Effects On Survey Estimates From Reducing Nonresponse

Adam Safir, Rebecca Steinbach, Timothy Triplett, and Kevin Wang
The Urban Institute

Key Words: Nonresponse Bias, Contactibility, Cooperation

1. Introduction

Using a variety of procedures designed to maximize response rates, survey organizations expend sometimes extraordinary efforts to minimize the potential for nonresponse bias. Given that nonresponse bias is a function of both the nonresponse rate and the difference between respondents and nonrespondents, maximizing response rates is a sensible approach to minimizing the potential for bias contributed by the first part of the equation.

The second part of the equation, the difference between respondents and nonrespondents, however, is by definition a more complicated component to address. A number of approaches have been suggested for measuring the size of this difference. One such method is to use difficult-to-interview respondents, obtained through increased call attempts, higher incentives, or an extended field period, as proxies for nonrespondents.

When the additional interviews obtained as a result of these efforts display characteristics similar to interviews already conducted with the "easier-to-interview" respondents, researchers may arrive at one of two conclusions: (1) the difficult-to-interview, or nonrespondents-by-proxy, do not differ in meaningful or systematic ways from other respondents, thus implying ignorable nonresponse, or (2) a core group of nonrespondents remain unmeasured, thus suggesting the potential for nonignorable nonresponse bias. Faced with either of these prospects, researchers may question the extent to which additional interviewing efforts are merited, given the absence of identifiable nonignorable nonresponse bias. For example, where there is little indication of a bias reduction resulting from extended efforts to obtain additional interviews, the survey organization may consider a redesign of expensive refusal reworking procedures (Scheuren, 2000).

This paper presents the results of research conducted to analyze the effects of efforts to minimize the potential for nonresponse bias in the 1999 round of the National Survey of America's Families (NSAF). In particular, questions about the efficacy of maximizing response rates on minimizing nonresponse bias motivated this research.

In the first major analysis component — level of effort — we compare the characteristics of the easy-to-interview to the characteristics of the difficult-to-interview, as defined by number of calls to contact and number of refusals. As the literature has suggested that the characteristics of noncontacts and refusals may differ substantively from each other as well as from the "average" respondent, particular emphasis was given to examining differences between the aggregate easy-to-interview group and each subset of the difficult-to-interview, the difficult-to-contact and the reluctant-to-participate. In addition to comparing these groups within the 1999 survey round, we also compare measures associated with varying levels contactibility and cooperation across survey rounds.

The second major analysis component — potential for nonresponse bias — focuses on assessing the potential for nonresponse bias due to non-measured sample elements, treating difficult-to-interview observations as informative of the non-interviewed. Within this analysis step, we also report on the results of a comparison of sampling frame data across easy-to-interview, difficult-to-interview, and non-interviewed households, defined by completion status in NSAF and a short follow-up survey. The data were compared across easy-to-interview, difficult-to-interview and non-interviewed households to assess the appropriateness of using difficult-to-interview respondents as proxies for the non-interviewed.

2. Data Sources

This research uses data from the 1999 and 1997 rounds of NSAF data collection, as well as data from a nonresponse follow-up survey to the 1999 NSAF survey. NSAF is a survey of the economic, health, and social characteristics of children, adults under the age of 65, and their families. The survey has a dual-frame design (random-digit-dial of telephone households and area sample of non-telephone households), features an oversample of low-income households with children, and is representative of the nation and of 13 states. The questionnaire consists of a short screening interview, used to determine household eligibility, and a longer extended interview, which gathers detailed information on the characteristics of sampled household members. The interview can lasts 30 to 45 minutes, and is conducted with the most knowledgeable adult (MKA) of

the sampled child/ren as well as a randomly sampled childless adult in a subset of households. NSAF uses standard survey methods to reduce nonresponse, such as multiple contact attempts and refusal conversion, as well as more extensive efforts, including monetary incentives and an extended field period. Westat conducted both rounds of data collection for NSAF.

The Nonresponse Follow-up Survey (NFS) consisted of a brief survey of 2,000 randomly selected NSAF respondents and nonrespondents. The questionnaire included selected questions from the NSAF instrument, as well as additional opinion questions about importance of surveys and data collection. The data collection for NFS was conducted by the University of Maryland Survey Research Center (SRC) during the later stages of the 1999 NSAF field period.

