21302	34332	2.718
MARGINAL	SUM_WTS	PERCENT	100	.	34332	34332	.

of the 1999 Methodology series it is not discussed in detail here. It is worth noting that the Taylor series procedures in SUDAAN do not fully account for complex weighting schemes such as nonresponse or raking, but the replication procedures that use appropriately prepared replicate weights do account for these methods.

For descriptive statistics, SUDAAN offers three procedures: PROC CROSSTAB for categorical variables, PROC DESCRIPT for both continuous and categorical variables, and PROC RATIO for ratios of either continuous or categorical variables. These procedures can compute statistics of interest, such as the sum of weights, means, totals, and percentages along with their corresponding standard errors, design effects, and confidence intervals. PROC DESCRIPT and RATIO estimate linear contrasts of parameter estimates and perform hypothesis tests. They can also reflect poststratification of the weights in the variance estimates by adding files of control totals. PROC CROSSTAB will produce chi-square goodness of fit tests for two-dimensional tables.

Exhibit 4-2 demonstrates how to compute sample estimates and their corresponding standard errors using SUDAAN. PROC CROSSTAB is used to produce estimates of proportions and totals with standard errors and design effects. The jackknife method is used. The output of a request used to produce Round 3 estimates from the national child data file for the POV200 variable follows.

### **4.3 Review of Software Output**

The two software programs for estimating sampling errors from complex samples were used to compute sampling errors from the 2002 NSAF child file to illustrate some of the issues discussed earlier. Table 4-1 summarizes the output from the programs. As noted previously, WesVar and SUDAAN produce the same point estimates and are based on the same sample size. The standard errors and design effect estimates may vary depending on how they are computed. However, in this case the standard errors are identical because both WesVar and SUDAAN use replication using the weights provided on the data file. The design effects are also equal because we chose the jackknife option for computing design effects from SUDAAN. Note that in Report No. 4 of the 1999 methodology series, the SUDAAN and WesVar sampling error estimates differed because the SUDAAN procedure shown there used the Taylor series method of variance estimation.

**Exhibit 4-2.**  
**Computing Sample Estimates and Their Corresponding Standard Errors Using SUDAAN**

**SUDAAN Estimates of Proportions and Totals Using PROC CROSSTAB**

```
43      /*****
44
45      Proc CROSSTAB.
46
47      *****/
48
49      proc crosstab data = r3_sud  design = JACKKNIFE  ;
50      weight CHW0;          /* final fs weight */
51      JACKWGTS CHW1-CHW60 /ADJACK=1;
52
53      subgroup POV200 ;
54      levels 2;
55      tables POV200 ;
56
57      print nsum wsum sewgt deffwgt totper setot defftot /style=nchs
58      wsumfmt=f10.0 setotfmt=f7.3 defftotfmt=f7.3;
59      run ;
```

Opened SAS data file R3\_SUD for reading.

NOTE: There were 34332 observations read from the data set WORK.R3\_SUD.

NOTE: The PROCEDURE CROSSTAB printed pages 1-2.

NOTE: PROCEDURE CROSSTAB used:

real time	1.98 seconds
cpu time	0.99 seconds

60

61

NOTE: SAS Institute Inc., SAS Campus Drive, Cary, NC USA 27513-2414

NOTE: The SAS System used:

real time	15.40 seconds
cpu time	3.34 seconds

**Exhibit 4-2.**

**Computing Sample Estimates and Their Corresponding Standard Errors Using SUDAAN (continued)**

**SUDAAN Output**

Project Name: NSAF 2002 - Variance Report SUDAAN runs

S U D A A N  
 Software for the Statistical Analysis of Correlated Data  
 Copyright      Research Triangle Institute      January 2003  
 Release 8.0.2

Number of observations read      : 34332      Weighted count : 72645014  
 Denominator degrees of freedom :      60

Date: 01-28-2004      Research Triangle Institute  
 Time: 15:20:12      The CROSSTAB Procedure

Variance Estimation Method: Replicate Weight Jackknife  
 by: POV200

---

POV200	Sample Size	Weighted Size	SE Weighted	Tot Percent	SE Tot Percent	DEFF Tot Percent #4
Total	34332	72645014	0.00	100.00	0.000	.
1	13030	27487924	313476.20	37.84	0.432	2.718
2	21302	45157090	313476.20	62.16	0.432	2.718

---

**Table 4-1.**  
**Estimates of Children Computed Using WesVar and SUDAAN**

<b>Variable</b>	<b>EST_TYPE</b>	<b>N</b>	<b>ESTIMATE</b>	<b>STDERROR</b>	<b>DEFF</b>
	<b>(1)</b>	<b>(2)</b>	<b>(3)</b>	<b>(4)</b>	<b>(5)</b>
<b>POV100</b>					
1	TOTAL	5,141	12,013,064	254,751	
2	TOTAL	29,191	60,631,950	254,751	
MARGINAL	TOTAL	34,332	72,645,014	0	
1	PERCENT	5,141	16.54	0.351	3.059
2	PERCENT	29,191	83.46	0.351	3.059
MARGINAL	PERCENT	34,332	100		
<b>POV200</b>					
1	TOTAL	13,030	27,487,924	313,476	
2	TOTAL	21,302	45,157,090	313,476	
MARGINAL	TOTAL	34,332	72,645,014	0	
1	PERCENT	13,030	37.84	0.432	2.718
2	PERCENT	21,302	62.16	0.432	2.718
MARGINAL	PERCENT	34,332	100		
<b>UCURNINS</b>					
0	TOTAL	31,449	65,757,047	278,757	
1	TOTAL	2,883	6,887,967	278,757	
MARGINAL	TOTAL	34,332	72,645,014	0	
0	PERCENT	31,449	90.52	0.384	5.89
1	PERCENT	2,883	9.48	0.384	5.89
MARGINAL	PERCENT	34,332	100		

