NSAF

Methodology Reports

NSAF Survey Methods and Data Reliability

Report No. 1

Prepared by:

Natalie Abi-Habib
Adam Safir
Timothy Triplett
The Urban Institute

Assessing the New Federalism
An Urban Institute Program to Assess Changing Social Policies

Preface

Report No. 1: NSAF Survey Methods and Data Reliability is the first report in a series describing the methodology of the National Survey of America's Families (NSAF). One component of the *Assessing the New Federalism Project* at the Urban Institute, conducted in partnership with Child Trends, 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 survey sample is representative of the nation as a whole and of 13 focal states, and therefore allows for both national as well as state-level analysis. Westat conducted data collection for the survey.

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 includes the following reports:

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	NSAF questionnaire

About this Report

This first report in the Methodology Series provides readers with an introduction to the National Survey of America's Families, its sample design, data collection techniques, and estimation methods. An overview is also provided describing the survey's dual-frame design, the format of interviews, and the types of questions asked. In addition, the

methods used to minimize errors and compensate for those that are unavoidable in data collection are described. Finally, the report presents information on the survey's resulting reliability—both in terms of sampling and nonsampling 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, D.C. 20037

E-mail: nsaf@ui.urban.org
Website: http://newfederalism.urban.org/nsaf

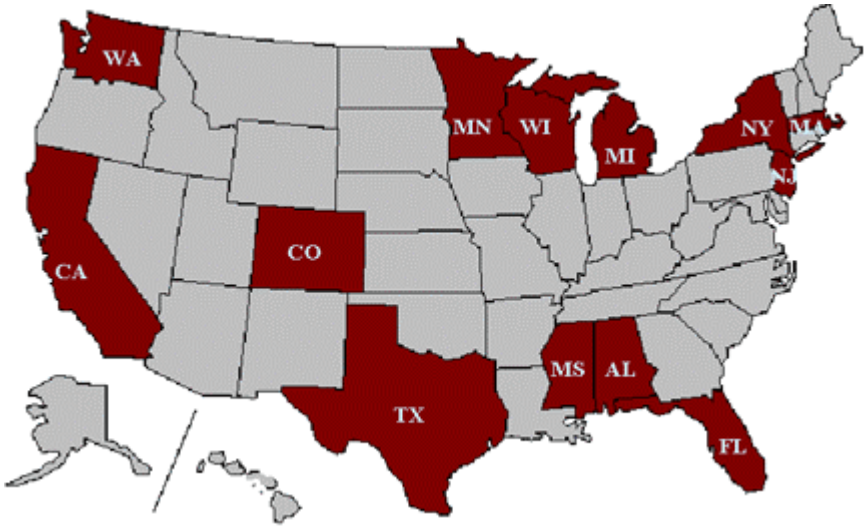
Adam Safir
and
Tim Triplett

Survey Methods and Data Reliability

Introduction

The National Survey of America's Families (NSAF) is a survey of the economic, health, and social characteristics of children, adults under the age of 65, and their families. NSAF data collection was conducted for the Urban Institute and Child Trends by Westat, a nationally renowned survey research firm. The survey was administered three times, in 1997, 1999, and 2002. In each round, interviews were conducted with over 40,000 families, yielding information on more than 100,000 persons under the age of 65. Data collection for the third round of the survey was conducted from February 2002 through October 2002.

Figure 1. Targeted NSAF States



The survey sample is representative of the civilian, noninstitutionalized population under the age of 65 in 13 states and the balance of the nation (figure 1).[1] The 13 focal states, or states selected for in-depth study, were selected for reasons related to how well they represented the nation as a whole on characteristics important to the goal of the survey. Collectively these 13 states account for over half of the U.S. population and represent a broad array of government programs, fiscal capacity, and demographic characteristics. As with virtually all household surveys, some important segments of the population (e.g., the

homeless) could not be sampled because of their living arrangements and therefore are not included in the survey results. A small fraction of the sample consisted of "linguistically isolated" households, where neither English nor Spanish was spoken by any person within the household. Individuals in these living arrangements were not interviewed.[2]

Survey Process

Sample Design. The survey features a dual-frame design, with sample drawn separately for each of the 13 state areas and for the balance of the nation, and an oversample of low-income families with children (see Report No. 2 in the NSAF Methodology Series for details of the oversample selection in each round). The main frame consisted of a random-digit dial (RDD) sample of telephone households. Serving as the supplementary frame, an area sample of nontelephone households was designed to capture respondents living in households without land-line telephones.