3. Prior Research

A number of nonresponse studies were conducted following the first round of NSAF data collection (1997) to learn more about the characteristics of NSAF nonrespondents and to assess the impact of missing data from unit nonresponse on survey estimates. We apply the same basic approach of these earlier nonresponse analyses to the 1999 NSAF data, comparing survey results among respondents by level of effort required to obtain an interview, with the assumption that the results of these comparisons would be informative of the differences between those interviewed and those not interviewed.

In the 1997 nonresponse analyses, it was expected that "any pattern for the socioeconomic indicators would be consistent with two hypotheses about the influences toward participation in the NSAF—that those receiving transfer payments would be at home more often (and thus more easily contacted and perhaps with lower time costs of participation) and, because of the topic of the survey, that those receiving transfer payments would be more interested in providing information to the interviewer. Both of these observations are important because they suggest the possibility of nonignorable nonresponse errors; that is, both for contact and for cooperation, the attribute of key interest is an indirect causal factor for response" (Groves and Wissoker, 1999).

For statistics computed on 1997 NSAF telephone households with children, little evidence of important nonresponse errors was observed. However, there was a small tendency for households with higher socioeconomic status to require more effort to obtain an interview. Additionally, NSAF nonrespondents tended to be black non-Hispanic (Groves and Wissoker, 1999). Overall, no evidence for a serious nonresponse bias arising from a large fraction of refusals was detected.

Other studies have found that reluctant respondents tended to be older, with somewhat lower socio-economic status, while difficult-to-contact respondents tended to be younger and more affluent (Curtin, et al, 2000).

4. Methods

4.1 Level of Effort

The level of effort analysis examined the effect on estimates from reducing nonresponse. As previously noted, this research was motivated by an interest in understanding the gains in nonresponse bias reductions realized from the additional effort expended to obtain interviews with the difficult-to-interview. Any large scale data collection effort with finite resources faces the same need to address the question of whether level of effort should be increased to improve data quality, albeit at a higher cost, or whether it can be reduced to minimize operational costs, without a corresponding risk to data quality.

The level of effort analysis file was restricted to 1999 NSAF children in telephone households. The total sample size was 34,831 sampled children. Comparison groups were formed by classifying the sample into hierarchies of contactibility based on number of calls before first screener contact (1, 2, 3 or 4, and 5+), and cooperation, based on number of refusals before completing the interview (0, 1, and 2+). We then compared the household and parental characteristics of children in difficult-to-contact and reluctant-to-participate households to those of children in the "average" responding household. We included a broad range of measures, intended to replicate the 1997 analyses by Groves and Wissoker, and also to reflect a variety of demographics and other survey items, such as age, race, ethnicity, education, income, employment, health insurance, program participation, family structure, and household tenure.

In addition to examining potential differences in these groups within the 1999 data, we also compare group differences across survey rounds. Using the results from the Groves and Wissoker analysis, we compare the differences between the groups in 1997 to the

Table 1. Estimates by Number of Calls Required for First Contact, RDD Cases with Children

Characteristic	Calls Before First Screener Contact: One	Two	3 or 4	5+	Total
MKA employed	71.66	71.72	72.13	79.56	72.33
s.e.	*0.53*	*1.50*	*1.31*	*1.51*	*0.51*
Spouse employed	87.32	88.03	90.82	89.70	88.07
s.e.	*0.47*	*1.06*	*0.99*	*1.29*	*0.33*
Both MKA and spouse employed	60.87	61.09	63.32	69.92	61.87
s.e.	*0.67*	*1.57*	*1.76*	*2.03*	*0.54*
No high school degree or GED (MKA)	10.70	9.50	7.67	5.59	9.70
s.e.	*0.43*	*0.67*	*0.87*	*0.79*	*0.31*
No high school degree or GED (SP)	11.55	9.67	8.20	6.04	10.39
s.e.	*0.51*	*0.92*	*0.96*	*1.01*	*0.36*
Black, non-Hispanic (MKA)	11.18	10.67	10.85	10.69	11.01
s.e.	*0.44*	*0.73*	*0.92*	*1.13*	*0.34*
Hispanic (MKA)	13.68	13.24	11.39	9.44	12.98
s.e.	*0.41*	*0.80*	*0.94*	*1.01*	*0.34*
Age of MKA	37.57	37.54	36.82	36.84	37.41
s.e.	*0.09*	*0.19*	*0.21*	*0.22*	*0.07*
Spouse present	80.75	79.64	78.45	74.70	79.79
s.e.	*0.50*	*0.90*	*1.05*	*1.27*	*0.38*
MKA has health insurance	84.81	86.85	87.08	89.26	85.81
s.e.	*0.44*	*0.77*	*0.99*	*1.09*	*0.31*

differences between the groups in 1999, looking at the difference of the differences with an interest in assessing any change in degree or direction.