**Table 4-1.**  
**Estimates of Children Computed Using WesVar and SUDAAN (continued)**

<b>Variable</b>	<b>EST_TYPE</b>	<b>N</b>	<b>ESTIMATE</b>	<b>STDERROR</b>	<b>DEFF</b>
	<b>(1)</b>	<b>(2)</b>	<b>(3)</b>	<b>(4)</b>	<b>(5)</b>
<b>JAFDCF</b>					
(missing)	TOTAL	32,108	67,631,557	193,938	
1	TOTAL	1,889	4,335,245	178,547	
2	TOTAL	335	678,212	73,435	
MARGINAL	TOTAL	34,332	72,645,014	0	
(missing)	PERCENT	32,108	93.1	0.267	3.808
1	PERCENT	1,889	5.97	0.246	3.696
2	PERCENT	335	0.93	0.101	3.793
MARGINAL	PERCENT	34,332	100		
<b>KLUNCH</b>					
(missing)	TOTAL	16,337	34,049,616	320,464	
1	TOTAL	9,299	20,200,429	318,093	
2	TOTAL	8,696	18,394,969	315,427	
MARGINAL	TOTAL	34,332	72,645,014	0	
(missing)	PERCENT	16,337	46.87	0.441	2.683
1	PERCENT	9,299	27.81	0.438	3.279
2	PERCENT	8,696	25.32	0.434	3.423
MARGINAL	PERCENT	34,332	100		
<b>JFSTAMPF</b>					
1	TOTAL	4,403	10,178,392	256,929	
2	TOTAL	29,929	62,466,622	256,929	
MARGINAL	TOTAL	34,332	72,645,014	0	
1	PERCENT	4,403	14.01	0.354	3.565
2	PERCENT	29,929	85.99	0.354	3.565
MARGINAL	PERCENT	34,332	100		

## REFERENCES

### Methodology References

- Brick, J. Michael, Pam Broene, David Ferraro, Tom Hankins, and Teresa Strickler. 2000. *1999 NSAF Sample Estimation Survey Weights*. Methodology Report No. 3.
- Brick, J. Micheal, Pam Broene, David Ferraro, Tom Hankins, Carin Rauch, and Teresa Strickler. 2000. *1999 NSAF Variance Estimation*. Methodology Report No. 4.
- Brick, J. Michael, Gary Shapiro, Ismael Flores-Cervantes, David Ferraro, and Teresa Strickler. 1999. *1997 NSAF Snapshot Survey Weighting*. Methodology Report No. 3.
- Brick, M. Forthcoming. *2002 NSAF Sample Estimation Survey Weights*. Methodology Report No. 3.
- Brick, Mike, David Ferraro, Teresa Strickler, and Benmei Liu. 2003. *2002 NSAF Sample Design*. Methodology Report No. 2.
- Converse, Nate, Adam Safir, Fritz Scheuren, Rebecca Steinbach, and Kevin Wang. 2001. *1999 NSAF Public Use File User's Guide*. Methodology Report No. 11.
- Flores-Cervantes, Ismael, J. Michael Brick, and Ralph DiGaetano. 1999. *1997 NSAF Variance Estimation*. Methodology Report No. 4.

### General References

- Brick, J. Michael, and G. Kalton. 1996. "Handling Missing Data in Survey Research." *Statistical Methods in Medical Research* 5:215–38.
- Brick, J. Michael, E. Tubbs, M. Collins, and M. J. Nolin. 1997. *Unit and Item Response, Weighting and Imputation Procedures in the 1993 National Household Education Survey*. Working Paper 97-05. U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.
- Broene, Pam, and K. Rust. 1998. *Strengths and Limitations of Using SUDAAN, Stata, and WesVarPC for Computing Variances from NCES Data Sets*. Working Paper. U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.
- Cohen, S. 1996. *An Evaluation of Alternative PC based software packages developed for the Analysis of Complex Survey Data*. Rockville, MD: Agency for Health Care Policy and Research.

- DiGaetano, R., J. Michael Brick, and Ismael Flores-Cervantes. 1998. "Preserving Degrees of Freedom in the Multi-Mode, Multi-Site Survey." In *Proceedings of the Survey Research Methods Section* (475–80). Alexandria, VA: American Statistical Association.
- Flores-Cervantes, I., Shapiro, G., and Brock-Roth, S. 1998. Effect of Oversampling by Poverty Status in an RDD Survey. In *Proceedings of the Survey Research Methods Section* (469–74). Alexandria, VA: American Statistical Association.
- Garrett, Bowen. 2000. BRMEAN, BRTAB, BRREG, BRLOGIT, BRMLOGIT. Software developed by the Urban Institute, Washington, D.C.
- Kish, L. 1965. *Survey Sampling*. New York: John Wiley and Sons.
- . 1992. "Weighting for Unequal Pi." *Journal of Official Statistics* 8:183–200.
- Lepkowski, J., and J. Bowles. 1996. "Sampling Error Software for Personal Computers." *The Survey Statistician* 35:10-16.
- Lundström, S., and C.-E. Särndal. 1999. "Calibration as a Standard Method for Treatment of Nonresponse." *Journal of Official Statistics* 15:305–27.
- Rust, K. F., and J. N. K. Rao. 1996. "Variance Estimation for Complex Surveys using Replication Techniques." *Statistical Methods in Medical Research* 5:282–310.
- Skinner, C. J., D. Holt, and T. M. F. Smith. 1989. *Analysis of Complex Surveys*. New York: John Wiley and Sons.
- Stata. 1999. *Stata User's Guide*, Release 6. College Station, TX: Stata Corporation.
- SUDAAN. 1998. *SUDAAN User's Manual*, Release 8.0.2. Research Triangle Park, NC: Research Triangle Institute.
- Verma, V., C. Scott, and C. O'Muircheartaigh. 1980. "Sample Designs and Sampling Errors for the World Fertility Survey." *Journal of the Royal Statistical Society A* 143:431–73.
- Waksberg, J., J. Michael Brick, Gary Shapiro, Ismael Flores-Cervantes, and B. Bell. 1997. "Dual-Frame RDD and Area Sample for Household Survey with Particular Focus on Low-Income Population." In *Proceedings of the Survey Research Methods Section* (713–18). Alexandria, VA: American Statistical Association.
- WesVar. 2000. *WesVar™ 4.0 User's Guide*. Rockville, MD: Westat.
- Wolter, K. M. 1985. *Introduction to Variance Estimation*. New York: Springer-Verlag.

