

# ASSESSING NONRESPONSE BIAS IN THE NATIONAL SURVEY OF AMERICA'S FAMILIES

Tamara Black and Adam Safir, The Urban Institute  
Adam Safir, UI, 2100 M Street NW, Washington, DC 20037

**Key Words:** Declining Unit Response Rates, Degree of Ignorability, Noncontact and Refusal Differences

## 1. INTRODUCTION

Sample surveys are attractive to social scientists in large part because of their low cost and convenience relative to other measurement options. However, the ability to make statistically reliable statements about population parameters, based on a smaller subset of the target population, is becoming increasingly threatened by the long-term, downward trend in survey participation.

Declining response rates are particularly troubling because of the uncertainty that surrounds their impact. Although response rates are frequently used by the general public as an indicator of survey quality and credibility, perhaps less commonly understood is that large nonresponse rates represent a non-trivial source of bias only to the extent that nonrespondents differ from respondents on characteristics of interest.

Using respondent information obtained in the second round of the National Survey of America's Families (NSAF), this analysis examines the extent and impact of nonresponse bias in the first round of the NSAF. While there is some discussion around the relative size of nonresponse, the primary focus concerns the degree of difference between respondents and nonrespondents. After presenting contextual background information on the NSAF, the discussion will review nonresponse from a general perspective before linking specific issues of interest to the analysis. Finally, findings and conclusions are discussed.

## 2. SURVEY BACKGROUND

The NSAF is a survey of the economic, health, and social characteristics of children, adults under the age of 65, and their families. Two rounds of interviews were conducted in 1997 and 1999, yielding information on over 40,000 households and 100,000 persons in each round. Westat conducted the data collection for the NSAF.

Large representative samples of households were taken in each of 13 targeted states plus the balance of the nation. The 13 states were Alabama, California,

Colorado, Florida, Massachusetts, Michigan, Minnesota, Mississippi, New Jersey, New York, Texas, Washington, and Wisconsin.

These 13 states account for over half of the U.S. population and have a broad array of government programs, fiscal capacity, and child well-being. The sample results from the first round provide a wide range of characteristics for each of the targeted study areas and for the country as a whole, in the period just before the era of the New Federalism (when major changes in U.S. federal and state policies occurred). The sample results from the second round provide information on the characteristics of the targeted study areas and for the country as a whole after many of the changes of New Federalism had been implemented. Collectively, they form a sound baseline from which many of the changes brought about by the New Federalism can be measured, assessed, and tracked.

The NSAF sample is representative of the civilian, noninstitutionalized population under age 65. The first round of data was obtained from February to November 1997, and the second round of data were collected from February to October 1999. The NSAF sample had two parts: the main sample consisted of an RDD survey of households with telephones. This was supplemented with a second (area probability) sample of households without telephones. Telephone households were sub-sampled, with the subsampling rates depending on the presence of children in the household and their response to a single household income-screening question. All households screened with children and classified as low-income were given a full interview, while higher-income households with children and all households without children (but with someone under 65) were sub-sampled before in-depth interviewing. Households with only adults age 65 and over were screened out of the survey.

## 3. NONRESPONSE

### 3.1 Defining Nonresponse

Nonresponse is one source of survey error that almost all sponsors pay particular attention to, and as such, the rate is an important quality indicator that is used to judge not only the specific survey but also the survey organization itself (Lyberg and Dean, 1992). As noted

by Cochran (1977), any sizable percentage of nonresponse makes the results open to question by anyone who cares to do so.

Nonresponse is defined as the failure to obtain complete measurements on the survey sample (Groves, 1989). There are two types of nonresponse;

Unit nonresponse results from the inability to obtain information from some elements of the population that were selected and designated for the sample (Churchill, 1999).

Item nonresponse refers to the failure to measure specific items of interest from an otherwise complete survey element.

Response rates refer to the ratio of the number of units interviewed to the eligible number of sampled units, weighted to represent the appropriate proportion of the population (Brick, et al, 1999).

Nonresponse rates for RDD surveys are driven by the inability to contact persons associated with sampled numbers within the time frame of the survey, the refusal by a sampled person to participate in the survey, or the inability of a sampled person to provide survey information due to a physical, mental, or language barrier. For example, the NSAF is conducted in English and Spanish only.

In recent years, more pronounced refusal rates have been fueled by factors related to the increased use of telephone interviewing as a polling, market research, and telemarketing tool. In addition, recent low unemployment rates have led to a highly competitive hiring environment that has made it more difficult for survey organizations to recruit and retain quality telephone interviewers. Finally, noncontact rates have also risen, in large measure due to the increased availability and use of mobile phones, answering machines, and incoming-call filtration systems such as Caller ID and Call Blocking (Brick, et al, 2000).

