

The Hidden 25 Percent: An Analysis of Nonresponse on the 1980 General Social Survey

TOM W. SMITH

NONRESPONSE can seriously bias survey estimates and distort inferences. The relationship between nonresponse and survey estimates is simple and well defined, but the actual impact is ill known. For the sample mean of a particular variable (Y) the association is

$$\bar{Y} = W_1\bar{Y}_1 + W_2\bar{Y}_2 \quad (1)$$

where W_1 and W_2 are the proportion respondents and nonrespondents. The relative bias (RB) of using the response mean to equal the sample means is

$$RB(\bar{Y}_1) = W_2 \frac{(\bar{Y}_1 - \bar{Y}_2)}{\bar{Y}} \quad (2)$$

We can see that the relative bias is serious only when the nonresponse rate (W_2) is large and the difference in the means is great. Given this simple formula we can easily measure the magnitude of the nonresponse bias. The problem is that while we know the nonres-

Abstract Methods for estimating nonresponse bias are reviewed and several methods are tried on the 1980 GSS. The results indicate that various estimating procedures are inappropriate and that even the more promising techniques can provide faulty estimates of nonresponse bias. By its nature, nonresponse bias is very difficult to assess accurately and no simple, certain method exists.

Tom W. Smith is Senior Study Director, National Opinion Research Center, University of Chicago. This research was done for the General Social Survey Project directed by James A. Davis. The project is supported by the National Science Foundation, Grant No. SOC77-03279. This is an abridged version of GSS Technical Report No. 25 published by NORC, 1981. The author wishes to thank James A. Davis, Howard Schuman, and Stanley Presser for their comments.

ponse rate, we do not know the nonresponse mean since we have no measure of Y among nonrespondents.

Two alternatives are usually presented in discussing nonresponse—how to minimize nonresponse and how to estimate and correct for differences between the respondents and nonrespondents. In this paper we ignore the first alternative, accepting that a nonresponse rate of .25 is typical for good, state-of-the-art surveys (Smith, 1978; Davis et al., 1980; and Groves and Kahn, 1979). Instead, we will review the various existing approaches to estimating the characteristics of nonrespondents and then apply several of the proposed approaches to nonresponse on the 1980 GSS.

Measuring Nonrespondents and Assessing Nonresponse Bias

Numerous methods have been proposed to estimate the attributes of nonrespondents (Daniel, 1975). Some are appropriate for certain types of surveys (e.g., list samples only) while others can be used with modification across various methods of administration with various sample frames (e.g., from mail lists to RDD telephone). Attention will focus primarily on methods that are appropriate, or at least have been offered as appropriate, for face-to-face, national surveys. Among other things, this eliminates list samples where information about the respondent is known prior to the survey. Our review of nonresponse studies found nine major approaches to assess and adjust for nonresponse:

1. External population checks
2. Geographic/aggregate level data
3. Interviewer estimates
4. Interviewing nonrespondents about nonresponse
5. Subsampling of nonrespondents
6. Substitution for nonrespondents
7. Politz-Simmons adjustment
8. Extrapolation based on difficulty
9. Conversion adjustments

Probably the simplest check is to compare sample estimates (usually distributions) to some universe figures or preferred sample estimates such as the U.S. Census or the Current Population Survey (Crossley and Fink, 1951; Stephen and McCarthy, 1958; Smith, 1979; and Presser, 1981). Strictly speaking, when using such a criterion comparison, one is not checking how much difference comes from nonresponse but how much comes from nonresponse *and* all other

sources (item unreliability, interviewer error, etc.). If one shows that differences are within sampling error, then either no noticeable non-response bias exists on the variable being compared or nonresponse bias is being offset by other, countervailing errors. Similarly, a large difference does not specify nonresponse as the cause. This imprecision is, of course, undesirable from the perspective of studying non-response per se, but since one typically wants to know primarily whether the survey is reliable and representative, such general checks often serve the ultimate purpose satisfactorily. Unfortunately, however, superior estimates are often unavailable and are at best usually limited to a few demographics. Since representativeness on one variable is not generalizable to all variables or relationships in question, the usefulness of this approach is limited.