## Appendix A

### Round 1 Design Effect Tables

**Table A-1.**  
**Average DEFF and DEFT for Estimates from the Round 1 1997 NSAF National Sample,**  
**by Type of Estimates**

<b>Type of estimate</b>	<b>All</b>				<b>Low-Income</b>			
	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>
<b>Child</b>								
All race	4.99	8.44	3.11	2.21	6.03	10.59	3.08	2.42
Black	5.87	9.90	1.48	2.39	6.05	9.36	1.79	2.43
Hispanic	3.78	5.70	1.86	1.93	3.81	6.06	1.93	1.93
<b>Adult</b>								
All race	5.01	7.10	2.69	2.23	5.65	9.12	3.51	2.36
Black	5.34	9.29	3.19	2.29	5.51	9.42	3.41	2.33
Hispanic	3.86	6.29	1.99	1.94	4.23	6.53	2.34	2.04
Option B adults	5.18	9.49	2.69	2.24	5.75	11.36	3.23	2.35
Households with children	5.48	8.64	2.83	2.32	6.13	11.27	3.59	2.44
Adult-only households	4.21	5.96	2.16	2.04	4.24	5.62	2.84	2.05

**Table A-2.**  
**Average DEFF and DEFT for Estimates from the 1997 NSAF National Child Sample, by Study Area**

Study area	All Children				Low-Income Children			
	Average	Maximum	Minimum	DEFT	Average	Maximum	Minimum	DEFT
Alabama	1.54	2.87	0.94	1.23	1.96	3.90	1.17	1.39
California	1.89	3.26	0.84	1.36	2.43	5.48	1.13	1.53
Colorado	1.60	2.69	1.10	1.26	1.89	2.97	1.05	1.37
Florida	1.74	3.79	0.84	1.30	2.06	4.75	0.63	1.41
Massachusetts	1.61	2.33	1.12	1.26	2.27	3.78	0.95	1.49
Michigan	1.71	5.15	1.03	1.29	2.04	6.69	0.61	1.40
Minnesota	2.21	8.91	1.08	1.44	2.92	13.20	1.23	1.64
Mississippi	2.03	5.54	0.95	1.40	2.34	6.58	1.21	1.50
New Jersey	1.56	2.35	0.92	1.24	2.12	3.40	1.12	1.44
New York	1.48	2.76	0.81	1.20	1.75	3.14	0.78	1.31
Texas	1.70	2.37	1.01	1.30	2.02	3.02	1.12	1.41
Washington	1.62	2.20	0.72	1.27	2.19	2.97	1.37	1.47
Wisconsin	1.88	3.54	0.71	1.35	2.26	3.65	1.16	1.49
Balance of nation	1.81	3.25	1.00	1.33	2.23	4.29	1.11	1.47
National	4.99	8.44	3.11	2.21	6.03	10.59	3.08	2.42

**Table A-3.**  
**Average DEFF and DEFT for Estimates from the 1997 NSAF National Adult Pair Sample, by Study Area**

Study area	All Adults				Low-Income Adults			
	Average	Maximum	Minimum	DEFT	Average	Maximum	Minimum	DEFT
Alabama	2.20	3.66	1.28	1.47	2.53	6.31	1.00	1.55
California	1.81	2.96	1.14	1.33	2.42	5.09	1.53	1.54
Colorado	1.87	2.90	1.20	1.36	2.12	2.86	1.44	1.45
Florida	2.22	3.37	1.22	1.48	2.70	3.83	1.66	1.63
Massachusetts	2.01	3.62	0.96	1.40	2.44	4.11	1.60	1.55
Michigan	1.92	4.96	0.95	1.36	2.37	4.32	1.33	1.52
Minnesota	2.17	4.06	1.51	1.46	2.89	6.88	1.57	1.67
Mississippi	2.25	4.14	1.45	1.49	2.51	5.44	1.41	1.56
New Jersey	1.95	4.30	1.15	1.38	2.73	3.76	1.86	1.65
New York	2.03	3.13	1.22	1.41	2.65	3.76	1.55	1.61
Texas	2.32	3.50	1.39	1.51	2.79	4.89	1.20	1.64
Washington	1.92	3.23	1.05	1.37	2.29	3.46	1.38	1.50
Wisconsin	1.95	2.91	1.18	1.38	2.37	3.94	1.15	1.53
Balance of nation	1.74	2.62	0.96	1.31	2.15	3.28	1.26	1.45
National	5.01	7.10	2.69	2.23	5.65	9.12	3.51	2.36

**Table A-4.**  
**Average DEFF and DEFT for Estimates from the 1997 NSAF Study Area Child Sample, by Study Area**

Study area	All Children				Low-Income Children			
	Average	Maximum	Minimum	DEFT	Average	Maximum	Minimum	DEFT
Alabama	1.63	2.73	0.92	1.27	1.77	3.04	1.00	1.32
California	1.62	2.91	0.78	1.26	1.90	3.12	1.11	1.36
Colorado	1.60	2.56	0.85	1.26	1.82	2.60	0.99	1.34
Florida	1.59	2.57	0.90	1.25	1.90	3.10	0.71	1.36
Massachusetts	1.51	2.28	0.88	1.22	1.98	3.50	0.96	1.39
Michigan	1.75	2.99	1.10	1.31	1.82	3.30	0.65	1.33
Minnesota	1.55	2.30	0.85	1.24	1.91	2.82	1.06	1.37
Mississippi	1.71	2.83	0.70	1.29	1.79	2.68	0.85	1.33
New Jersey	1.61	2.52	1.02	1.26	2.25	3.54	1.03	1.49
New York	1.52	2.71	0.78	1.22	1.80	3.00	0.77	1.33
Texas	1.57	2.24	1.00	1.24	1.66	2.43	1.07	1.28
Washington	1.62	2.61	0.76	1.26	2.15	3.65	1.31	1.46
Wisconsin	1.91	3.49	0.89	1.36	2.27	3.67	1.18	1.50
Balance of nation	1.57	2.20	1.00	1.25	1.86	2.65	1.25	1.36
National	4.56	8.10	2.92	2.12	5.13	9.27	3.16	2.24

**Table A-5.**  
**Average DEFF and DEFT for Estimates from the 1997 NSAF Study Area Child Sample, Hispanic Children, by Study Area**

Study area	All Children				Low-Income Children			
	Average	Maximum	Minimum	DEFT	Average	Maximum	Minimum	DEFT
Alabama	1.60	1.92	0.76	1.33	1.70	1.98	1.21	1.30
California	1.70	2.84	0.93	1.29	1.84	3.38	1.13	1.34
Colorado	1.67	2.39	1.08	1.29	1.68	2.65	1.07	1.29
Florida	1.54	2.16	0.78	1.23	1.67	2.46	0.98	1.28
Massachusetts	1.86	4.15	0.86	1.34	1.95	4.65	0.71	1.37
Michigan	1.59	2.26	1.13	1.26	1.51	2.01	0.94	1.22
Minnesota	1.79	2.82	0.90	1.32	2.12	3.24	1.14	1.44
Mississippi	1.62	1.87	1.30	1.27	n/a	n/a	n/a	n/a
New Jersey	1.92	2.72	0.76	1.37	2.17	2.94	0.81	1.46
New York	1.52	2.58	0.72	1.22	1.57	2.61	0.76	1.24
Texas	1.63	2.41	1.02	1.27	1.59	2.52	0.87	1.25
Washington	1.81	2.53	0.85	1.33	1.98	2.86	1.00	1.39
Wisconsin	2.17	3.23	1.25	1.47	2.53	3.57	1.34	1.58
Balance of nation	1.68	2.66	0.81	1.28	1.77	2.85	1.08	1.32
National	3.62	5.16	1.91	1.89	3.68	5.43	1.75	1.90

*Note:* In MS for low-income children, all sample sizes are less than 10, which is the threshold we are using for reporting design effects.

