Nonresponse Bias from Incomplete Interviews in the National Survey of America's Families

Kevin Wang, Urban Institute
Kevin Wang, Urban Institute, 2100 M Street, NW, Washington DC, 20037

Key Words: Nonresponse bias, partial interviews, respondent burden

In weighting adjustments for unit nonresponse, the weights of respondents are adjusted to compensate for nonrespondents that are similar to respondents on various characteristics, either from the survey itself or from other data sources. While we generally think of noncooperation as refusing to do any part of a survey, some individuals classified as nonrespondents may actually answer significant portions of the survey before they break off the interview. In practice, there may be differences among nonrespondents in the information that is available for weighting adjustments in general. However, the data requirement for nonresponse adjustments is that information used to create comparable groups of respondents and nonrespondents must be available for both groups. Thus, additional information that is available for some nonrespondents but not all nonrespondents is often ignored in weighting adjustments for unit nonresponse.

In this paper, I use data from round 2 of the National Survey of America's Families (NSAF) in order to explore whether break-offs should be treated as a distinct form on nonresponse in post-survey adjustments for nonresponse. Nonresponse bias may arise as a problem to the extent that respondents who only complete part of the survey are different from respondents who complete the survey as well as those who do not complete any part of the survey (except for a short screening interview to determine eligibility for the extended interview). This paper uses information about break-offs to assess the degree to which such biases might exist in estimates produced from NSAF analytic files and whether such biases are reduced through nonresponse weighting adjustments that do not explicitly take into account the characteristics of incomplete interviews. In addition, I examine some of the characteristics of break-offs to gain further understanding of why a significant portion of respondents do not complete the NSAF interview.

Nonresponse Adjustments for Break-offs

There is little guidance on what constitutes the difference between a partial complete and a break-off. AAPOR guidelines on response rates note that there are three widely used criteria for distinguishing a partial complete from a break-off: (1) the proportion of all applicable items answered, (2) the proportion of key items answered and (3) the proportion of all applicable items asked (American Association for Public Opinion Research, 2000). Ultimately, the differences between a complete, a partial interview and a break-off will depend on the purpose of the survey as well as the goal of the analysis.

Break-offs are an example of what Brick and Kalton (1996) refer to as partial nonresponse, a type of nonresponse that falls somewhere in between unit nonresponse and item nonresponse. In cases of partial nonresponse, we have more data available than in archetypical cases of unit nonresponse but much more missing data on individual items than usually implied by item nonresponse.

Partial nonresponse poses special problems for standard methods designed to adjust for unit and item nonresponse. One possibility is to treat partial nonresponse as unit nonresponse and employ weighting adjustments in which partial nonrespondents are treated as if they were unit nonrespondents. But this necessitates throwing away data gathered from the respondent. Another possibility is to treat partial nonresponse as a problem of item nonresponse and impute for the missing data. But an imputation approach may be impractical to implement when many items are not answered.

Instead, break-offs may be more usefully viewed as a distinct form of noncooperation. Noncooperation on surveys is usually thought of as a more or less discrete choice by the respondent on whether to participate in a survey. The phenomena of initial cooperators who break-off the survey suggests that while there is a crucial decision on whether or not to participate in a survey, there are also a series of decisions made after the survey has begun to continue participation.

In some surveys, no distinctions are made between noncontacts and refusals during the creation of nonresponse adjustments. If different processes generate noncontacts and refusals, adjustment procedures which ignore these differences may not

reduce nonresponse bias. Groves and Couper (1998) suggest that separate adjustments should be carried out to reflect differences between noncontacted cases and refusals. A process that generates noncontact may bear little resemblance to a process that generates noncooperation, conditional upon contact.

We can carry this idea a step further by hypothesizing that the process generating break-offs of interviews may be different from the initial decision to participate in a survey. If respondents who break off the interview are very different from those who complete the survey or those who refuse to cooperate at all, weighting adjustments which ignore this may not compensate for nonresponse bias. In most surveys, break-offs are not adjusted for separately. Instead, if a nonresponse adjustment models noncontacts as distinct from refusals, break-offs are treated as refusals in such adjustments.