Two related means of assessing bias are the geographic/aggregate-level approach (Houseman, 1953; Dunkelberg and Day, 1973; Hawkins, 1975; DeMaio, 1980; House and Wolf, 1978). Since the geographic location of the sample household is known, one can code for all cases certain aggregate-level data such as (1) region and city type, (2) figures for census tract or other units, (3) interviewer description of neighborhood, and (4) interviewer description of dwelling unit. These can be recorded for all households and they can include many contextual variables that are commonly used in survey analysis. However, they still cover only a small fraction of the variables of interest on a typical survey and suffer from the problem of the nongeneralizability of representativeness across unexamined variables. In particular, they do not apply to individual level attributes. On the other hand, since we have complete and accurate observations, we have a precise measure of nonresponse bias for the covered variables.

Going a step beyond the geographic/aggregate-level approach, interviewers can make observations and estimates about households and individual respondents (DeMaio, 1980; Moser and Kalton, 1972; Lansing et al., 1971). The advantage is that one can expand the range of comparable variables and that one can include individual-level variables. One problem is that it is not possible to get complete information. On the 1980 GSS and the 1968 Michigan election survey no estimates were possible on race for .248 and .215 of nonrespondents, respectively, or on income for .364 and .294. In addition, interviewer estimates are usually more error-prone than directly acquired data. Finally, there is a clear limit to what variables can be checked.

To get into nonrespondents' minds, special field procedures are sometimes used: interviewing nonrespondents, intensive follow-ups

on a subsample of nonrespondents, and substituting for nonrespondents. In the nonrespondent interviewing approach refusers are asked a few questions about why they refused to participate and/or a few demographics. The Bureau of Social Science Research (1981) was able in a recent study to get a noninterview interview with 53 percent of refusers, and Wilcox (1977) got some demographics for 29 percent of refusers, for example. The problem of this approach is that it applies only to refusers rather than not-at-homes and others,¹ information is available for only some refusers, and it is difficult to pick up much substantive information pertinent to a study in a necessarily short nonrespondent interview.

The subsampling of nonrespondents is used to get an estimate of nonrespondents by making extraordinary efforts to interview a representative sample of them (Lundberg and Larsen, 1949; Lagay, 1969–70). This method is often used successfully in mail surveys. After several mailings and reminders, a sample of nonrespondents is drawn and this group is approached via some more persuasive medium, such as telephone calls or personal visits (Hansen and Hurwitz, 1946; Kish, 1965:556; Ognibene, 1971). It is difficult to use and not particularly cost-efficient in a first-rate, face-to-face full probability sample. When substantial efforts are made to get all cases, the returns from extraordinary efforts beyond the standard procedures are likely to be small.

Another field procedure that has been proposed is substitution. Under this method alternative households are added to the sample to replace nonrespondent households. While this method is useful to achieve a prescribed sample size, it tends to replace nonrespondents with people who resemble respondents rather than nonrespondents, and this approach does not appear to be widely used to deal with nonresponse bias. However, there are related techniques that do somewhat the same thing. Block quota samples require interviewers to fill a certain quota, for example, so many employed/unemployed females and so many young/old males from a given block. If no one is home, a household refuses, or no one in the household fits the remaining quota slots, then the interviewer proceeds to the next house. In effect, the block quota sampling uses a substitution procedure, passing over unavailable households and substituting available ones instead. The quotas are designed to insure that the hard-to-get groups are represented, so in effect one does not substitute easy for hard households/respondents, but merely gets those easy and hard

¹ It might be possible to reach the not-at-homes with a similar nonresponse interview by leaving a mail-back version.

households/respondents that are available at the moment. On its face this sampling approach seems likely to increase nonresponse error, but experiments between full probability and block quota surveys show few differences (Stephenson, 1979).

Another superficially related approach is Kish's replacement procedure (Kish and Hess, 1959; Kish, 1965:560–62), which substitutes previous nonrespondent households (from an earlier and similar survey) for nonrespondents to the current survey, assuming that they are reasonable replacements and that it will be possible to secure interviews in a high percentage of them. Administrative problems and the difficulty of getting a high response rate from the replacement households have severely limited the actual use of this approach. The Current Population Survey does use a related technique, however. Unlike most panel surveys, the CPS does not exclude nonrespondents from subsequent waves. Not-at-homes are continued in the sample and an attempt is made to reach them on all waves despite absences on earlier waves. Refusers are approached a second month, and only after repeated refusals are households dropped from subsequent waves. Thus, each CPS cross-section includes "replacement" households that were nonrespondents on earlier surveys.