**Table A-6.**  
**Average DEFF and DEFT for Estimates from the 1997 NSAF Study Area Child Sample, Black Children, by Study Area**

Study area	All Children				Low-Income Children			
	Average	Maximum	Minimum	DEFT	Average	Maximum	Minimum	DEFT
Alabama	1.69	2.74	0.79	1.29	1.66	2.65	0.98	1.28
California	1.68	2.86	1.03	1.29	1.66	2.57	0.97	1.27
Colorado	1.96	2.89	0.96	1.39	1.91	3.05	1.25	1.37
Florida	2.03	4.33	0.87	1.40	2.08	4.57	0.91	1.42
Massachusetts	2.03	3.80	1.09	1.41	1.80	3.07	0.85	1.32
Michigan	1.72	2.69	1.08	1.30	1.63	2.70	1.12	1.27
Minnesota	1.69	2.47	0.65	1.29	1.69	2.65	0.70	1.29
Mississippi	1.79	2.80	0.85	1.32	1.79	2.89	0.94	1.32
New Jersey	1.76	2.93	1.05	1.32	1.96	4.15	0.73	1.38
New York	1.58	2.94	0.83	1.24	1.65	2.89	0.94	1.27
Texas	1.52	2.30	0.59	1.22	1.62	2.29	0.67	1.26
Washington	1.75	2.65	0.77	1.31	2.16	3.08	0.95	1.46
Wisconsin	2.29	5.49	0.90	1.49	2.49	6.68	0.96	1.54
Balance of nation	1.68	2.80	0.83	1.28	1.64	2.72	0.96	1.27
National	5.45	8.77	1.64	2.31	5.53	8.47	1.99	2.33

**Table A-7.**  
**Average DEFF and DEFT for Estimates from the 1997 NSAF Study Area Adult Pair Sample, by Study Area**

Study area	All Adults				Low-Income Adults			
	Average	Maximum	Minimum	DEFT	Average	Maximum	Minimum	DEFT
Alabama	2.16	3.43	1.36	1.46	2.45	5.62	1.45	1.54
California	1.81	2.74	1.14	1.34	2.26	3.47	1.47	1.49
Colorado	1.88	2.80	1.05	1.36	2.14	3.20	1.45	1.45
Florida	2.44	3.66	1.64	1.55	2.99	4.81	1.67	1.71
Massachusetts	2.03	3.25	0.98	1.41	2.31	3.90	1.42	1.51
Michigan	2.13	5.51	0.97	1.43	2.51	3.96	1.23	1.57
Minnesota	1.93	3.52	1.21	1.37	2.38	3.51	1.73	1.53
Mississippi	2.22	3.25	1.45	1.48	2.39	3.99	1.24	1.53
New Jersey	2.05	4.50	1.15	1.42	2.78	3.80	1.78	1.66
New York	2.10	3.07	1.35	1.44	2.72	3.86	1.59	1.64
Texas	2.25	3.11	0.94	1.48	2.53	3.79	0.82	1.57
Washington	1.94	3.50	0.99	1.38	2.34	3.35	1.10	1.51
Wisconsin	2.06	3.38	1.20	1.42	2.45	3.96	1.32	1.55
Balance of nation	1.69	2.54	0.91	1.29	1.99	3.55	1.26	1.40
National	4.87	6.06	2.59	2.20	5.14	6.20	3.63	2.26

**Table A-8.****Average DEFF and DEFT for Estimates from the 1997 NSAF Study Area Adult Pair Sample, Hispanic Adults, by Study Area**

<b>Study area</b>	<b>All Adults</b>				<b>Low-Income Adults</b>			
	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>
Alabama	1.94	2.71	1.33	1.40	1.85	2.03	1.54	1.36
California	2.15	3.27	0.88	1.45	2.45	4.70	1.43	1.55
Colorado	1.86	3.25	0.65	1.34	2.18	4.08	0.86	1.44
Florida	2.64	6.61	1.65	1.60	3.17	7.66	1.20	1.75
Massachusetts	2.64	4.22	1.51	1.61	2.51	5.17	0.85	1.56
Michigan	1.85	2.69	1.00	1.34	2.15	3.53	1.07	1.45
Minnesota	1.74	2.71	0.80	1.30	2.02	3.83	1.20	1.40
Mississippi	1.31	1.62	0.78	1.14	1.14	1.37	0.88	1.06
New Jersey	2.24	4.23	0.87	1.47	2.43	4.21	0.99	1.54
New York	2.74	5.40	1.34	1.63	3.03	4.41	1.67	1.73
Texas	2.18	3.92	0.67	1.45	2.30	4.98	0.78	1.49
Washington	2.17	3.45	1.06	1.45	2.52	4.34	0.70	1.55
Wisconsin	3.86	7.71	1.75	1.91	4.97	10.83	0.86	2.14
Balance of nation	1.69	2.38	0.79	1.29	1.88	2.90	0.96	1.36
National	3.86	5.85	2.27	1.95	4.39	5.69	2.54	2.08

**Table A-9.**  
**Average DEFF and DEFT for Estimates from the 1997 NSAF Study Area Adult Pair Sample, Black Adults, by Study Area**

Study area	All Adults				Low-Income Adults			
	Average	Maximum	Minimum	DEFT	Average	Maximum	Minimum	DEFT
Alabama	2.03	2.79	1.09	1.42	2.02	2.82	1.14	1.41
California	1.90	3.28	0.53	1.35	2.08	3.62	0.92	1.41
Colorado	1.55	2.39	0.44	1.22	1.47	2.02	0.84	1.21
Florida	2.84	5.45	1.67	1.66	2.71	8.07	1.67	1.61
Massachusetts	1.99	3.07	0.93	1.40	2.05	3.45	1.09	1.42
Michigan	2.13	4.45	0.73	1.43	2.44	4.20	0.46	1.52
Minnesota	1.54	2.41	0.84	1.23	1.44	2.07	0.94	1.20
Mississippi	2.05	3.32	1.10	1.42	2.05	3.59	1.19	1.42
New Jersey	1.70	2.58	1.00	1.30	2.02	3.59	1.05	1.40
New York	1.89	2.62	0.95	1.36	1.98	3.35	1.32	1.39
Texas	2.15	5.11	1.19	1.44	2.07	2.94	0.88	1.43
Washington	1.58	2.18	1.18	1.25	1.57	2.24	1.07	1.25
Wisconsin	2.05	5.59	1.35	1.41	1.96	2.46	1.49	1.39
Balance of nation	1.69	3.49	1.06	1.28	1.65	2.88	0.73	1.27
National	4.88	8.26	3.40	2.19	4.88	6.57	2.68	2.20