### 3.2 Nonresponse Bias

Nonresponse introduces error into survey estimates to the extent that nonrespondents differ from respondents on dimensions measured by the survey. For many descriptive statistics (e.g., the sample mean estimate of the population mean), the nature of nonresponse error in a simple random sample can be expressed as:

$$\bar{y}_r = \bar{y}_n + \frac{m}{n}(\bar{y}_r - \bar{y}_m),$$

where the  $\bar{y}_r$  reflects the values on the respondents, the  $\bar{y}_n$  the values on the entire sample,  $m$  is the number of nonrespondent cases,  $n$  the total sample size, and  $\bar{y}_m$  reflects the values on the nonrespondent cases. In this context, reducing nonresponse can have a dampening effect on both bias and variance.

### 3.3 Early NSAF Nonresponse Studies

In spite of extraordinary efforts that were conducted to attain high response rates in the first round of the survey, the overall response rate in the 1997 NSAF, comparable to other surveys, was about 70 percent (Scheuren and Wang, 1999). This is consistent with the experiences of other large telephone surveys. The overall response rate is the product of the component rates. In the case of adults, the overall response rate for estimates of adults is the product of the extended adult response rate and the screener adult response rate.

To assess the possibility that nonresponse in the NSAF contributed to bias on the survey statistics, a variety of studies have been conducted to gain insight into the characteristics of nonrespondents. The findings showed that for a set of key statistics computed on the population of telephone households with children, there was little evidence of important nonresponse errors (Groves and Wissoker, 1997).

## 4. METHODS

### 4.1 Analysis

The objective of this analysis was to assess the presence of bias due to unit nonresponse in the first round of the NSAF. The analysis used the results from both the first and second round of the NSAF. Specifically, second round respondents were grouped according to first round disposition (e.g., complete, refusal, or noncontact) and tested for differences in demographic, socioeconomic, health coverage, and program participation characteristics.

The following two hypotheses were tested:

1. There is no difference between the characteristics of completes and refusals.
2. There is no difference between the characteristics of completes and noncontacts.

Should the null hypothesis that completes are not significantly different from either refusals or noncontacts be rejected, one may surmise that nonresponse bias due to that component of nonresponse is non-ignorable. On the other hand, should either hypothesis fail to be rejected, one may conclude that

nonresponse bias due to that component of nonresponse is of a trivial nature and may be ignored.

The analysis took advantage of an overlapping sample design feature of the NSAF. That is, a portion of the starting sample of telephone numbers used in the second round came from a pool of numbers used in the first round. This feature was included in the sample design in order to reduce the variances of the estimates of change.

#### 4.2 Comparison Groups

The control group, called **Completes**, was composed of sampled adults 18-64 in households in which a completed interview was obtained in both rounds (see Table 1). The test groups were composed of adults in households for which an interview was not obtained in the first round, but for whom a completed interview was obtained in the second round. The first test group, called **Refusals**, was composed of adults for whom first round nonresponse was due to a household refusal. The second test group, called **Noncontacts**, was composed of adults for whom first round nonresponse was due to noncontact (no answers, non-residential, and non-working numbers).

It is important to note that while the terms completes, refusals, and noncontacts are being employed to describe the control and test groups, the terms refer ONLY to the disposition of the household in the first round, and that for ALL units, completed interviews were obtained in the second round.

Table 1. Analysis Groups

Group	Round 1 Disposition	Round 2 Disposition
Completes	Complete	Complete
Refusals	Refusal	Complete
Noncontacts	Noncontact	Complete

However, there is the risk that using telephone number as the primary sampling unit may allow for persons associated with a newly assigned telephone number to be included in the analysis (i.e., a working telephone number in the first round may have been reassigned to a completely new household in the second round), thus effectively undermining the unambiguous link between telephone number and sampled persons across rounds.

To control for this possible "noise," only those adults in households reporting they had the same telephone number as two years ago were eligible for the analysis.

As can be seen in the Table 2, completes and refusals showed substantial telephone stability across rounds.

Table 2. Results of Controlling for Same Telephone Number

RD 1 Disposition	Total RD 2 Adults	RD 2 Adults with Same Phone Num	
Complete	17,965	16,989	94.6%
Refusal	2,646	2,365	89.4%
Noncontact	7,614	3,920	51.5%
Total	28,225	23,274	82.5%

However, as expected, a substantial number (almost half) of the noncontacts from the first round who were contacted in the second round were eliminated from the analysis because the residents reported that the telephone number where they were reached was not the same telephone number they had two years ago.