4.2 Potential for Nonresponse Bias

The appropriateness of using the difficult-to-interview as a proxy for the non-interviewed relies on the validity of the assumption that the difficult-to-interview characteristically resemble the non-interviewed. To test this assumption, we use exchange-level sampling frame data to examine differences between households that completed the NSAF (Group AB, n=993), households that completed NFS but not NSAF (Group C, n=231), and households that did not complete either NSAF or NFS (Group D, n=562).

Exhibit 1. Sampling Frame Data Comparison Groups

Interview Status		NFS: Yes	NFS: No
NSAF	Yes	A	B
	No	C	D

The exchange-level data were provided on the Genesys Sampling Systems sample data file. The exchange-level data are limited for 1999 purposes, as they are 1997 FIPS county projections for dominant exchanges. However, as the inputs used to form nonresponse weighting adjustment classes in NSAF, they have a direct relevance to the survey in question. Additionally, the exchange-level data feature a desirable level of geographic specificity. T-tests were used to compare the mean characteristics of Group AB (the easy-to-interview) to Group C (the difficult-to-interview) and Group D (the non-interviewed), as well as to compare the mean characteristics between Groups C and D. Under the assumption that the difficult-to-interview are informative of the non-interviewed, we expected to see small or no differences between Groups C and D, and larger differences between either or both of these two groups and the easy-to-interview group, Group AB. We include "B" in the easy-to-interview group because although interviews were attempted but not obtained in NFS, we acknowledge that the length of the NSAF interview likely had an effect on the decision to participate in the follow-up (for respondents who completed NSAF but not NFS).

Comparison measures included average rent, median income, median home value, percent age 0-17, percent black, percent Hispanic, percent renters, percent listed, percent income 0-10K, percent income 11-15K, and percent income 16-25K. The results of the sampling

Table 2. Estimates by Refusal Status, RDD Cases with Children

Characteristic	Refusal Status None	One	2+	Total
MKA employed	72.34	72.21	72.54	72.33
s.e.	*0.61*	*0.96*	*1.47*	*0.51*
Spouse employed	87.05	90.29	88.29	88.07
s.e.	*0.43*	*0.68*	*1.21*	*0.33*
Both MKA and spouse employed	61.11	62.85	63.73	61.87
s.e.	*0.74*	*1.13*	*1.59*	*0.54*
No high school degree or GED (MKA)	10.07	8.91	9.55	9.70
s.e.	*0.45*	*0.56*	*0.84*	*0.31*
No high school degree or GED (SP)	10.71	9.31	11.30	10.39
s.e.	*0.48*	*0.68*	*1.08*	*0.36*
Black, non-Hispanic (MKA)	10.45	10.51	15.55	11.01
s.e.	*0.40*	*0.63*	*1.41*	*0.34*
Hispanic (MKA)	14.25	11.31	9.65	12.98
s.e.	*0.46*	*0.61*	*0.84*	*0.34*
Age (MKA)	37.12	37.79	38.05	37.41
s.e.	*0.10*	*0.13*	*0.22*	*0.07*
Spouse present	78.92	81.13	81.56	79.79
s.e.	*0.47*	*0.73*	*1.34*	*0.38*
MKA has health insurance	84.98	87.51	86.43	85.81
s.e.	*0.42*	*0.61*	*1.03*	*0.31*

frame data would be used to draw conclusions about the potential for nonresponse bias due to the unobserved sample elements.

5. Results

5.1 Level of Effort

The additional effort expended to interview difficult-to-contact and reluctant-to-participate respondents yielded respondents whose characteristics and circumstances were relatively similar to the more easily interviewed, with some notable differences. Those easily contacted were more likely to be not in the labor force or unemployed, and more likely to receive benefits from the government, while those more difficult-to-contact tended toward higher socioeconomic status and education levels. Larger differences were seen in some measures, particularly on demographics such as race and age.