RDD Sample. Telephone households were screened to determine eligibility for an extended interview. All households screened as low-income and as having children were administered an extended interview. Higher-income households with children and all households without children but with a resident under the age of 65 were subsampled at varying rates prior to in-depth questioning (subsampling based on predetermined characteristics is a standard and efficient means of ensuring adequate sample sizes for elements from populations of interest). Households without any residents under the age of 65 were excluded from the survey. In all, 133,503 telephone households were screened in 2002, resulting in 43,157 extended interviews in telephone households.

Area Sample. Selection of the area sample began with an area probability frame of residential census block groups, in which households were given in-person screening interviews to determine nontelephone status. All households screened as not having a land-line telephone were administered the extended interview, regardless of family income or presence of children. For the purposes of the survey, wireless telephone-only households were treated as nontelephone households. Because only a small

fraction of households do not have a telephone, block groups from the 1990 Census with a very high percentage of telephone households were not included in the area sample frame. A special coverage adjustment was made during the weighting process to account for persons in nontelephone households in these block groups that by design were excluded from the frame.

In 2002, the size of the area sample was substantially reduced. A total of 730 area sample households were screened, yielding completed interviews in 578 area sample households. The decision to reduce the size of the area sample was based on considerations such as cost, stability of the variance estimates, new research on statistical adjustment methods, and lower estimates of the percentage on households without telephones. By comparison, in 1999, interviews were completed in 1,486 area sample households, and in 1997, interviews were completed in 1,163 area sample households.

In response to the sample size reduction, it became necessary to create separate estimation domains:

- A national-level sample that includes nontelephone households, and
- A state-level sample that intentionally excludes nontelephone households.

To account for these differences, a national-level weight is provided for the purpose of producing national-level estimates (which include nontelephone interviews). A separate state-level weight, which includes a model-based adjustment to account for the exclusion of nontelephone households, is provided for producing state-level estimates. Additional information about the use of the survey estimation weights can be found in Report No. 11: Public Use User's Guide.

In 1999, there was a partial overlap of 1997 sampling units; approximately 60 percent of the starting sample for the 1999 RDD sample was drawn from the 1997 sample. In general, a sample overlap is expected to simultaneously improve the precision of change estimates and introduce the possibility of a new source of bias in change estimates. The precision of change estimates will be improved because of the induced correlation over time. On the other hand, biases could result from increased nonresponse and respondent

conditioning (for more details on overlap sampling methodology, see Report No. 2 in the 1999 NSAF Methodology Series). NSAF telephone numbers associated with households that had successfully completed screener interviews in 1997 were retained for extended interviews at a higher rate than telephone numbers that had been finalized in 1997 with a noncontact, refusal, or nonworking screener result code. A small number of households from the 1997 area sample frame were also retained, using a slightly different methodology. It should be noted that use of the overlap in 1999 does not imply a longitudinal design; in fact, no attempt was made across rounds to interview the same respondent associated with a common telephone number or street address, nor is any means of identifying overlap households provided. Unlike in 1999, the 2002 sample design did not incorporate an overlap component.

Data Collection. For both the RDD and area sample, interviewing was conducted in two stages. First, a short, five-minute screening interview was administered to determine household eligibility. Following the screener interview, a more detailed, 27- to 50-minute extended interview was administered to eligible households. Telephone interviewers located in centralized facilities conducted all interviews using computer-assisted telephone interviewing (CATI). In an effort to minimize mode effects, in-person interviewers provided cellular telephones to respondents in nontelephone households to connect them with interviewing centers for the CATI interview. As such, interviews were conducted in essentially the same way in both telephone and nontelephone households.

Adults with Children. In households with children under the age of 18, up to two children were sampled for in-depth study: one under the age of 6 and another between the age of 6 and 17. Interviews were conducted with the most knowledgeable adult (MKA)—that is, the adult in the household who was most knowledgeable about the health care, education, and well-being of the sampled child. Interviews with MKAs collected data on the sampled children, the MKA and his or her spouse/partner, and their families. It was possible to have different MKAs for each child, although it was more common to have a single respondent (usually the mother) for both children. In each round of the NSAF, over 28,000 extended interviews were conducted with the primary caregivers of children.

