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 less than perfect survey participation rates.

The second part of the equation is by definition a more complicated component to address. A number of approaches have been suggested for measuring the size of the difference between respondents and nonrespondents, despite the unobserved status of the latter. 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.

Although it is assumed that additional efforts to obtain interviews with the difficult-to-interview will improve precision and reduce nonresponse bias (Lynn, et al, 2002), when the interviews obtained as a result of these efforts display characteristics similar to interviews already conducted with 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 (1) or apparent ineffectiveness (2). 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, this research was motivated by questions about the efficacy of maximizing response rates on minimizing nonresponse bias.

In the first major analysis component — level of effort — we address the effect that increasing the level of effort expended to increase participation rates has on reducing nonresponse bias, by comparing the characteristics of persons in easy-to-interview households to the characteristics of persons in difficult-to-interview households. These groups are 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 each subset of the difficult-to-interview, the difficult-to-contact (5+ calls to contact) and the reluctant-to-participate (2+ refusals), and the average interviewed household. In addition to comparing these groups within the 1999 survey round, we also compare measures associated with varying levels of 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 these three groups 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, used to gather detailed information on the characteristics of sampled household members. On average, the interview lasts 30 to 45 minutes, and is conducted with the most knowledgeable adult (MKA) of the sampled child/ren and/or 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 second data source consisted of a short follow-up survey (SFS) conducted with a random selection within predefined strata of NSAF respondents and nonrespondents. The starting sample size was 2,000 finalized NSAF telephone numbers, of which 1,788 were determined to be eligible for interview or reinterview. The questionnaire included selected items from the NSAF instrument, as well as additional opinion questions about the importance of surveys and research. The data collection for SFS 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, and compare 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. Level of Effort

4.1 Methods

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 reduction realized from the additional effort expended to obtain interviews with the difficult-to-interview. Most large scale data collection efforts with limited resources face the same need to address 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 included both screener and extended refusals in the total number of refusals; while there may be some differences between the two types, our primary interest was in the presence of a reluctant household member, regardless of the type of reluctance (initial refusal vs. second

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
Foreign-born person lives in household	16.22	13.41	13.38	12.51	15.06
s.e.	*0.47*	*0.79*	*1.09*	*1.15*	*0.39*
Household income below 200% poverty level	35.26	33.66	32.90	26.78	33.99
s.e.	*0.75*	*1.06*	*1.30*	*1.46*	*0.53*
Received food stamps in 1998	10.76	10.11	8.17	9.79	10.23
s.e.	*0.46*	*0.57*	*0.75*	*1.32*	*0.34*
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*
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*
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*

interview refusal).

The analysis 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, using selection weights. The selection weights excluded any adjustments for nonresponse or poststratification. 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. We expected the findings to inform an assessment of the relative utility of additional contact and conversion attempts, although a final determination is limited by the lack of revised weights to compare estimates with and without the cases obtained as a result of additional efforts.

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 differences between the groups in 1999, looking at the difference of the differences to gauge change in degree or direction.

4.2 Results

The additional effort expended to interview difficult-to-contact and reluctant-to-participate respondents yielded respondents whose characteristics and circumstances were relatively similar to those of the more easily interviewed, with some notable differences.

Table 1 compares attributes of children in households easily contacted to the attributes of those in difficult-to-contact households. Parents in the difficult-to-contact households tend to be less poor, less Hispanic, more insured, younger, and less foreign-born. Further, parents of children in easily contacted households were more likely to be unemployed or not in the labor force, and more likely to receive benefits from the government. Parents in the more difficult-to-contact households tended toward higher socioeconomic status and education levels. While intuitive and consistent with prior research, these results do re-emphasize the importance of additional contact attempts, particularly as children living with employed adults tended to have been contacted only with a greater number of calls.

Table 2 presents a comparison of children in households interviewed without a refusal, those interviewed after one refusal, and those interviewed

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

Characteristic	Refusal Status None	One	2+	Total
Foreign-born person lives in household	16.07	13.75	12.40	15.06
s.e.	*0.45*	*0.66*	*0.83*	*0.39*
Homeowner lives in household	70.76	75.64	76.95	72.72
s.e.	*0.59*	*0.75*	*1.25*	*0.36*
Household income below 200% poverty level	35.61	31.19	31.46	33.99
s.e.	*0.65*	*0.93*	*1.37*	*0.53*
Received food stamps in 1998	10.88	8.77	9.99	10.23
s.e.	*0.42*	*0.64*	*1.15*	*0.34*
Confident medical care available if needed	92.99	93.76	93.01	93.2
s.e.	*0.36*	*0.49*	*0.66*	*0.3*
Unable to pay mortgage, rent, or utilities, past year	15.41	14.03	14.06	14.9
s.e.	*0.53*	*0.76*	*1.03*	*0.42*
Ever skip meals because money unavailable	10.77	10.49	9.88	10.6
s.e.	*0.44*	*0.55*	*0.92*	*0.36*
Black, non-Hispanic (MKA)	10.45	10.51	15.55	11.01
s.e.	*0.40*	*0.63*	*1.41*	*0.34*
Age (MKA)	37.12	37.79	38.05	37.41
s.e.	*0.10*	*0.13*	*0.22*	*0.07*