**Table A-10.**  
**Average DEFF and DEFT for Estimates from the 1997 NSAF Study Area Adult Pair Sample, “Option B” Adults,**  
**by Study Area**

<b>Study area</b>	<b>All Adults</b>				<b>Low-Income Adults</b>			
	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>
Alabama	1.55	2.36	1.00	1.24	1.70	2.70	0.96	1.29
California	1.65	2.65	0.92	1.27	1.86	2.65	0.99	1.35
Colorado	1.66	3.29	1.03	1.27	1.86	3.57	1.17	1.35
Florida	1.89	2.48	0.95	1.36	2.11	2.96	0.94	1.44
Massachusetts	1.59	2.79	0.77	1.25	1.77	2.52	0.82	1.31
Michigan	1.71	2.83	0.49	1.29	1.89	2.73	0.61	1.35
Minnesota	1.81	3.30	0.86	1.33	2.14	3.42	1.02	1.44
Mississippi	1.66	4.15	0.98	1.27	1.46	2.03	0.90	1.20
New Jersey	1.79	3.46	0.98	1.32	2.15	3.07	0.95	1.45
New York	1.75	2.97	0.73	1.31	2.14	3.87	0.84	1.44
Texas	1.72	2.58	0.71	1.29	2.01	3.39	0.60	1.38
Washington	1.73	2.50	0.96	1.30	1.99	3.07	1.07	1.39
Wisconsin	1.82	3.19	0.82	1.33	1.99	3.62	0.97	1.39
Balance of nation	1.54	2.20	0.93	1.23	1.83	3.49	1.06	1.34
National	4.56	7.76	2.72	2.12	4.80	8.71	2.99	2.17

**Table A-11.**  
**Average DEFF and DEFT for Estimates from the 1997 NSAF Study Area Adult Pair Sample, Adults in Households**  
**with Children, by Study Area**

Study area	All Adults				Low-Income Adults			
	Average	Maximum	Minimum	DEFT	Average	Maximum	Minimum	DEFT
Alabama	1.91	2.82	0.93	1.37	2.17	3.10	1.18	1.46
California	1.97	4.11	1.21	1.39	2.24	4.39	1.33	1.48
Colorado	1.93	3.29	0.85	1.38	2.15	3.71	1.08	1.45
Florida	2.43	3.77	1.61	1.54	3.01	4.88	1.64	1.72
Massachusetts	2.01	2.89	1.11	1.41	2.44	3.67	1.35	1.55
Michigan	2.41	4.96	0.48	1.52	2.43	4.80	0.70	1.53
Minnesota	1.88	3.27	1.26	1.36	2.21	3.84	1.14	1.47
Mississippi	2.15	3.98	1.24	1.45	2.14	4.18	1.03	1.44
New Jersey	1.99	3.30	1.10	1.39	2.30	3.76	1.22	1.50
New York	2.55	4.47	1.42	1.57	2.96	5.18	1.29	1.69
Texas	1.95	3.17	0.84	1.38	2.22	3.27	0.84	1.46
Washington	2.09	3.01	0.88	1.43	2.23	3.40	1.17	1.47
Wisconsin	2.07	3.53	1.01	1.41	2.21	3.51	1.24	1.47
Balance of nation	1.69	2.53	1.10	1.29	2.00	2.97	1.14	1.40
National	4.97	7.63	2.77	2.21	5.22	8.05	3.35	2.27

**Table A-12.**  
**Average DEFF and DEFT for Estimates from the 1997 NSAF Study Area Adult Pair Sample, Adults in Households  
with No Children, by Study Area**

<b>Study area</b>	<b>All Adults</b>				<b>Low-Income Adults</b>			
	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>
Alabama	1.68	3.30	1.19	1.29	1.65	3.53	0.90	1.27
California	1.70	2.50	1.00	1.29	1.72	2.26	1.13	1.31
Colorado	1.45	2.37	0.87	1.20	1.56	2.24	0.92	1.24
Florida	1.65	2.95	0.94	1.27	1.78	3.36	0.96	1.32
Massachusetts	1.67	2.54	0.91	1.28	1.62	2.37	1.16	1.27
Michigan	1.64	2.65	0.87	1.27	2.13	3.57	1.06	1.44
Minnesota	1.67	3.77	0.75	1.27	2.00	3.59	0.95	1.40
Mississippi	1.89	2.67	0.73	1.36	1.93	2.67	1.03	1.38
New Jersey	1.69	2.88	0.75	1.29	2.26	3.51	1.23	1.49
New York	1.51	2.34	0.77	1.22	1.80	2.59	1.15	1.33
Texas	1.82	3.07	0.91	1.33	1.92	3.86	0.89	1.36
Washington	1.62	2.85	1.06	1.26	1.74	2.85	0.96	1.31
Wisconsin	1.70	2.76	0.80	1.29	2.03	3.93	0.70	1.40
Balance of nation	1.63	2.37	0.87	1.27	1.68	2.39	1.10	1.29
National	4.26	5.91	2.34	2.05	4.07	5.61	2.71	2.01

## Appendix B

### Round 2 Design Effect Tables

**Table B-1.**  
**Average DEFF and DEFT for Estimates from the Round 2 1999 NSAF National Sample, by Type of Estimates**

Type of estimate	All				Low-Income			
	Average	Maximum	Minimum	DEFT	Average	Maximum	Minimum	DEFT
<b>Child</b>								
All race	3.99	7.20	2.65	1.99	4.15	7.00	2.59	2.03
Black	4.64	6.93	1.72	2.13	5.08	7.53	2.04	2.23
Hispanic	3.10	4.37	1.54	1.75	3.22	4.42	1.78	1.78
<b>Adult</b>								
All race	4.39	7.35	2.39	2.06	4.25	6.24	2.63	2.05
Black	5.69	10.24	1.91	2.34	4.54	7.91	2.07	2.10
Hispanic	3.57	5.20	1.63	1.87	3.53	5.52	1.48	1.85
Option B adults	4.12	6.69	1.84	2.00	3.73	5.70	1.93	1.91
Households with children	4.66	8.08	2.29	2.13	4.08	7.70	1.78	1.99
Adult-only households	3.39	5.11	1.91	1.82	3.85	5.29	2.00	1.95