Finally, there are several methods for estimating the effects of nonresponse directly from respondents, by the Politz-Simmons approach, by extrapolation based on difficulty, and by convertibility. In the times-at-home or Politz-Simmons approach (Politz and Simmons, 1949; Kish, 1965:559–60; Moser and Kalton, 1972:178–81), respondents are asked how many times they were at home at the time of interview during the last x number of days. This is taken as representing their probability of availability. Respondents are then weighted according to the inverse of the number of days they were home. This method assumes that nonresponse is basically a function of availability (as the block quota samples do) and that weighting adequately adjusts for the probabilities of availability. In two empirical tests (Durbin and Stuart, 1954; Simmons, 1954) this was not found to perform as well as callbacks, however. The technique is still commonly employed; Gallup, for example, included a times-at-home weight as part of its surveys from 1960 to 1967 and since then has included a weight factor that apparently combines a times-at-home weight with a poststratification weight (Gaertner, 1976).

The difficulty method uses some measure of how hard it was to get an interview from a respondent. This might be the number of mailings, visits, or telephone calls, how long it took to get a response, or some measure derived from these indicators of difficulty. While there are numerous variations, the basic approach determines whether a

particular variable is related to difficulty (e.g., the proportion employed rising with the number of attempts). If a linear or some other regular relationship is found, then this association is used to impute the distribution of the variable among nonrespondents. This approach has the advantage of allowing an estimate for every variable contained in the survey and thus avoids the problem of nongeneralizability that plagues several of the methods considered above. It rests on the premise that difficulty is related to final nonresponse. If the final nonrespondents differ from the merely difficult, then this procedure will obviously misestimate the attributes of the final nonrespondents. This approach is one of the most frequently employed in estimating nonresponse and has shown some impressive results, especially in mail surveys where known attributes from a list sample could be compared to estimates from the difficulty extrapolations (e.g., Hilgard and Payne, 1944; Crossley and Fink, 1951; Hendrick, 1956; Mayer, 1964; Dunkelberg and Day, 1973; Cranberg, 1975; Filion, 1975-76; Armstrong and Overton, 1977). At the same time, however, there have been a number of criticisms of the method and cautions about its general application. Stephan and McCarthy warn that "great care would have to be exercised in carrying out this extrapolation, and its use is not recommended except under exceptional circumstances" (1958:257). (See also criticism and subsequent rebuttal by Ellis et al., 1970, and Filion, 1976.)

Unlike the difficulty approach, which is aimed at all nonrespondents, the convertibility approach uses converted refusals as estimates for final refusals. Usually the converted are seen as substitutes for the final refusals although it is possible to do an extrapolation with the first group being the respondents who never refused, the converts making the second point, and the final refusers as the last group² (Benson et al., 1951; Stinchcombe et al., 1981; Robbins, 1963; O'Neil, 1979; DeMaio, 1980; Andersen et al., 1979). As in the difficulty approach, all variables can be studied and the appropriateness of the technique rests on the supposition that final refusers are like temporary refusers (in the case of the substitution approach) or at least more like temporary refusers than cooperative respondents (in the case of extrapolation).

² So far this approach uses only two types of respondents, cooperators and converts. This means that there is either no significant difference between the two groups or a linear relationship to extrapolate to final refusers. If temporary refusers were further subdivided into easy or hard to convert, or some other refinement, such as number of attempts needed to convert, then it would be possible for other relationships to emerge. Apparently no one has attempted such a refinement and the usually small number of total converts would make such refined analysis difficult.

In brief, a number of procedures have been proposed to assess the impact of nonresponse. Some methods, such as using geographic/aggregate-level data, allow a complete assessment of nonresponse bias for a limited number of variables; other techniques, such as difficulty extrapolation, permit estimates of nonresponse bias for all variables. None of the methods permit the complete determination of nonresponse bias for all variables.

Analysis of Nonresponse on 1980 General Social Survey

To assess the impact of nonresponse on the 1980 GSS we selected four of the more promising and widely applied techniques: (1) geographic/aggregate-level analysis, (2) interviewer estimates, (3) extrapolation for difficulty, and (4) convertibility. The 1980 GSS was a multistage, full-probability sample of the contiguous United States. Households were sampled according to NORC's equal probability selection procedures and a Kish table was used to choose a respondent from the designated households (King and Richards, 1972; GSS, 1980). Interviewers kept a record of calls, noting each attempt to contact the household or respondent by the date and time of the attempt, method of contact (personal/telephone), and the outcome (not-at-home, temporary refusal, interview, etc.). If a contact could be made, a household enumeration folder was filled out listing all household members along with their relationship to head of household, age, sex, marital status, and location (staying at household/staying elsewhere). If an interview was secured, a questionnaire was completed. For nonrespondents a noninterview report recorded the reason for nonresponse, descriptions of why the nonresponse occurred (e.g., why a person was never found at home), and interviewer estimates of the family income, race of household, number of adults, number of adult males, presence of married couple, and age of head of household.