Obviously, break-offs may not make up a large proportion of nonresponse and it may not be worth it to carry out a separate nonresponse adjustment for break-offs. In most surveys, particularly for telephone surveys, refusals often take place in the introduction. In panel surveys, nonresponse due to panel attrition between panels is probably a greater source of nonresponse in subsequent waves than break-offs during the course of the interview. But, if break-offs are in sufficient numbers relative to initial refusals, it may make sense to conduct a separate adjustment for such cases.

One situation in which we might expect differences between break-offs and completes is when respondents face differential burdens of answering survey questions. One facet of respondent burden is the length of the interview. If a questionnaire is designed in such a way that certain types of respondents are asked more questions than others, those receiving more questions may be more likely to break off the interview.

Another factor which may lead to differential rates of breaking off is different reactions of respondents to the questions being asked as the survey progresses. For example, in many public opinion surveys, demographic items such as race, ethnicity and income are asked at the end of the survey. Sensitive items such as these are asked towards the end of the survey since such questions may lead respondents to break off the survey. If a survey is divided up into modules or sections of questions with different subject matter in each section, respondents may break off the survey differentially, depending upon the subject matter.

The NSAF Design and Estimation

The NSAF, a survey of the well-being of children, adults under age 65, and their families, is a component of Assessing the New Federalism (ANF), an Urban Institute project designed to examine the impact of recent shifts (devolution) of much of the responsibility from the federal level to state and local governments for programs designed to assist low-income families. [1] Low-income families (defined as a family income below 200 percent of the federal poverty level) are of particular interest because policy changes are expected to affect these families the most. The first round of the survey was conducted in 1997, the second round in 1999 and a third is planned for 2002. Westat conducted the survey for rounds 1 and 2 and will conduct the survey for round 3 as well.

The NSAF has large probability samples in each of 13 focal states, as well as a sample for the balance of the country, and it produces reliable state and national estimates of child and adult characteristics. The NSAF uses a dual-frame sample design. A random digit dialing (RDD) frame of telephone households is supplemented with an area probability sample of nontelephone households.

The NSAF interview consists of a short, 5 minute, screener interview followed by an extended interview. In the screener, questions are asked to determine the eligibility of the household. The NSAF includes only households with at least one person under 65. In addition, the NSAF oversamples households with children and families with incomes below 200 percent of the poverty threshold.

The NSAF extended interview ranges from 30 to 45 minutes, depending on whether the questions are intended to ask about a single adult or about children in the household. Interviews about adults without children (referred to as childless adult) are conducted with a randomly selected adult over 18 years old and last about 30 minutes. Interviews about children and

[1] The Assessing the New Federalism Project is funded by a consortium of private foundations including the Annie E. Casey Foundation, the Robert Wood Johnson Foundation, the Henry J. Kaiser Family Foundation, the Ford Foundation, the David and Lucile Packard Foundation, the John D. and Catherine T. MacArthur Foundation, the Charles Stewart Mott Foundation, and others.

families are conducted with the person who knows the most about the health and well-being of the children (Most Knowledgeable Adult or MKA). MKA interviews last about 45 minutes. Sections of the NSAF extended interview are as follows:

Section	Content
B	Health Status and Satisfaction
C	Child Education
D	Household Roster
E	Health Care Coverage
F	Health Care Use and Access
G	Child Care
H	Nonresidential Parents
I	Employment and Earnings
J	Family Income
K	Welfare Program Participation
L	Adult Education and Training
M	Housing and Hardship
N	Issues, Problems, Social Services
O	Race, Ethnicity, Nativity
P	Closing

Cases that complete the entire survey are classified as CO (complete). Cases that break off the survey after completing section K (Welfare Program Participation) are coded as partial completes. Partial completes (CP cases) along with the completes constitute cases in NSAF analytic files for both rounds 1 and 2. Cases that complete the household roster (section D) but which break off prior to completing section K are coded as CD cases, which I refer to in this paper as break-offs.

Table 1 shows the distribution of extended interview final dispositions in round 2 of the NSAF for the RDD sample. Since the NSAF extended interview follows a screener interview, noncontact is not a major source of nonresponse. Instead, most of the nonresponse on the extended interview is from noncooperation. Overall, about 45 percent of the extended interview nonresponse is from incomplete interviews.