B-1

**Table B-2.**  
**Average DEFF and DEFT for Estimates from the 1999 NSAF National Child Sample, by Study Area**

Study area	All Children				Low-Income Children			
	Average	Maximum	Minimum	DEFT	Average	Maximum	Minimum	DEFT
Alabama	1.71	2.59	1.07	1.30	1.88	2.94	1.09	1.36
California	1.71	2.92	0.98	1.30	1.95	2.95	1.00	1.39
Colorado	1.77	3.02	0.77	1.32	1.98	3.20	0.75	1.39
Florida	1.89	4.94	0.82	1.35	2.02	4.79	0.85	1.40
Massachusetts	1.63	2.71	1.08	1.27	1.83	2.78	1.20	1.34
Michigan	2.21	9.05	1.00	1.44	2.19	7.75	1.04	1.44
Minnesota	2.03	5.49	0.99	1.39	2.35	5.85	1.02	1.50
Mississippi	1.67	5.36	0.73	1.26	1.76	5.32	0.95	1.30
New Jersey	1.81	2.67	1.23	1.34	1.82	2.65	0.89	1.34
New York	1.70	2.36	0.92	1.30	1.82	2.90	1.05	1.34
Texas	1.87	3.22	0.91	1.35	2.02	3.16	1.17	1.41
Washington	1.68	2.87	0.92	1.29	1.78	2.43	1.03	1.33
Wisconsin	2.04	3.11	1.18	1.42	2.22	4.04	1.01	1.47
Balance of nation	1.70	2.99	1.10	1.29	1.86	3.06	1.00	1.35
National	3.99	7.20	2.65	1.99	4.15	7.00	2.59	2.03

**Table B-3.**  
**Average DEFF and DEFT for Estimates from the 1999 NSAF National Adult Pair Sample, by Study Area**

Study area	All Adults				Low-Income Adults			
	Average	Maximum	Minimum	DEFT	Average	Maximum	Minimum	DEFT
Alabama	2.12	4.27	0.95	1.43	2.46	6.36	1.18	1.54
California	1.84	3.01	1.12	1.35	1.99	3.49	1.22	1.40
Colorado	3.06	8.86	0.73	1.70	2.23	4.31	0.99	1.47
Florida	2.39	3.78	1.31	1.53	2.41	3.26	0.95	1.54
Massachusetts	2.32	4.14	1.44	1.51	2.45	4.19	1.37	1.55
Michigan	2.20	3.64	1.16	1.47	2.20	3.83	1.15	1.46
Minnesota	1.98	3.38	1.22	1.39	2.13	3.24	1.45	1.45
Mississippi	1.90	4.91	0.95	1.35	1.94	5.83	1.02	1.36
New Jersey	2.69	4.81	1.52	1.62	2.50	4.48	1.31	1.56
New York	2.35	3.81	1.33	1.52	2.32	3.56	0.74	1.51
Texas	3.11	5.41	1.56	1.74	3.23	6.80	1.44	1.77
Washington	2.20	4.29	1.33	1.47	2.13	3.31	1.27	1.45
Wisconsin	2.86	7.52	1.18	1.65	2.67	4.36	1.60	1.61
Balance of nation	1.82	3.81	1.06	1.32	1.96	2.59	0.95	1.39
National	4.39	7.35	2.39	2.06	4.25	6.24	2.63	2.05

**Table B-4.**  
**Average DEFF and DEFT for Estimates from the 1999 NSAF Study Area Child Sample, by Study Area**

Study area	All Children				Low-Income Children			
	Average	Maximum	Minimum	DEFT	Average	Maximum	Minimum	DEFT
Alabama	1.84	2.94	1.02	1.34	1.88	2.90	0.84	1.36
California	1.82	3.64	0.96	1.34	2.05	3.58	0.84	1.42
Colorado	1.72	2.56	0.84	1.30	1.82	2.77	0.88	1.34
Florida	1.83	2.96	0.88	1.34	1.93	2.86	0.82	1.38
Massachusetts	1.68	2.85	1.13	1.29	1.83	2.94	1.18	1.34
Michigan	1.96	2.96	1.15	1.39	1.79	2.41	1.12	1.33
Minnesota	1.72	2.95	1.01	1.30	1.93	3.43	1.12	1.37
Mississippi	1.59	2.91	0.74	1.25	1.60	2.67	0.92	1.25
New Jersey	1.84	2.65	1.23	1.35	1.84	2.86	0.86	1.35
New York	1.78	2.81	0.92	1.32	1.83	2.94	1.05	1.34
Texas	1.73	2.82	0.93	1.30	1.77	2.77	1.11	1.32
Washington	1.75	2.83	0.82	1.31	1.74	2.40	1.00	1.31
Wisconsin	2.07	3.42	1.15	1.43	2.15	4.46	1.06	1.45
Balance of nation	1.65	2.93	0.99	1.27	1.66	2.87	1.10	1.28
National	4.13	6.98	2.48	2.02	3.97	6.14	2.97	1.98

B-4

**Table B-5.****Average DEFF and DEFT for Estimates from the 1999 NSAF Study Area Child Sample, Hispanic Children, by Study Area**

Study area	All Children				Low-Income Children			
	Average	Maximum	Minimum	DEFT	Average	Maximum	Minimum	DEFT
Alabama	2.18	3.57	0.86	1.55	1.64	2.12	1.07	1.27
California	2.23	4.29	0.78	1.47	2.37	4.18	0.85	1.52
Colorado	1.71	2.69	0.94	1.30	1.80	2.61	0.81	1.33
Florida	1.59	2.63	0.78	1.25	1.61	2.44	0.73	1.26
Massachusetts	1.66	2.38	0.73	1.28	1.69	2.57	0.78	1.29
Michigan	1.60	2.36	1.02	1.26	1.60	2.35	0.82	1.25
Minnesota	1.78	2.61	0.85	1.32	1.68	2.36	1.16	1.29
Mississippi	1.23	1.24	1.22	1.11	n/a	n/a	n/a	n/a
New Jersey	1.66	2.76	0.83	1.28	1.67	2.68	0.89	1.28
New York	1.65	2.83	1.07	1.28	1.61	2.74	0.88	1.26
Texas	1.65	2.81	0.77	1.28	1.60	2.46	0.81	1.26
Washington	1.86	3.30	0.92	1.35	1.75	3.07	0.85	1.31
Wisconsin	2.00	3.21	0.43	1.39	2.26	3.73	0.56	1.48
Balance of nation	1.49	2.17	0.88	1.21	1.39	1.82	0.86	1.18
National	3.16	4.79	1.51	1.76	3.25	4.68	1.67	1.79

*Note:* In MS for low-income children, all sample sizes are less than 10, which is the threshold we are using for reporting design effects.