Adults without Children. In households without children, depending on the size of the household, one or two adults between the age of 18 and 64 were randomly selected for the extended interview. Similarly, in households with children, in addition to the MKA, one or two additional adults under the age of 65 who did not have any of their own children under the age of 18 living with them were sampled and interviewed. The childless adult interview, for the most part, covered the same questions as the MKA interview except that detailed questions on children were not asked. In each round, over 15,000 extended interviews were conducted with adults between the age of 18 and 64, who did not have any of their own children under the age of 18 living with them.

Questionnaire Content. The complete questionnaire for both 1997 and 2002 can be found in Report No. 12 of the 1997 and 2002 NSAF Methodology Series. The 1999 questionnaire is included in Report No. 1 of the 1999 NSAF Methodology Series. Table 1 illustrates the broad range of topics covered by the NSAF questionnaire.

Table 1. NSAF Question Topics

Child well-being	**Economic security**
Child behavior problems	Child support
Child care use	Employment and Earnings
Child education and cognitive development	Family income
Child social and positive development	Food security
Family environment	Housing and economic hardship
Family stress	Welfare program participation
Family structure	**Other areas**
Parent/adult psychological well-being	Attitudes on welfare, work, and raising children
Health and health care	Demographics
Health care coverage	Household composition
Health care use and access	Social services issues
Health status/limitations	

Some questions ask about family circumstances at the time of the survey; others ask about the 12 or more months prior to the interview. The 1999 and 2002 questionnaires include revisions, additions, and deletions to the original survey instrument in response to changes in public programs between rounds, such as the introduction of the State Children's Health Insurance Program (SCHIP) and the subsequent expansion of the

program to include adults. In addition, some questions were modified if not operating as expected. A significant improvement was made in 1999, for instance, in the survey's approach to asking about nativity, leading to better estimates of immigration status. Documentation of questionnaire changes can be found in Report No. 1 in the 1999 NSAF Methodology Series and Report No. 12 in the 2002 NSAF Methodology Series.

Estimation Methods

Weights. Responses to NSAF items were weighted to provide approximately unbiased aggregate estimates for each study area and for the country as a whole. The NSAF survey weights are designed to

- Compensate for differential probabilities of selection for households and persons;
- Reduce biases occurring where nonrespondents have different characteristics than respondents;
- Adjust, to the extent possible, for undercoverage in the sampling frames and in the conduct of the survey; and
- Reduce the variance of the estimates by using auxiliary information.

The weighting process can be described as involving three stages for both the RDD and the area sample:

1. The first stage is the computation of the base weight. The base weight is the inverse of the probability of selection, which accounts for the unequal screening rates. This weight also includes an adjustment for the planned exclusion of nontelephone households from the Census block groups excluded due to high telephone penetration rates and for the subsampling of persons in selected households (see Triplett and Abi-Habib 2003).
2. The second stage is an adjustment for unit nonresponse (of entire households and persons who did not respond to the survey). In this stage, the weights of

respondents in particular groups are adjusted to account for the nonrespondents in their group.

3. In the third stage, the nonresponse-adjusted weights are post-stratified and raked to agree with independent population totals derived from U.S. Census Bureau sources on the dimensions of age, education, ethnicity, gender, race, and housing tenure.[3] This is done for each study area and the nation as a whole.

The three stages listed above incorporate data from the screener interview to create household-level weights and data from the extended interview to create person- and family-level weights. Each set of weights is produced twice, once for national-level estimates and once for state-level estimation.

Item Nonresponse. A separate estimation issue concerns the handling of item nonresponse—that is, when an interview was obtained but some specific questionnaire items are missing. For most questions the item nonresponse rates were very low, often less than 1 percent. As is the case with any household survey containing questions about sensitive information (such as income and mortgage amounts), the NSAF occasionally encounters significant levels of item nonresponse. In particular, nonresponse rates for items related to income are generally higher (approaching 20 percent in some cases); however, these rates are consistent with those of other major household surveys.