Table 1: 1999 NSAF Extended Interview Final Dispositions – Child Unit of Analysis

Final Result	Frequency	% Overall
CO (complete)	34,253	80.3
CP (partial)	537	1.3
CD (break-offs)	3,495	8.2
Other refusals	3,108	7.3
Other non-response	1,278	3.0

Table 2 shows the distribution of cases that broke off after completing the household roster section of the survey as well as the percentage of cases that broke off the survey for each section, given that the interview had not already been broken off. In the NSAF, break-offs are scattered throughout the survey.

Table 2: 1999 NSAF Number of Break-offs and Partial Completes , by Section

Section in which break-off occurred		MKA Interviews N (%)	Childless Adult Interviews N (%)
E	Health Care Coverage	539 (1.64)	222 (1.24)
F	Health Care Use and Access	663 (2.06)	448 (2.53)
G	Child Care	288 (0.91)	NA (NA)
H	Nonresidential Parents	114 (0.36)	NA (NA)
I	Employment and Earnings	715 (2.29)	289 (1.67)
J	Family Income	302 (0.99)	132 (0.78)
K	Welfare Program Participation	230 (0.76)	81 (0.48)
L	Adult Education and Training	114 (0.38)	47 (0.28)
M	Housing and Hardship	119 (0.40)	90 (0.54)
N	Issues, Problems, Social Services	166 (0.56)	60 (0.36)
O	Race, Ethnicity, Nativity	29 (0.10)	14 (0.08)
P	Closing	22 (0.07)	17 (0.10)
Break-offs and Partial Completes (CDs + CPs)		3,301	1,400

Weighting for the NSAF is achieved in several stages. The initial household level base weight is subjected to a nonresponse adjustment for nonresponse on the screener interview followed by poststratification to control totals on the number of households. The probability of selecting the child within the household is then applied to the poststratified household weight to arrive at a child base weight. This weight is then subjected to a nonresponse adjustment to account for extended interview nonresponse. Finally,

poststratification to person level control totals is conducted to arrive at the final estimation weights.[2]

Results

Table 3 presents estimates for child characteristics for break-offs (CDs) and those for completed cases in the main NSAF data (CO and CP cases). The first three columns of estimates use the child base weights (which exclude person level nonresponse and poststratification adjustments). The final two columns show estimates based on the extended interview nonresponse adjusted weight and the final weight which includes poststratification at the person level. Since extended nonresponse (including break-offs) for area sample cases is virtually nonexistent, the analysis here focuses solely on telephone sample cases.

Significance tests were carried out comparing child base weighted estimates for break off (CD) cases and cases in NSAF analytic files (CO and CP cases).[3] Compared to cases in NSAF data files, break-off cases have lower incomes, are more likely to be uninsured and more likely to participate in programs such as Medicaid, food stamps and Temporary Assistance for Needy Families (TANF) or Aid to Families with Dependent Children (AFDC). But the magnitude of the differences along with the relative numbers of break off cases does not appear to indicate the presence of bias in NSAF estimates from not including break off cases in NSAF analysis files.[4] Overall, the estimates that include the CD cases look very similar to estimates that exclude such cases that have been weighted for nonresponse.

In examining the degree to which the extended nonresponse adjustment and poststratification address possible bias from not getting complete interviews for all respondents, it is important to consider differences between nonrespondents due to breaking off and nonrespondents on the extended for other reasons, including immediate refusal. In other words, can break-offs in the NSAF serve as proxies for other types of nonrespondents?

Table 4 compares the characteristics of all cases sampled for the extended interview on screener income status and metro status. With regard to screener income, nonrespondents who complete at least through section D (break-offs) are more likely to report a screener income below 200 percent poverty than other refusals. In other words, respondents who will not take part in the extended interview at all (or at least through section D) tend to have higher incomes than respondents who break off the interview. On screener income, the other refusals look more similar to completed cases than break-offs. This is not the case with metro status, where break-offs do not appear to be very different from both completes and other refusals.