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. Larger, significant differences were seen in some measures, particularly on the demographic items such as race and age. However, it is important to note that these observed, larger demographic/socioeconomic differences did not translate into significant differences in other important outcome items, 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 data. For example, in 1997, the percentage point difference between past year TANF receipt in 2+ refusal households and the average was -0.6 (6.1 vs. 6.7) or -9.0%. In 1999, this difference shrank to -0.4 percentage points (4.2 vs. 4.6) or -8.1%[1], resulting in a 0.2 point decrease and 0.9 percentage point decrease (-9.0% vs. -8.1%) in the difference of these groups from 1997 to 1999.

Overall, the comparisons in table 3 show that the degree of difference between difficult-to-interview and average households has decreased, with one or two exceptions, such as race and ethnicity. It should be noted that although there were some differences between the incentive structure used for refusal conversion in round 1 and the incentive structure used in round 2, it is our belief that the increased use of incentives in the second round would have served to increase the estimated difference between the difficult-to-interview and the average household (due to higher conversion rates among the more reluctant sampled households). In fact, the opposite was observed.

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, leading to a capture of a greater number of difficult-to-interview households which more closely resemble easier-to-interview cases (and are therefore potentially less informative of nonrespondents), and proportionately fewer difficult-to-interview households which are more similar to nonrespondents (and who are arguably more informative of nonrespondents).

5. Potential for Nonresponse Bias

5.1 Methods

The appropriateness of using the difficult-to-interview

[1] As a result of rounding, some differences may appear to be slightly higher or lower than the difference of the reported rates.

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 last year	-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
2+ Refusals vs. Average						
Household income below 200% poverty level	-3.5	-9.9	-2.5	-7.4	-1.0	-2.4
Received TANF last year	-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
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 *

* Indicates significant differences at the .05 level.

as a proxy for the non-interviewed rests 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), households that completed SFS but not NSAF (Group C), and households that did not complete either NSAF or SFS (Group D).

The exchange-level data were provided on the Genesys Sampling Systems sample data file. As projections based on FIPS county projections for dominant exchanges, the exchange-level data have a certain level of coarseness, but are still useful data to analyze, particularly since they are used to form the nonresponse weighting adjustment classes in NSAF. Additionally, the exchange-level data feature a desirable level of geographic specificity.

Exhibit 1. Sampling Frame Data Comparison Groups

Interview Status		**SFS**	
		Yes	No
NSAF	Yes	Group A (n=675)	Group B (n=318)
	No	Group C (n=231)	Group D (n=562)

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 SFS, 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 SFS).

Comparison measures included average rent, median income, median home value, percent age 0-17, percent black non-Hispanic, percent Hispanic, percent renters, percent listed, percent income 0-10K, percent income 11-15K, and percent income 16-25K. The results of the sampling frame data would be used to draw conclusions about the appropriateness of making statements about the potential for nonresponse bias due to unobserved sample elements.

5.2 Results

At the exchange-level, Group C (difficult-to-interview) respondents tended to live in exchanges with a higher percentage of black non-Hispanics 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 non-Hispanic, higher percent Hispanic, 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.

Table 4. Exchange-level Household Estimates by Comparison Group

	Group		
Exchange Characteristic	AB	C	D
Average Monthly Rent (in dollars)	464	476	503
Median Income (in dollars)	42,002	40,919	44,528
Percent Black	11.4	14.8	12.1
Percent Hispanic	8.3	9.6	10.0
Percent Listed Telephone Numbers	38.6	35.9	36.9
Percent Renters	33.7	37.4	34.8
Percent Age 0-17	25.7	25.7	25.2

6. Conclusions

The results of our research indicate that on average, the 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 larger differences were seen the demographic make-up of the groups (for example, in education and employment between contactibility groups, and in race and ethnicity between cooperation groups), these differences were not observed to carry over into important outcome measures such as confidence in medical care and food insecurity. While these results were encouraging, we acknowledge certain analytic limitations, such as the lack of revised selection weights to compare estimates with and without the difficult-to-interview cases.

We also note a perceptible decrease in the degree of difference between these groups over time, although again, 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 potential for bias, 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 their utility 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. While this points positively to the use of the difficult-to-interview as proxies for the non-interviewed, the coarseness of the sampling frame data limit our 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.

"No. 7: Early Nonresponse Studies of the 1997 National Survey of America's Families", Robert Groves and Doug Wissoker, March 1999, The Urban Institute website.

"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.

"The Effects of Extended Interviewer Efforts on Nonresponse Bias", Peter Lynn, Clarke, P., Martin, J. and Sturgis, P. (2002), chapter 9, 135-147, Survey Nonresponse (eds. R.M. Groves, D.A. Dillman, J. Eltinge and R.J.A. Little), John Wiley.

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