Differences varied by group. Table 1 compares attributes of children in households easily contacted to the attributes of those in difficult-to-contact households. The difficult-to-contact tend to be less poor, more employed, more educated, less Hispanic, more insured, younger, and less foreign-born. Additionally, children living with employed adults tend to have been contacted with a greater number of calls. Table 2 presents a comparison of children in households which provided the interview without a refusal, those providing it after one refusal, and those providing it after two or more refusals. Reluctant-to-participate households tend to be more likely to include homeowners, be slightly less poor, more Black non-Hispanic, older, and less foreign-born.

It is important to note, however, that overall, the larger demographic differences did not translate into significant differences in other survey items. Few meaningful differences were seen on important outcome measures, such as access to care, food security, or economic hardship.

With respect to changes in the degree and direction of differences over time, table 3 compares the difference between groups in the 1997 data to the difference between groups in the 1999. For example, in 1997, the difference between the percent of MKAs in 2+ refusal households and the average was 4.4 percentage points (75.2 vs. 70.8) or 6.2%. In 1999, this difference shrank to 0.2 percentage points (72.5 vs. 72.3) or 0.3%, resulting in a 4.2 point decrease and 5.9 percentage

Table 3. Degree and Direction of Differences in Survey Estimates by Level of Effort and Survey Round

	Round 1		Round 2		R2 Diff vs. R1 Diff	
5+ Calls-to-Contact vs. Average	**Pt. Diff**	**% Diff**	**Pt. Diff**	**% Diff**	**Pt. Diff**	**% Pt. Diff**
Homeowner lives in household	1.1	1.5	0.0	0.0	-1.1	-1.5
Household income below 200% poverty level	-8.7	-24.6	-7.2	-21.2	-1.5	-3.4
Received TANF in 1998	-2.5	-37.3	0.8	16.8	*-1.7*	*-20.5*
Confident medical care available if needed	2.2	2.4	0.6	0.6	-1.6	-1.8
Ever skip meals because money unavailable	-1.5	-12.5	-1.0	-9.2	-0.5	-3.3
Biological mother lives in household	-1.8	-2.0	-0.1	-0.2	-1.7	-1.8
MKA employed	12.1	17.1	7.2	10.0	-4.9	-7.1
No high school degree or GED (MKA)	-3.5	-34.3	-4.1	-42.4	0.6	8.1
Black, non-Hispanic (MKA)	-0.6	-5.8	-0.3	-2.9	-0.3	-2.9
Hispanic (MKA)	-4.1	-36.3	-3.5	-27.3	-0.6	-9.0
MKA has health insurance	2.9	3.4	3.5	4.0	0.6	0.6
2+ Refusals vs. Average						
Homeowner lives in household	3.0	4.1	4.2	5.8	1.2	1.7
Household income below 200% poverty level	-3.5	-9.9	-2.5	-7.4	-1.0	-2.4
Received TANF in 1998	-0.6	-9.0	-0.4	-8.1	-0.2	-0.9
Confident medical care available if needed	0.3	0.3	-0.2	-0.2	*-0.1*	*-0.1*
Ever skip meals because money unavailable	-1.5	-12.5	-0.7	-6.8	-0.8	-5.7
Biological mother lives in household	-1.3	-1.4	-0.6	-0.7	-0.7	-0.8
MKA employed	4.4	6.2	0.2	0.3	-4.2	-5.9
No high school degree or GED (MKA)	1.6	15.7	-0.1	-1.5	*-1.5*	*-14.1*
Black, non-Hispanic (MKA)	0.2	1.9	4.5	41.2	4.3	39.3
Hispanic (MKA)	0.5	4.4	-3.3	-25.7	*2.8*	*21.2*
MKA has health insurance	2.1	2.4	0.6	0.7	-1.5	-1.7

Figures in italics are significant at the .05 level

point decrease (6.2% vs 0.3%) in the difference in of these groups from 1997 to 1999.

Overall, the comparisons in table 3 show that the degree of difference has decreased between difficult-to-interview and average households. We speculate that the decrease in the degree of difference may be attributable to a number of factors. For example, due to social or telephony changes, the "pool" of those harder-to-interview may be increasing to include those whose characteristics and circumstances are more similar to the easier-to-interview.