Nearly all questions on earnings, employment, and family income were imputed when missing, as were selected items on health insurance coverage, health care use and access, and housing and economic hardship. The imputation of missing responses makes the data easier to use; for example, imputing missing income responses permits the calculation of family income and poverty measures for all sampled families. Also, imputation partially adjusts for sample bias, since the characteristics of nonrespondents may differ from those of respondents.

All missing responses were imputed at the person level, except for the housing and economic hardship items, which were imputed at the household level. A standard “hot deck” method was used to impute missing responses in all three rounds of the NSAF. In a

hot-deck imputation, a respondent who fails to respond to a particular question is given a value reported by a different respondent with similar characteristics. This approach is the most common method used to assign values for missing responses in large-scale household surveys. For example, the March Current Population Survey (CPS), the source of official annual estimates of poverty, uses hot-deck imputation.

Sampling Errors

The standard error of an estimate measures the uncertainty of that estimate due to the fact that the sample selected is just one of the many possible samples that could have been chosen (thus, it is often called sampling error).[4] The uncertainty in survey estimates due to sampling can also be expressed as a margin of error. Table 2 presents the average margins of error for NSAF estimates, calculated at the 95 percent confidence level. That is, there is a 95 percent probability that the true population value will fall within the interval specified by the margin of error.

Table 2. Margin of Error and Sample Size Comparisons

	1997		1999		2002	
	Margin of error (%)	*Sample size*	*Margin of error (%)*	*Sample size*	*Margin of error (%)*	*Sample size*
Children						
National						
All	1.17	34,439	1.03	35,938	0.94	34,332
Low-income	1.79	17,471	1.70	13,615	1.53	13,030
State						
All	2.53	2,388	2.65	2,414	2.57	2,200
Low-income	3.87	1,205	4.40	904	4.40	826
Adults						
National						
All	0.80	75,525	0.74	74,719	0.79	70,577
Low-income	1.34	29,677	1.31	23,511	1.40	22,401
State						
All	1.95	5,190	2.16	4,955	2.22	4,497
Low-income	3.39	2,029	3.83	1,529	4.03	1,401

Nationally, the margins of error for the 1999 and 2002 estimates were designed to be somewhat lower than those for 1997.[5] At the state level, the overall decline in poverty rates led to a decrease in the oversample of low-income children and an increase in the average margin of error. Even so, the relatively small average margin of error for state-level estimates of low-income children and adults remains one of the main achievements of the NSAF design across all rounds.

Nonsampling Errors

While sampling error is usually associated with the precision of survey estimates, nonsampling errors are usually (but not exclusively) associated with bias in survey estimates. Bias is simply the degree to which an estimated survey statistic (such as a proportion, total, or mean) differs from the true population value. Biases in survey estimates can arise for many reasons. Below, we provide some examples of NSAF methodological work focusing on three potential sources of bias: undercoverage, nonresponse, and problems relating to measurement.

Undercoverage. Survey estimates may be biased if certain subgroups of the population are not given the opportunity to be sampled for the survey. For example, a survey that relies on telephones to conduct interviews would by design exclude households without telephones, and for the purposes of estimating to the general population, would exhibit a potential undercoverage bias relative to nontelephone households. Because the NSAF has a primary focus on low-income families, of which a disproportionate percentage are nontelephone households, the sample is designed to include nontelephone households.

NSAF estimates, like those from all major household surveys, do suffer from undercoverage when compared to Census Bureau estimates. However, the effects appear to be consistent with those seen in other large-scale household surveys. To assess the rate of coverage, estimates for different population groups in the NSAF are compared to those of the Census Bureau. Ideally, such coverage ratios should be close to 100 percent. In 2002, before adjustment, the coverage ratio of children is 97 percent, somewhat higher than the comparable 1999 coverage ratio of 90 percent and 1997 coverage ratio of 94 percent. For adults, the coverage ratio is 90 percent, an increase from 87 percent in 1999 and 86 percent in 1997.

As mentioned earlier, adjustments made to Census Bureau population control totals in the weighting process mitigates the potential bias resulting from undercoverage. For more information on undercoverage concerns in the NSAF, see Report Nos. 3 and 14 in the 1997 NSAF Methodology Series and Report No. 3 in the 1999 and 2002 NSAF Methodology Series.