The 1980 GSS had a net sample of 1,931. There were 1,468 completed cases, 315 refusals, 66 not-at-homes, 78 others (mostly not mentally or physically capable of participating, but also including administrative errors), and 4 lost documents.³ This gives a response rate of .760, a refusal rate of .163, a not-at-home rate of .034, and an "other" rate of .042 (including the four unclassified cases).

In the subsequent analysis we examine nonresponse in general as

³ These figures differ slightly from those presented in Davis et al., 1980. The difference comes from internal inconsistencies between the record-of-calls and the noninterview report forms.

well as the main types of nonresponse-refusals, not-at-homes, and others. Both past research and findings from this study suggest that these groups are quite different in their motivations for nonresponding, in their demographic profile, and in other notable ways (Stinchcombe et al., 1980; Kish, 1965; O'Neil, 1979; Bebbington, 1970).

Geographical/Aggregate-Level Analysis

The geographic/aggregate-level analysis was restricted to measures of community type and region. As Table 1 shows, there are large differences in the response rates across city types and regions. Response rates are lowest in central cities, rise moderately in suburbs and exurbia within metropolitan areas, and increase substantially in rural areas. This urban-rural difference replicates similar findings from numerous other studies (Lansing et al., 1971; Moser and Kalton, 1972; DeMaio, 1980; House and Wolf, 1978; Groves and Kahn, 1979) and derives mainly from variation in the refusal rate. The not-at-homes follow roughly the same pattern as refusals, and the others

Table 1. Response/Nonresponse by City Type and Region^a

<i>Response/Nonresponse</i>	<i>Completed Case</i>	<i>Refusal</i>	<i>Not-at-Home</i>	<i>Other</i>	
SRCBELT					
Central city of 12 largest SMSAs	.665	.216	.065	.054	(185)
Central city 13-100 largest SMSAs	.713	.212	.039	.036	(307)
Suburb of 12 largest SMSAs	.697	.211	.038	.054	(185)
Suburb of 13-100 largest SMSAs	.751	.190	.027	.032	(221)
Other urban	.791	.144	.031	.034	(731)
Other rural	.849	.081	.020	.050	(298)
	$\chi^2 = 44.6$ prob. = .0001 ^b				
REGION					
New England	.716	.230	.041	.014	(74)
Midatlantic	.691	.234	.038	.038	(346)
East North Central	.760	.160	.053	.028	(400)
West North Central	.819	.118	.024	.039	(127)
South Atlantic	.767	.124	.046	.063	(348)
East South Central	.854	.131	.000	.015	(130)
West South Central	.789	.132	.026	.053	(152)
Mountain	.806	.153	.020	.020	(98)
Pacific	.758	.171	.016	.056	(252)
	$\chi^2 = 50.5$ prob. = .0012				

^a Analyses were also conducted on two other community type variables: SIZE—population at local community, and XNORCSIZ—city, suburb, town, rural typology.

^b Here and elsewhere we used calculations based on SRS assumptions. Since the sample was clustered this underestimates the true sampling variance.

appear to be scattered across city types. Regional response rates tend to be highest in the South and lowest in the Northeast, although the pattern is not completely uniform. Most previous studies have found some regional differences (except House and Wolf, 1978), but there is disagreement on where nonresponse is highest. Love and Turner (1975) found response rates lowest in the Northeast, and results from Schuman and Gruenberg (1970) and Dunkelberg and Day (1973) suggest a similar conclusion, but DeMaio (1980) found that refusals were lowest in the Northeast and highest in the West. As with city size, most variation is in the level of refusals.

Both city type and region exercised independent effects on response rates. When we controlled for city type, the South had a response rate 8.5 percentage points above the Northeast. With region controlled, suburbs and exurbia had a response rate 6.3 percentage points above central cities (Davis, 1975).