#### 4.3 Limitations

This analysis acknowledges some important limitations with regard to the findings. Clearly, the results cannot be used to make statements about persons associated with households which either refused or were not contacted in both rounds. The results of the analysis are only valid for those first round refusals and noncontacts who were contacted and interviewed in the second round.

Further, the methodology uses information obtained from respondents in 1999 to make statements about the characteristics of those same respondents in 1997. The number of changes that may have occurred over a two-year span, particularly in a period of strong economic growth, may confound the analysis. A substantial change in circumstance or socioeconomic make-up (e.g., in employment, income, or some other set of characteristics) may have caused nonrespondents to look more similar to completes in the second round than they did in the first round. Thus, if anything, this would tend to dampen any differences between completes, refusals, and noncontacts.

A related area of concern is that this same change in circumstance might also lead to a greater likelihood of contact and cooperation in the second round. While the NSAF cannot directly measure a change in circumstance for nonresponding first round households, some analysis on this was conducted for first round completed households. Finally, there is the problem that household composition is not static. As dynamic arrangements, households gain and lose members over time. To the extent that members who have joined a

particular household after the first round differ from existing household members, the overall demographic and socioeconomic make-up of the household may have changed accordingly.

To better understand the phenomenon of dynamic household membership, sampled adults in completed second round households were matched to sampled adults from the same household in the first round. Adults who matched by name, age, and sex between rounds were labeled as **Matchers**, and all unmatchable, presumably new household members were labeled as **Joiners**.

Clearly, this procedure could only be done for completes, as the necessary information was not available for first round refusals and noncontacts. However, if the assumption is made that the nature of the difference between matchers and joiners remains stable for responding and nonresponding households, an understanding of the difference between matchers and joiners for completes would allow for a corresponding adjustment in the interpretation of effects among refusals and noncontacts.

As expected, joiners were found to be proportionately more likely to have an other relative (e.g., in-law, grandparent) or nonrelative (e.g., boarder) relationship to the householder than matchers (see Table 3). On characteristics of interest, further analysis showed that joiners were more likely to be younger, renters, poor, uninsured, not married, and not U.S. citizens.

Based on these findings, the analysis moved forward with the assumption that while the degree of difference between matchers and joiners across analysis groups may vary, the shape and direction of this difference will remain constant.

Table 3. Distribution of Matchers and Joiners

Relationship to Householder	Matchers %	Joiners %
Householder	48.9	32.2
Spouse	32.8	19.1
Child (18+)	14.4	20.7
Other relative	2.2	11.3
Nonrelative	1.7	16.7

## 5. FINDINGS

### 5.1 General Findings

Using WesVar to account for the complex survey design, chi-square tests of significance were conducted

at the 0.05 level to compare completes with refusals and noncontacts on a broad array of variables measuring demographic, socioeconomic, health coverage, and program participation characteristics. Specifically, the groups were compared on the following dimensions: gender, presence of children, age, race/ethnicity, household size, education level, U.S. citizenship, marital status, household status (own/rent), employment status, family income as a percent of the federal poverty level (FPL), Aid to Families with Dependent Children (AFDC) receipt, telephone interruption, insurance status, and food concerns (see Table 4).

The findings showed completes to be significantly different from both refusals and noncontacts with respect to age and race/ethnicity. Complete households were also larger, more likely to include minor children, and less likely to have experienced food concerns than other types of households.

### 5.2 Completes vs. Refusals

Overall, completes and refusals were very similar. The two groups were identical with respect to marital status, and although completes were slightly more likely than refusals to have a Bachelor's degree, to be female, employed, above 200% poverty, without health insurance, and own their homes, and slightly less likely to be U.S. citizens, none of these differences were significant.

### 5.3 Completes vs. Noncontacts

There were several significant differences between completes and noncontacts. In addition to differences in mobility and citizenship, as compared to completes, noncontacts were more likely to have experienced phone interruptions, to live in 1- or 2-person households, and to be renters. They were also less likely to be U.S. citizens.

The two groups also differed with respect to health coverage and participation in social welfare programs. Compared to completes, noncontacts were more likely to be without health insurance and to be receiving AFDC.

Finally, completes and noncontacts were dissimilar in terms of educational attainment, employment and poverty. Of the two groups, noncontacts were less likely than completes to have earned a bachelor's degree or to be currently employed. Noncontacts were also more likely to have family incomes at or below 200% of the federal poverty level.