Nonresponse. Unit nonresponse occurs when sampled units (such as households or families) do not respond to a survey. As a practical matter, unit nonresponse increases the cost of obtaining the same number of completed interviews. From a statistical standpoint,

unit nonresponse raises concerns about possible bias when respondents and nonrespondents systematically differ from one another on key survey measures.

It is important to note that a low response rate is not itself a direct indicator of the actual magnitude of nonresponse bias. Estimates will not suffer from nonresponse bias when there are no differences between respondents and nonrespondents. Typically, survey methodologists focus on response rates as proxy indicators of the presence of nonresponse bias because there is usually little information available on nonrespondents that can be used to judge the magnitude of differences between respondents and nonrespondents.

The response rates achieved in the NSAF are presented in table 3. Since more than one person is sampled per household, several response rates are reported: child, adult, and household. The household response rate, however, is the calculation typically reported for household surveys. The 1997 NSAF screener response rate is 77.4 percent, the 1999 screener response rate is 76.7 percent, and the 2002 screener response rate is 66.0 percent. The 1999 response rate was lower than the 1997 response rate in part due to the overlap between rounds.

Table 3. NSAF Response Rates (percent)

	1997	1999	2002
Screener	77.4	76.7	66.0
Extended			
Children	84.1	81.4	83.5
Adults	79.9	77.5	78.7
Overall			
Children	65.1	62.4	55.1
Adults	61.8	59.4	51.9
Area sample	85.4	86.1	79.4

Note: Data collection periods (in weeks) are 42 in 1997, 35 in 1999, and 38 in 2002.

The rather large decline for 2002 is not unexpected since RDD response rates have been sharply declining across the board in recent years for many surveys, and the NSAF appears to be no exception. In an effort to reduce unit nonresponse, several incentive procedures were employed during data collection with varying degrees of success (see Cantor et al. 2003). Nonetheless, concerns about potential nonresponse bias led to a number of methodological studies with the objective of understanding the impact of nonresponse on statistics of interest. Four such studies are cited below.

- For the 1997 survey, initially interviewed NSAF households were compared with the characteristics of nonresponding households that were later reached with supplemental efforts. These analyses showed no evidence of large or systematic nonresponse errors in the 1997 NSAF statistics examined (see Report No. 7 in the 1997 NSAF Methodology Series).

- A second effort to assess nonresponse bias matched overlapping households from the 1997 and 1999 samples to assess differences in characteristics associated with households that completed an interview in 1999 but not in 1997 (due to refusal or noncontact) and households that completed an interview in both rounds. The results indicate that for the items examined, NSAF nonresponse was largely ignorable—that is, it led to virtually no bias after adjustment (Black and Safir 2000).

- In a third study, a modeling approach was used to estimate potential nonresponse bias. The study concludes that for nearly two-thirds of the households that refuse an interview at screening, the nonresponse may be ignored after adjustment and does not lead to bias. Based on other evidence, the remaining one-third of the refusal cases probably contribute no or minimal bias as well (Scheuren 2000).

- Finally, a fourth study that analyzed the results of a short nonresponse follow-up survey found that the characteristics of respondents and nonrespondents differed little on survey items of interest (Triplett et al. 2002).

Measurement Error. Measurement error arises due to the imperfect nature of the data collection process in surveys. The interviewer, the respondent, the questionnaire, and the mode of data collection are all potential sources of measurement error.

Interviewer Error. Interviewers can introduce measurement error if, for example, they are inconsistent in their delivery of questions to respondents and in the way they record the answers obtained. For all three NSAF rounds, heavy reliance was placed on extensive interviewer training and monitoring procedures. For example, approximately 10 percent of each interviewer's work was silently monitored for quality control purposes.

Respondent Error. Measurement error may also arise if respondents differ in their abilities and motivation to answer specific survey questions. For example, due to the use of a partial overlap of units sampled in the 1997 NSAF, some of the 1999 NSAF respondents participated in both rounds of the survey. Their previous participation could have introduced a panel effect, causing them to either change the behavior being measured or to provide responses that are in some way different from what they would have been had the respondents not participated previously.