**Table B-6.**  
**Average DEFF and DEFT for Estimates from the 1999 NSAF Study Area Child Sample, Black Children, by Study Area**

Study area	All Children				Low-Income Children			
	Average	Maximum	Minimum	DEFT	Average	Maximum	Minimum	DEFT
Alabama	1.80	2.87	0.57	1.33	1.84	2.82	0.58	1.34
California	1.98	3.21	0.91	1.39	1.87	3.07	0.50	1.34
Colorado	1.64	2.04	1.26	1.28	1.73	2.10	1.03	1.31
Florida	2.05	2.87	0.96	1.42	1.99	3.11	0.94	1.40
Massachusetts	1.97	3.96	0.69	1.38	2.15	4.23	0.84	1.44
Michigan	1.67	2.36	0.97	1.29	1.49	2.02	0.91	1.22
Minnesota	2.02	2.99	0.97	1.40	1.90	2.90	1.10	1.36
Mississippi	1.55	2.48	0.75	1.24	1.50	2.32	0.86	1.22
New Jersey	2.01	3.16	1.25	1.40	1.93	2.80	0.90	1.38
New York	1.77	2.71	0.95	1.32	1.86	3.13	0.98	1.35
Texas	1.98	3.19	0.65	1.39	1.94	3.05	0.68	1.38
Washington	1.67	2.56	1.19	1.28	1.71	3.14	1.20	1.30
Wisconsin	2.07	2.99	0.85	1.42	1.96	3.02	0.87	1.38
Balance of nation	1.72	3.00	1.08	1.30	1.65	2.84	1.13	1.28
National	4.57	7.73	1.69	2.12	4.79	7.58	1.99	2.17

**Table B-7.**  
**Average DEFF and DEFT for Estimates from the 1999 NSAF Study Area Adult Pair Sample, by Study Area**

Study area	All Adults				Low-Income Adults			
	Average	Maximum	Minimum	DEFT	Average	Maximum	Minimum	DEFT
Alabama	2.26	4.34	1.13	1.48	2.56	5.62	1.31	1.58
California	1.87	3.42	1.34	1.35	1.95	2.85	1.30	1.39
Colorado	3.16	9.21	0.66	1.73	2.24	3.92	1.16	1.47
Florida	2.63	4.21	1.30	1.61	2.47	3.56	0.95	1.56
Massachusetts	2.45	4.34	1.68	1.55	2.56	3.99	1.48	1.58
Michigan	2.40	3.84	1.37	1.53	2.56	3.87	1.33	1.58
Minnesota	1.94	2.96	1.19	1.38	2.03	3.55	1.30	1.41
Mississippi	1.90	2.97	0.94	1.36	1.85	3.34	1.00	1.34
New Jersey	2.81	5.02	1.86	1.66	2.65	4.30	1.36	1.61
New York	2.40	3.70	1.50	1.54	2.37	3.64	0.73	1.52
Texas	3.34	5.86	1.51	1.80	3.35	7.15	1.41	1.79
Washington	2.36	4.00	1.34	1.52	2.18	3.19	1.15	1.46
Wisconsin	2.99	6.44	1.36	1.70	2.81	4.90	1.68	1.66
Balance of nation	1.88	3.78	1.06	1.35	1.96	2.88	1.33	1.39
National	4.57	7.94	2.45	2.10	4.19	5.83	2.65	2.04

**Table B-8.****Average DEFF and DEFT for Estimates from the 1999 NSAF Study Area Adult Pair Sample, Hispanic Adults, by Study Area**

<b>Study area</b>	<b>All Adults</b>				<b>Low-Income Adults</b>			
	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>
Alabama	3.30	6.72	1.19	1.72	2.09	3.00	1.27	1.43
California	1.96	2.94	0.90	1.38	2.13	3.32	1.14	1.44
Colorado	2.66	4.37	1.06	1.61	2.10	3.63	1.18	1.43
Florida	3.37	5.30	1.66	1.81	2.80	4.56	1.48	1.66
Massachusetts	2.50	4.67	1.21	1.55	2.40	4.30	1.01	1.52
Michigan	2.20	3.10	1.53	1.48	1.92	2.66	1.13	1.38
Minnesota	1.95	3.09	0.99	1.37	1.71	3.34	0.86	1.28
Mississippi	2.47	3.74	1.01	1.52	1.87	2.42	1.32	1.35
New Jersey	2.39	3.70	0.97	1.52	2.07	3.72	0.43	1.42
New York	2.46	4.20	1.45	1.55	2.49	4.41	1.44	1.56
Texas	3.69	6.42	1.59	1.89	3.63	6.71	1.03	1.86
Washington	2.18	4.17	1.09	1.44	1.91	3.53	0.94	1.36
Wisconsin	2.95	6.38	0.65	1.64	2.22	4.11	0.78	1.45
Balance of nation	2.09	4.07	1.05	1.43	1.83	3.74	0.99	1.33
National	3.93	5.94	2.08	1.96	3.76	5.89	1.73	1.91

**Table B-9.**  
**Average DEFF and DEFT for Estimates from the 1999 NSAF Study Area Adult Pair Sample, Black Adults, by Study Area**

Study area	All Adults				Low-Income Adults			
	Average	Maximum	Minimum	DEFT	Average	Maximum	Minimum	DEFT
Alabama	2.02	2.95	0.95	1.41	2.09	3.19	1.20	1.43
California	1.62	3.08	0.73	1.24	1.56	3.08	0.67	1.22
Colorado	2.03	4.43	0.32	1.33	1.30	2.12	0.67	1.12
Florida	2.67	4.23	1.26	1.62	2.16	3.81	1.09	1.45
Massachusetts	2.84	5.14	1.41	1.67	2.91	5.86	1.24	1.67
Michigan	2.23	3.95	0.68	1.45	2.23	5.20	0.73	1.45
Minnesota	2.02	3.10	0.92	1.40	2.22	5.43	1.07	1.45
Mississippi	1.82	2.97	0.96	1.33	1.69	3.03	0.98	1.29
New Jersey	2.01	3.18	1.25	1.40	2.40	4.83	1.15	1.52
New York	2.48	5.75	1.11	1.53	1.94	3.17	1.22	1.38
Texas	2.44	5.46	1.27	1.54	2.13	3.21	1.37	1.45
Washington	2.29	3.18	1.57	1.50	1.69	2.87	0.94	1.29
Wisconsin	2.06	4.13	0.64	1.41	1.87	2.94	1.05	1.35
Balance of nation	2.27	4.45	1.03	1.48	1.77	3.38	0.87	1.31
National	5.73	10.90	2.78	2.35	4.31	8.23	2.27	2.05