Interviewer Estimates

Next, nonresponse bias was examined by having interviewers estimate the following characteristics of the nonresponding households: race, family income, number of adult males, number of adults, presence of married couple, and age of head of household. Estimates were possible in two-thirds to three-quarters of households and were more often available for refusers than for others or not-at-homes (see Table 2). The absence of estimates from a substantial minority of nonrespondents as well as the probable unreliability of some of the estimates necessarily hampers the use of these interviewer estimates to study nonresponse bias.

Looking first at those nonrespondents with available information, we find no significant differences between completed cases and esti-

Table 2. Item Nonresponse Among Completed Cases and Respondents
(Proportion Missing)

Variables	Completed Cases	Nonrespondents			
		All	Refusals	Not-at- Homes	Others
Race	.000	.248	.203	.348	.346
Income	.075	.364	.314	.500	.449
Number of adult males	.000	.296	.273	.379	.321
Number of adults	.000	.303	.283	.379	.321
Married couple	.000	.320	.305	.424	.295
Age of head	.007	.285	.232	.439	.372
	(1,468)	(459)	(315)	(66)	(78)

mated nonrespondents on race, marital status, and number of adult males (see Table 3). Nonrespondent households have, however, older heads of household, fewer adults, and less middle income and poor.

We also discover that the profile of each type of nonrespondent is quite different. Refusers are somewhat more likely to be married, to have a middle income, and to be over 30 years old than respondents. Not-at-homes tend to be isolated individuals, less likely to be married, and more likely to live alone. The others are also isolated individuals, but in addition they are typically old and poor as well. These differences generally follow those in earlier studies.⁴

Table 3. Selected Characteristics Among Completed Cases and Estimatable Nonresponse Cases

Variables	Com- pleted Cases	Nonresponse Cases				Probability	
		All	Refusals	Not-at-		Completed Cases vs. Non- response	Completed Cases, Refusals, Not-at- Homes, Others
				Homes	Others		
Race							
White	.898	.910	.928	.837	.882		
Black	.095	.084	.064	.163	.118		
Other	.007	.006	.008	.000	.000	.789	.411
	(1,468)	(345)	(251)	(43)	(51)		
Income							
Less than \$7,000	.204	.226	.153	.212	.605		
\$7,000–19,000	.412	.541	.597	.455	.326	.000	.000
\$20,000+	.384	.233	.250	.333	.070		
Number of adult males							
None	.193	.251	.223	.244	.377		
One	.695	.684	.725	.634	.547	.066	.167
Two+	.113	.065	.052	.122	.066		
Married couple							
Yes	.606	.641	.721	.447	.455	.243	.000
No	.394	.359	.279	.553	.545		
Number of adults							
One	.252	.359	.305	.488	.491		
Two	.598	.509	.575	.341	.358	.008	.016
Three+	.151	.131	.119	.170	.151		
Age of head							
Under 30	.200	.076	.087	.081	.020		
30–64	.607	.591	.649	.703	.224	.000	.000
Over 65	.193	.332	.264	.216	.755		

⁴ Most previous research finds final nonresponse to be highest among the older ages (Lowe and McCormick, 1955; Lansing et al., 1971; Weaver et al., 1975; Hawkins, 1975;

The proportion of nonrespondents with estimates is lowest in large central cities and their suburbs and in rural areas, with smaller central cities, their suburbs, and small towns having estimates for a significantly higher proportion of cases. Missing estimates also appear to be highest in the West and among nonwhite households. These results suggest that the estimated nonresponse households are not typical of all nonrespondents and thus do not give an unbiased estimate of them.

Difficulty

Next, we examined the association between difficulty in completing an interview, as measured by the number of attempts, and respondent characteristics. We believed that difficulty resulted from three factors: availability (essentially the probability of a respondent being home at a given time), contactability (the probability of some other responsible household member being home at a given time), and reluctance (respondent's and/or informant's willingness to cooperate). The following groups were anticipated to have high availability: nonmembers of the labor force (especially homemakers); women; members of the labor force working few hours, not traveling, and not self-employed; widowed people; older people, infirm, and physically restricted; and low socioeconomic status people. Households with high contactability were anticipated to include: married couples and households with more than one adult, young children, and a spouse not in the labor force. Reluctant households were presumed to be urban, fearful of crime, and mistrustful of people. In addition we included race because of its close association with several of the preceding variables.