To test for panel effects, the 1999 survey estimates of those respondents who were interviewed in both rounds of the survey were compared with survey estimates of respondents from the nonoverlap sample (Wang, Cantor, and Safir 2000). Basic demographic and household estimates were adjusted to known control totals from the 2000 Census. In general, little evidence of panel conditioning effects was found. Even in those cases where statistically significant differences occur, the magnitude of difference was small enough that estimates for the overall sample for 1999 were largely unaffected.

Questionnaire Error. The NSAF questionnaire is modeled heavily on the instruments of its survey predecessors. This approach has resulted, for the most part, in reliable results comparable to other national surveys on similar topics. Every effort was made to keep question wording unchanged from round to round to improve estimates of change. As noted earlier, however, improvements were made to some questions between the 1997 and 1999 surveys, and again between the 1999 and 2002 surveys. For more information on these issues, see Report No. 12 in the 1997 and 2002 NSAF Methodology Series and Report No. 1 in the 1999 NSAF Methodology Series.

Mode Error. The mode of data collection can lead to measurement error. For example, respondents may be less likely to provide truthful responses to sensitive questions in telephone interviews than in self-administered surveys (see de Leeuw and van der Zouwen 1988, for example). For the NSAF, it is plausible that the use of cellular telephones to conduct interviews with nontelephone households reduced the potential for measurement error due to mode differences between the two parts of the survey.

External Validation

Despite the extensive efforts made in the NSAF (only some of which have been described here), nonsampling errors could not be eliminated entirely. The ability to quantify the effect of these errors within the survey itself was very limited, unlike the case for sampling errors. Consequently, an extensive effort was made to compare the NSAF with other surveys whenever possible. Overall, nonsampling errors in the NSAF are reasonably well controlled for and do not appear to have led to more than minor inconsistencies between the NSAF and other surveys designed to obtain similar information.

To illustrate this, we compare NSAF estimates for the distribution of employment earnings for nonelderly adults (18 to 64 years of age) in all three rounds against the corresponding CPS estimates (table 4).[6] A remarkable degree of closeness exists, given that there are sampling and nonsampling errors in both surveys.

Table 4. Earnings from Employment Last Year, Adults Age 18 to 64 (percent)

	1997 (1996 Earnings)		1999 (1998 Earnings)		2002 (2001 Earnings)	
Earnings	**NSAF**	**CPS**	**NSAF**	**CPS**	**NSAF**	**CPS**
Under $10,000	22.1	23.0	19.7	19.8	19.7	16.9
$10,000–14,999	11.9	11.9	10.4	10.8	10.4	9.2
$15,000–24,999	21.1	21.6	20.9	20.4	20.9	19.4
$25,000–34,999	16.7	16.3	17.0	17.2	17.0	17.1
$35,000–49,999	15.2	13.9	16.2	15.5	16.2	16.3
$50,000–74,999	9.0	8.6	10.6	10.4	10.6	12.4
$75,000+	4.0	4.7	5.3	6.0	5.3	8.8
Total	100.0	100.0	100.0	100.0	100.0	100.0

Sources: Current Population Survey (CPS) and National Survey of America's Families (NSAF) information from Urban Institute tabulations.
Note: Data may not add to totals due to rounding; data do not include subjects reported as earning zero or negative money.

In summary, for these and many other comparisons not shown (family composition, work experience, income, and poverty by key demographic characteristics), NSAF estimates approach CPS estimates—for the most part, well within normal sampling variation.

Concluding Comments

A brief overview of NSAF data collection and estimation procedures has been presented here. We have described both the methods used to minimize errors and those to compensate for errors inherent in data collection. We concluded by discussing the

survey's reliability—both in terms of sampling and nonsampling errors. We are particularly encouraged by the degree of comparability between estimates derived from the CPS and those from the NSAF. For the most part, the calculated sampling margins of error provide a guide to the reliability of NSAF data.

For other information on the NSAF design, data collection procedures, measures of quality, and guidelines on how to best use the data, please refer to the series of methodological reports that are available online.

Notes and Acknowledgments

This summary of NSAF survey methods and data reliability parallels work done earlier by Genevieve Kenney, Fritz Scheuren, and Kevin Wang to accompany the 1997 and 1999 *Snapshots* reports. Mike Brick of Westat contributed importantly to both. The authors also thank Laura Wherry for her helpful review and comments.