## 6. DISCUSSION

The evidence did not suggest that refusals differed meaningfully from completes. Rather, relatively few statistically detectable differences were found between the two groups. This result is consistent with the conjecture in Scheuren (2000) that a sizable percentage of NSAF refusals are an ignorable form of missingness. With respect to the second research hypothesis, many significant differences were uncovered between completes and noncontacts on the variables of interest. Therefore, while nonresponse bias due to refusals may be ignorable, concern persists about bias resulting from noncontacts.

However, once the relatively small volume of eligible noncontact cases is taken into consideration (removing out-of-scope dispositions normally excluded by response rate calculations, such as non-residential and nonworking numbers), the resulting bias takes on a more trivial nature.

## 7. REFERENCES

Brick, J. M., I. Flores-Cervantes, and D. Cantor. (1999). *1997 NSAF Response Rates and Methods Evaluation*. In the National Survey of America's Families (NSAF) Methodology Series, No. 8

Brick, J.M., J. Montaquila, and F. Scheuren. (2000). "Estimating Residency Rates For Undetermined

Numbers in RDD," Paper presented at American Association for Public Opinion Research Conference.

Churchill, G.A. (1999). *Marketing Research: Methodological Foundations*. Fort Worth: Harcourt Brace.

Cochran, W.G. (1977). *Sampling Techniques, Third Edition*. 1977 NY: Wiley.

Groves, R. (1989). *Survey Errors and Survey Costs*. NY: Wiley.

Groves, R., and Wissoker, D. (1999). *1997 Early Response Studies*. In the National Survey of America's Families (NSAF) Methodology Series, No. 7

Lyberg, L., and Dean, P. (1992), "Methods for Reducing Nonresponse Rates: A Review," Paper presented at American Association for Public Opinion Research Conference.

Scheuren, F., and Wang, K. (1999). *The National Survey of America's Families: A Mid-Term Status Report*. In the National Survey of America's Families (NSAF) Methodology Series, No. 16

Scheuren, F. (2000), "Quality Assessment of Quality Assessments: New Tools for a New Focus?" Paper presented at American Statistical Association Joint Statistical Meetings.