Finally, it was hypothesized that respondents with lower incomes were more likely to break off the survey than those with higher incomes since they are asked more questions than higher income respondents. I ran a logistic regression in which whether or not the extended interview was broken off (with CD and CP cases being classified as break-offs) is regressed on variables used in the extended nonresponse adjustment as well as variables thought to be associated with interview length. In addition to a dummy variable indicating screener income below 200 percent poverty, other variables that were used in the extended nonresponse adjustment and included in the model are 1) in a central city, 2) not in a central city but in an MSA, 3) if screener income is unknown (answered "don't know" or refused and 4) whether the case had a completed screener interview in round 1 or was from the new RDD sample for round 2. As additional indicators of respondent burden based on length of interview, I include household size, the presence of a spouse/partner and whether two or one focal child is asked about in the interview. Larger households, cases where there is a spouse/partner present and interviews where there are two focal children should take longer to complete and therefore, should have higher rates of breaking off, all else being equal.[5]

---

[2] For a fully elaborated description of the weighting process, see Brick, Broene, Ferraro, Hankins and Strickler (2000).

[3] Standard errors were estimated using a jackknife replicate variance technique.

[4] Differences for items in section J (receipt of AFDC, food stamps, owning a car and family income) are probably overstated since these variables are only observed for break offs that managed to reach these questions. We do not know the answers to these questions for respondents that broke off prior to J. If we restrict the analysis of other items to cases that completed section J, many of the other differences already reported grow even larger. For example, among those break-offs that completed section J but broke off in section K, the percentage of children covered by Medicaid increases to 29 percent.

[5] This analysis was done using the household as the unit of observation rather than the focal child. Household replicate weights (which include household level poststratification to household control totals) were used in the estimation.

Table 3: Comparison of Break-offs (CD) vs. Completed cases (CO, CP) - RDD Cases Only

		Child Base weighted			Extended Nonresponse adjusted	Final Poststratification
Section	Variable	CD	CO, CP	Overall	CO, CP	CO, CP
Screener	Screener - income below 200% poverty	*38.6*	31.9	32.5	32.5	35.0
	Owner	70.3	75.2	74.8	74.6	71.3
	Not all family members insured	17.4	16.0	16.1	16.4	17.5
B	Fair or poor health	4.9	4.1	4.2	4.2	4.5
	Heard of Medicaid	*85.6*	89.6	89.3	89.6	89.4
	Heard of CHIP	*38.0*	42.0	41.6	42.4	43.3
	Not confident in getting care	*9.2*	6.9	7.1	7.0	7.3
	Limiting condition	*6.8*	8.6	8.4	8.7	8.5
C	Positive school engagement	40.8	40.4	40.4	40.2	40.1
	Negative school engagement	21.3	18.6	18.8	18.7	18.9
	Suspended or expelled	11.9	12.5	12.4	12.6	13.2
D	Biological father in HH	*62.4*	69.9	69.2	69.4	68.0
	Biological mother in HH	88.5	90.3	90.1	90.2	90.2
E	Currently uninsured	*13.3*	9.7	10.0	9.9	10.6
	Currently covered by Medicaid	*16.4*	12.7	13.0	13.1	14.6
F	Postponed medical care	2.6	2.8	2.8	2.8	2.8
	Postponed dental care	*4.4*	6.6	6.5	6.7	6.7
I	MKA currently working	71.5	71.5	71.5	71.2	70.9
	Spouse/partner working	84.8	87.4	87.4	87.2	86.6
	Both MKA and Spouse working	61.2	60.8	60.8	60.4	59.5
J	Received AFDC last year	*18.0*	4.3	4.4	4.5	5.0
	Received Food stamps last year	*26.3*	10.4	10.5	10.7	12.1
	Own a car	*81.3*	92.0	91.9	91.5	90.1
	Extended - income below 200% poverty	*48.3*	33.2	33.4	34.0	36.6

CD - complete through section D, broke off prior to completing section K; cases not in NSAF analysis files; CO - completed interviews; CP - complete at least through section K, broke off prior to completing interview; cases are in NSAF analysis files

Entries in italics indicate differences in estimates using child base weight between CD and CO, CP cases are statistically significant at .1 level (two tailed tests)

Table 4: 1999 NSAF Screener Income and Metro Status by Extended Result
(RDD cases, child base weight)