Endnotes

1. In rounds 1 and 2, Milwaukee was also designated as a study area in its own right, so for 1997 and 1999 data Wisconsin can be viewed as consisting of two study areas, Milwaukee and the balance of the state. In 2002, Milwaukee is not sampled as a separate study area but is incorporated as part of the statewide Wisconsin study area.

2. However, NSAF field procedures did allow for the use of proxy respondents or "facilitators" to interview sampled respondents who could not be interviewed in English or Spanish. There still was a loss of between 1 and 2 percent of the sample due to this barrier. From the interviewer notes, these were often Chinese-, Korean-, or Russian-speaking households.

3. The totals used were based on the 2000 decennial census counts, carried backward to 1997 and 1999 and forward to 2002 by the Census Bureau using birth and death records, plus information on net migration. Other Census Bureau population estimates controlled were the percentage of persons by home ownership (used in the derivation of the weights for children and adults) and education level (used in the derivation of the weights for adults). Occasionally, the variables needed in this weighting were missing. When this occurred, the missing responses had to be imputed, although this was rarely necessary. Race for persons of Hispanic origin is an exception here and had to be imputed quite frequently, since many Hispanic respondents answered the race and ethnicity questions with the designation “Hispanic.” For this reason, while we are comfortable with the race

data overall and with the designation of Hispanic origin, we do not recommend that the race data for Hispanics be used separately.

4. The sampling error introduced because of the imputation of missing responses to specific questions is not estimated by the jackknife method used elsewhere in the survey. See Report No. 4 in the 1997 NSAF Methodology Series for more details on the jackknife. For the most part, though, this understatement should be very small, as is discussed in Report No. 10 in the 1997 NSAF Methodology Series.

5. By design, the overall national estimates were improved between the two rounds. This was done chiefly by substantially increasing the sample taken in the balance of the United States. Thus, nationally, the margin of error for low-income children dropped slightly, even though a drop was not achieved in most of the target states.

6. For details on the extensive external validation efforts undertaken, see Report Nos. 6 and 15 in the 1997 NSAF Methodology Series and Report No. 6 in the 1999 Methodology Series.

References

1997, 1999, and 2002 National Survey of America's Families Basic Snapshot I-III Tables. Washington, DC: The Urban Institute.

1997, 1999, and 2002 National Survey of America's Families Methodology Reports. Washington, DC: The Urban Institute.

Black, Tamara, and Adam Safir. 2000. “Non-Response Bias in the NSAF.” *Proceedings of the Survey Research Methods of the American Statistical Association*, Alexandria, VA.

Cantor, David, Pat Cunningham, Tim Triplett, and Rebecca Steinbach. 2003. “Comparing Incentives at Initial and Refusal Conversion Stages on a Screening Interview or a Random Digit Dial Survey.” *Proceedings of the Survey Research Methods of the American Statistical Association*, Alexandria, VA.

Cantor, David, Kevin Wang, and Natalie Abi-Habib. 2003. “Comparing Promised and Pre-paid Incentives for an Extended Interview on a Random Digit Dial Survey.” *Proceedings of the Survey Research Methods of the American Statistical Association*, Alexandria, VA.

de Leeuw, Edith D., and Johannes van der Zouwen. 1988. “Data Quality in Telephone and Face-to-Face Surveys: A Comparative Meta-Analysis.” *Telephone Survey Methodology*, John Wiley and Sons.

Scheuren, Fritz. 2000. “Quality Assessment of Quality Assessment.” *Proceedings of the Survey Research Methods of the American Statistical Association*, Alexandria, VA.

Triplett, Tim, and Natalie Abi-Habib. 2003. “Determining the Probability of Selection for a Telephone Household in a Random Digit Dial Sample Design Is Becoming More Difficult.” *Proceedings of the Survey Research Methods of the American Statistical Association*, Alexandria, VA.

Triplett, Tim, Adam Safir, Kevin Wang, Rebecca Steinbach, and Simon Pratt. 2002. “Using a Short Follow-up Survey to Compare Respondents and Nonrespondents.” *Proceedings of the Survey Research Methods of the American Statistical Association*, Alexandria, VA.

Wang, Kevin, David Cantor, and Adam Safir. 2000. “Panel Conditioning in a Random Digit Dial Survey.” *Proceedings of the Survey Research Methods of the American Statistical Association*, Alexandria, VA.