Alternatively, it may mean that we are capturing more of the difficult-to-interview households which more closely resemble easier-to-interview cases, and fewer difficult-to-interview households which are more similar to nonrespondents.

5.2 Potential for Nonresponse Bias

At the exchange-level, Group C (difficult-to-interview) respondents tended to live in exchanges with a higher percentage of blacks and renters and a lower percentage of listed telephone numbers than Group AB respondents (easier-to-interview). Alternatively, Group D households (non-interviewed) were shown to live in exchanges with a significantly higher median income, higher average rent, higher percent black, and lower percentage of listed telephone numbers than did Group AB respondents (see table 4).

Overall, Group AB households are more similar to Group D households and less similar to Group C households; however, Group C and D households exhibit smaller between group differences, and both exhibit larger differences as compared to Group AB households. The implication, based on the available data, is that Group C households, and respondents living in such households, may be viewed as reasonable proxies for respondents living in Group D households. However, the differences between the three groups is almost negligible across most measures. Given such small differences, the utility of the exchange-level data may be limited in arriving at actionable conclusions.

6. Conclusions

The results of our research indicate that on average, the

Table 4. Exchange-level Household Estimates by Comparison Group

Exchange Characteristic	Group AB	Group C	Group D	% Diff AB vs. C	% Diff AB vs. D	% Diff C vs. D
Average Monthly Rent (in dollars)	465	481	503	-3.5	*-8.2*	-4.6
Median Income (in dollars)	42,105	41,427	44,464	1.6	*-5.6*	*-7.3*
Percent Black	11.7	14.2	12.4	-21.2	-5.5	12.9
Percent Hispanic	8.0	9.9	9.7	-23.6	*-21.7*	1.5
Percent Listed Telephone Numbers	38.4	35.3	36.2	*8.2*	*5.7*	-2.8
Percent Renters	34.3	37.5	35.8	*-9.2*	-4.1	4.6
Percent Age 0-17	25.7	25.6	25.1	0.6	*2.3*	1.7
Sum of Squared Differences				11.7	6.5	2.8

Figures in italics are significant at the .05 level

characteristics of children in difficult-to-contact and reluctant-to-participate households do not differ in meaningful ways from those of children in average households. Although some differences do exist in the demographic items, these differences do not carry over into important outcome measures. We also note a perceptible decrease in the degree of difference between these groups over time, although we lack sufficient data to draw substantive conclusions from this finding. We speculate that as the pool of more difficult-to-interview households grows "passively" due to the increased availability and use of telephony barriers (e.g., Caller ID, dual voice-computer lines), "new" difficult-to-interview households may exhibit fewer differences as compared to easier-to-interview households. While this may suggest that some increase in absolute nonresponse may not translate to a monotonic increase in bias potential, it does raise the specter of a core group of difficult-to-interview households that have become even more difficult to identify and interview within a now larger difficult-to-interview respondent pool. Additionally, the increase of difficult-to-interview households which now more characteristically resemble easier-to-interview cases may further undermine the assumption that the difficult-to-interview are informative of the non-interviewed, thereby diminishing the utility of the difficult-to-interview as proxies for the non-interviewed.

With respect to using the difficult-to-interview as proxies for the non-interviewed, the analysis of sampling frame data showed that difficult-to-interview and non-interviewed households were more similar at the exchange-level, and each less similar to the easier-to-interview. However, the coarseness of the sampling frame data limit the ability to examine these findings in more detail.

References

"The Effects of Response Rate Changes on the Index of Consumer Sentiment", Richard Curtin, Stanley Presser, and Eleanor Singer; POQ Volume 64 / Winter 2000 / Number 4

"Consequences of Reducing Nonresponse in a National Telephone Survey", Scott Keeter, Carolyn Miller, Andrew Kohut, Robert M. Groves, and Stanley Presser; POQ Volume 64 / Summer 2000 / Number 2

"No. 7: Early Nonresponse Studies of the 1997 National Survey of America's Families", Groves, R., and Wissoker, D., March 1999.

"Quality Assessment of Quality Assessment", Fritz Scheuren, American Statistical Association 2000 Proceedings of the Section on Government Statistics and Section on Social Statistics