	Screener Income						Metro Status					
Extended Result	Above 200% poverty		Below 200% poverty		Unknown		In Central City		In MSA, not in Central City		Not in MSA	
	%	SE	%	SE	%	SE	%	SE	%	SE	%	SE
Completes	64.2	0.5	31.9	0.5	3.9	0.2	31.7	0.5	47.9	0.4	20.3	0.4
Break-offs	54.9	2.0	38.6	1.8	6.5	1.0	36.2	1.6	46.8	1.8	17.0	1.5
Other Refusals	62.0	2.0	28.1	1.9	9.9	1.1	34.4	1.8	51.5	1.9	14.1	1.3
Other Nonresponse	47.4	2.7	42.6	2.8	10.0	1.6	44.3	3.2	44.1	2.8	11.6	1.7

Results of the regression are shown in Table 5. All of the variables in the model that were used as part of the extended nonresponse in creating the final analysis weights have statistically significant effects on breaking off. Household size and the presence of two focal children do not appear to be related to breaking off.

Upon further reflection, the fact that the presence of a spouse/partner has a negative effect on breaking off is not all that surprising. While the presence of a spouse/partner may increase the length of the interview for some sections (e.g. sections I, on employment and earnings and section L, on adult education and training), for other sections of the survey, the presence of a spouse/partner may make the interview shorter. For example, if both biological or adoptive parents of the focal child are in the household, section H, on nonresidential parents, is skipped entirely.

Table 5: Logistic Regression of Breaking Off Extended Interview (CD or CP vs. CO)

Variable	Est.	SE
Intercept	-1.914	0.142
In Central City	0.376	0.081
In MSA, not in Central City	0.274	0.086
Screener income below 200% poverty	0.296	0.074
Screener income unknown	0.569	0.137
Round 1 Complete or New Sample	-0.530	0.081
Household Size	-0.009	0.026
Presence of spouse/partner	-0.189	0.082
Two focal children	0.094	0.076

$F = 14.672$ (d.f. = 14,47), $p < .001$

Discussion

The analysis provides little evidence of significant nonresponse bias due to the exclusion of break-off cases from NSAF analysis files. While there are some significant differences on survey items between break-offs and completed cases, the magnitude of these differences combined with the fairly low nonresponse rate on the extended interview do not suggest the presence of serious biases in NSAF estimates due to excluding break-off cases. In general, estimates that include the break-off cases appear very similar to fully weighted estimates that exclude break-off cases. This would suggest that there is little reason to revise NSAF extended interview nonresponse adjustments to account for break-off cases more explicitly.

While a weighting approach to dealing with partial nonresponse in the NSAF seems unnecessary, adopting an imputation approach to the problem would be desirable from the standpoint of response rates. Since the NSAF is composed of a separate screener and extended interview, the final response rate is the product of the screener response rate and the extended interview response rate. If all break-off cases were considered partial completes, this would add about 3-4 percentage points to the NSAF final response rate.

However, adopting an imputation strategy for dealing with partial nonresponse in the NSAF presents some practical problems that are difficult to overcome. The NSAF questionnaire features fairly elaborate skip patterns. For example, many, if not most of the questions that are asked about focal children are age-specific. The key age intervals for focal children in terms of questionnaire items are 0-2, 3-4, 5, 6-11, 12, 13-17. The NSAF also features questions whose displays vary by state in that the names for health insurance programs and cash assistance programs are unique to each state. It may be especially difficult to preserve associations between survey variables in the NSAF in carrying out imputations for missing data in break-off cases.

References

American Association for Public Opinion Research (2000) *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys.* Ann Arbor, Michigan: AAPOR.

Brick, J.M., P. Broene, D. Ferraro, T. Hankins, and T. Strickler (2000) *1999 NSAF Sample Estimation Survey Weights*, The Urban Institute, Washington, DC.

Brick, J.M., and G.H. Kalton (1996) Handling missing data in survey research, *Statistical Methods in Medical Research*, 5: 215-238.

Groves, R.M., and M.P. Couper (1998) *Nonresponse in Household Interview Surveys.* John Wiley and Sons, New York.