**Table B-10.**  
**Average DEFF and DEFT for Estimates from the 1999 NSAF Study Area Adult Pair Sample,**  
**“Option B” Adults, by Study Area**

<b>Study area</b>	<b>All Adults</b>				<b>Low-Income Adults</b>			
	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>
Alabama	1.96	2.56	1.19	1.39	2.06	2.75	1.35	1.43
California	1.85	2.99	0.85	1.34	1.86	3.08	0.96	1.35
Colorado	2.00	3.56	0.95	1.40	1.90	3.06	0.88	1.37
Florida	2.02	3.36	0.94	1.40	2.07	3.25	0.90	1.42
Massachusetts	1.75	2.57	0.95	1.31	1.71	2.94	0.88	1.29
Michigan	2.18	3.56	1.08	1.46	2.16	3.78	1.08	1.45
Minnesota	1.96	3.73	1.22	1.39	1.96	2.91	1.10	1.39
Mississippi	1.79	2.84	1.08	1.33	1.66	2.48	1.12	1.28
New Jersey	2.24	3.59	1.14	1.48	1.92	2.90	0.96	1.37
New York	2.12	3.79	0.92	1.44	1.85	2.81	0.81	1.34
Texas	2.10	3.59	1.13	1.43	2.14	3.22	1.27	1.45
Washington	1.94	3.70	0.84	1.37	1.97	3.33	0.88	1.39
Wisconsin	2.22	3.23	0.86	1.47	2.19	3.26	1.20	1.46
Balance of nation	1.78	2.57	0.83	1.32	1.66	2.39	0.80	1.27
National	4.25	6.41	1.92	2.04	3.70	4.87	1.85	1.91

**Table B-11.**  
**Average DEFF and DEFT for Estimates from the 1999 NSAF Study Area Adult Pair Sample,**  
**Adults in Households with Children, by Study Area**

<b>Study area</b>	<b>All Adults</b>				<b>Low-Income Adults</b>			
	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>
Alabama	2.05	2.86	0.97	1.42	2.05	3.22	1.31	1.42
California	1.91	3.04	0.81	1.36	1.96	3.19	0.89	1.38
Colorado	2.41	4.84	1.55	1.53	2.16	4.13	1.22	1.45
Florida	2.83	5.71	1.41	1.66	2.53	4.38	1.14	1.57
Massachusetts	1.97	4.78	0.95	1.38	2.19	5.41	0.80	1.44
Michigan	2.31	4.10	1.12	1.50	2.30	3.88	1.42	1.50
Minnesota	2.15	4.21	1.33	1.45	2.05	4.19	0.98	1.41
Mississippi	1.91	3.88	1.10	1.36	1.74	3.00	1.12	1.31
New Jersey	2.67	4.91	1.56	1.61	2.06	3.58	0.85	1.41
New York	2.70	4.63	1.99	1.63	2.33	4.86	0.92	1.50
Texas	2.51	4.28	1.01	1.56	2.42	4.12	1.36	1.53
Washington	2.15	4.18	1.20	1.45	2.06	3.07	0.85	1.42
Wisconsin	2.41	3.55	1.40	1.54	2.42	3.69	1.07	1.54
Balance of nation	2.00	3.14	1.23	1.40	1.81	2.61	0.95	1.33
National	4.82	8.10	2.50	2.17	4.02	5.48	1.92	1.98

**Table B-12.**  
**Average DEFF and DEFT for Estimates from the 1999 NSAF Study Area Adult Pair Sample,**  
**Adults in Households with No Children, by Study Area**

<b>Study area</b>	<b>All Adults</b>				<b>Low-Income Adults</b>			
	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>	<b>Average</b>	<b>Maximum</b>	<b>Minimum</b>	<b>DEFT</b>
Alabama	1.69	3.06	0.65	1.28	2.06	3.03	0.89	1.42
California	1.41	2.90	0.78	1.17	1.56	2.75	0.83	1.24
Colorado	1.70	5.94	0.56	1.27	1.47	2.61	0.65	1.19
Florida	1.69	2.65	0.95	1.29	1.74	2.61	1.13	1.31
Massachusetts	1.51	2.45	0.94	1.22	1.84	2.68	1.00	1.34
Michigan	1.71	2.79	0.81	1.29	2.11	3.34	1.03	1.44
Minnesota	1.45	2.33	0.65	1.19	1.69	2.64	0.90	1.29
Mississippi	1.55	2.09	0.58	1.23	1.72	2.91	0.82	1.29
New Jersey	1.45	2.00	0.94	1.20	1.60	3.20	0.89	1.25
New York	1.57	2.55	1.03	1.24	1.78	2.66	1.19	1.33
Texas	1.70	3.17	0.80	1.28	1.93	4.35	1.04	1.36
Washington	1.45	2.19	1.05	1.20	1.44	2.09	0.80	1.19
Wisconsin	1.86	4.14	0.57	1.33	1.99	3.49	0.85	1.39
Balance of nation	1.67	2.88	0.81	1.28	2.06	3.25	1.09	1.42
National	3.55	5.89	1.87	1.86	3.86	5.46	2.09	1.95

## Appendix C

### Design Effect Statistics

**Table C-1.**  
**Statistics Used to Compute Estimates of Average Design Effects**

<b>Child</b>	<b>Adult</b>
JAFDCF = 1	JOWNCAR = 1
JFSTAMPF = 1	UMARSTAT_R = 1
KBREAK = 1	IEMPNOW = 1
KLUNCH = 1	U_FTPT = 1
MCUTMEAL = 1	UHINS4_R = 1
MLESSRNT = 1	JFSTAMPF = 1
MPAYRENT = 1	MPAYRENT = 1
MPUBHOUS = 1	MTELTRAN = 1
MTELTRAN = 1	MCUTMEAL = 1
UCNGHL = 1	UCURMCD = 1
UCURMCD = 1	MFDWORRY = 1
UCURNINS = 1	UNOCON = 1
UENGNEG = 1	U_LFSR = 1
UNOCON = 1	MFDLACK = 1
UOUTNEG = 1	LWHCRDT = 1
UREADNEG = 1	MPUBHOUS = 1
USRC_NO = 1	MLESSRNT = 1
FWHDEN = 1	LWHTRN = 1
FWHMED = 1	U_USBORN = 2
NVOLUNT_R = 1	ITEMP = 1
NRELIG_R = 1	
UFAMSTR = 2	
UFAMSTR = 3	
UFAMSTR = 4	
UACTNEG = 1	
UAGGNEG = 1	
UMH2NEG = 1	
U_USBORN = 2	