**Table 4. Comparison of Completes vs. Refusals and Noncontacts**

	Completes		Refusals			Noncontacts		
	%	N	%	N	% Diff	%	N	% Diff
<b>Gender</b>								
Males	48.4	7,788	49.4	1,112	2.2	48.4	1,810	0.0
Females	<u>51.6</u>	<u>9,201</u>	<u>50.6</u>	<u>1,253</u>	-2.1	<u>51.7</u>	<u>2,110</u>	0.0
Total (Chi-sq p-value)	100	16,989	100	2,365	(0.4675)	100	3,920	(0.9889)
<b>Presence of Children</b>								
Yes	54.9	14,113	43.2	1,691	-21.3	48.2	2,914	-12.1
No	<u>45.1</u>	<u>2,876</u>	<u>56.8</u>	<u>674</u>	25.9	<u>51.8</u>	<u>1,006</u>	14.7
Total (Chi-sq p-value)	100	16,989	100	2,365	(0.0000)	100	3,920	(0.0133)
<b>Age</b>								
18-24	14.1	1,294	12.9	186	-8.0	18.2	518	29.1
25-34	18.8	3,382	15.3	397	-18.5	26.6	1,169	41.2
35-44	28.7	6,758	28.6	860	-0.3	27.0	1,226	-6.0
45-54	24.9	4,115	22.9	613	-8.2	18.0	733	-27.7
55-64	<u>13.5</u>	<u>1,440</u>	<u>20.3</u>	<u>309</u>	49.7	<u>10.3</u>	<u>274</u>	-23.9
Total (Chi-sq p-value)	100	16,989	100	2,365	(0.0098)	100	3,920	(0.0000)
<b>Race/Ethnicity</b>								
Black Non-Hispanic	10.7	1,447	9.8	157	-7.8	14.1	563	32.3
Hispanic	9.5	1,505	5.8	124	-39.5	15.6	593	64.1
Other Non-Hispanic	<u>79.8</u>	<u>14,037</u>	<u>84.4</u>	<u>2,084</u>	5.8	<u>70.2</u>	<u>2,764</u>	-12.0
Total (Chi-sq p-value)	100	16,989	100	2,365	(0.0283)	100	3,920	(0.0000)
<b>Household Size</b>								
1	5.5	502	6.8	122	22.9	11.8	279	114.5
2	21.9	1,851	37.5	460	71.2	30.6	638	39.8
3	25.4	3,807	18.5	546	-27.2	22.4	1,029	-11.7
4+	<u>47.2</u>	<u>10,829</u>	<u>37.3</u>	<u>1,237</u>	-21.1	<u>35.2</u>	<u>1,974</u>	-25.5
Total (Chi-sq p-value)	100	16,989	100	2,365	(0.0000)	100	3,920	(0.0000)
<b>Education Level</b>								
No HS Diploma or GED	11.6	1,777	8.2	175	-29.4	16.1	567	39.7
HS Diploma or GED	63.7	10,374	67.3	1,475	5.7	61.0	2,317	-4.3
Bachelor's Degree +	<u>24.7</u>	<u>4,838</u>	<u>24.5</u>	<u>715</u>	-0.9	<u>22.9</u>	<u>1,036</u>	-7.3
Total (Chi-sq p-value)	100	16,989	100	2,365	(0.0507)	100	3,920	(0.0013)
<b>US Citizenship</b>								
Yes	95.1	16,064	95.5	2,285	0.5	86.9	3,459	-8.6
No	4.9	827	4.5	72	-8.8	13.1	450	166.8
Total (Chi-sq p-value)	100	16,961	100	2,357	(0.5972)	100	3,909	(0.0000)
<b>Marital Status</b>								
Married w/Spouse in HH	64.0	12,703	64.0	1,724	0.1	53.5	2,424	-16.3
Other	<u>36.1</u>	<u>4,285</u>	<u>36.0</u>	<u>641</u>	-0.2	<u>46.5</u>	<u>1,495</u>	29.0
Total (Chi-sq p-value)	100	16,988	100	2,365	(0.9721)	100	3,919	(0.0001)
<b>Household Status</b>								
Own	81.2	13,746	80.7	1,968	-0.6	52.4	2,255	-35.5
Rent	<u>18.8</u>	<u>3,243</u>	<u>19.3</u>	<u>397</u>	2.6	<u>47.6</u>	<u>1,665</u>	153.4
Total (Chi-sq p-value)	100	16,989	100	2,365	(0.8345)	100	3,920	(0.0000)
<b>Currently Employed</b>								
Yes	78.0	13,420	75.9	1,857	-2.7	73.9	2,941	-5.3
No	<u>22.0</u>	<u>3,569</u>	<u>24.1</u>	<u>508</u>	9.6	<u>26.2</u>	<u>979</u>	19.0
Total (Chi-sq p-value)	100	16,989	100	2,365	(0.2668)	100	3,920	(0.0014)
<b>Total CPS Family Income</b>								
Less or Equal to 200% FPL	20.6	4,793	21.7	444	5.5	34.5	1,411	67.6
Greater than 200% FPL	79.4	12,196	78.3	1,921	-1.4	65.6	2,509	-17.5
Total (Chi-sq p-value)	100	16,989	100	2,365	(0.6263)	100	3,920	(0.0000)
<b>AFDC Receptiency</b>								
Yes	1.3	314	1.3	35	0.0	2.3	119	80.0
No	<u>98.7</u>	<u>16,675</u>	<u>98.7</u>	<u>2,330</u>	0.0	<u>97.7</u>	<u>3,801</u>	-1.1
Total (Chi-sq p-value)	100	16,989	100	2,365	(0.9978)	100	3,920	(0.0093)
<b>Telephone Interruption</b>								
Yes	2.7	647	2.2	69	-18.1	6.7	319	152.5
No	<u>97.4</u>	<u>16,339</u>	<u>97.8</u>	<u>2,295</u>	0.5	<u>93.3</u>	<u>3,594</u>	-4.1
Total (Chi-sq p-value)	100	16,986	100	2,364	(0.4777)	100	3,913	(0.0000)
<b>Currently Uninsured</b>								
Yes	12.4	2,036	10.5	198	-15.4	20.5	731	64.8
No	<u>87.6</u>	<u>14,953</u>	<u>89.5</u>	<u>2,167</u>	2.2	<u>79.5</u>	<u>3,189</u>	-9.2
Total (Chi-sq p-value)	100	16,989	100	2,365	(0.2305)	100	3,920	(0.0000)
<b>Food Concerns</b>								
Yes	14.5	3,281	11.0	333	-23.8	26.5	1,030	82.7
No	<u>85.5</u>	<u>13,703</u>	<u>89.0</u>	<u>2,023</u>	4.0	<u>73.5</u>	<u>2,888</u>	-14.0
Total (Chi-sq p-value)	100	16,984	100	2,356	(0.0221)	100	3,918	(0.0000)