We tested the relationship between these independent variables and difficulty by using one-way analysis of variance. As Table 4 indicates, the chosen indicators of availability were related to difficulty as anticipated. Availability had a basically linear relationship with labor force participation, socioeconomic status, life stage (age and marital status), health, and sex. It was unrelated to race and self-employment. Too few occupations were identifiable as involving ex-

and DeMaio, 1980). Nonresponse was found higher for whites in three studies (Schuman and Gruenberg, 1970; Weaver et al., 1975; and Hawkins, 1975) and not significantly different across races in two studies (DeMaio, 1980; Lansing et al., 1971). Middle income groups are usually found to have the highest nonresponse (Lansing et al., 1971; DeMaio, 1980), but this may well be a function of estimating error. Studies are also divided as to whether nonrespondents vary by sex. Crossley and Fink (1951), Hawkins (1975), and DeMaio (1980) found no difference, but Bartholomew (1961), Lowe and McCormick (1955), and Smith (1979) found an underrepresentation of men.

Table 4. Analysis of Variance of Number of Attempts by Selected Variables^a

Variable	Probabilities		Deviation from Linearity
	Between Groups	Linearity	
Labor force	.0000	.0000	.0044
Hours	.0000	.0000	.4541
Self-employed	.2345	—	—
Marital status	.0069	.0012	.1774
Age	.0000	.0000	.1569
Health	.0000	.0000	.7900
Education	.0000	.0000	.0300
Prestige	.0000	.0000	.0510
Income	.0000	.0000	.0002
Sex	.0024	.0001	.2876
Adults	.0356	.2475	.0332
Children under 18	.0691	—	—
Number of children	.0009	.0000	.2569
Spouse working	.2371	—	—
SRCBELT	.0031	.0000	.7117
SIZE	.9822	—	—
XNORCSIZ	.0402	.0009	.7209
FEAR	.2214	—	—
TRUST	.4500	—	—

^a A parallel analysis using nonparametric methods (Davis, 1975) showed similar results except for marital status. The proportion unmarried did not significantly vary but the proportion postmarried varied linearly.

tensive traveling away from home to permit analysis of this factor. Contactability, on the other hand, did not show the anticipated relationships. The presence of a spouse and/or children at home were unrelated to difficulty, and the number of adults had a weak and uninterpretable nonlinear association. Reluctance showed intermediate results. Community type (SRCBELT, XNORCSIZ) showed the expected associations between urbanness and difficulty, but neither fear of crime nor mistrust of people was related to reluctance.

A stepwise multiple regression equation further clarified these relationships. Labor force participation is the strongest correlate of difficulty. High socioeconomic status also meant more difficulty, probably because of more social and occupational activities outside the home. Some of this might result, however, from the growing proportion of people in the upper social rankings "protected" from interviewers by doormen, security systems, and other barriers. The young also proved to be more difficult to reach, probably because of more socializing outside of the home. Urban dwellers were also

harder to reach. Part of this difficulty seems to result from the need for more calls to persuade reluctant respondents, but there may also be a greater tendency for urbanites to spend more time away from the homes.⁵ Finally, we find that people with children were easier to reach, probably because there is usually someone at home to contact (for details see Smith, 1981).

In order to evaluate the effect of nonresponse on nondemographics we ran attempts by a wide range of attitudes, behaviors, and socio-psychological scales. Only 27.4 percent showed significant variation with number of attempts. Of the significant relationships, most were linear (67.4 percent linear, no significant deviation; 15.2 percent linear, with significant deviation; and 17.4 percent nonlinear).

The hard-to-gets appeared to have three main characteristics—liberal political views (e.g., pro-abortion, civil rights, tolerance), high socioeconomic status (e.g., members of professional groups, never received governmental aid), and active and youthful life style (having received a traffic ticket, members of youth groups, watch less television, favor legalization of marijuana). Of all of these, only the life style variables added explanatory power to the basic demographic model (hours, education, SRCBELT, age, number of children). Favoring the legalization of marijuana and having received a traffic ticket were significantly related, and drinking just missed the cutoff.⁶

While this means that very few nondemographics are independently related to difficulty, many variables (27 percent on the GSS) are closely enough related to the independent variables to vary notably with number of attempts. This means that not only will variables directly related to availability be affected, but many attitudinal and behavioral variables will also be affected. We also took a purposive sampling of 21 bivariate relationships and examined whether they varied by number of attempts. We found few significant differences and no clearly discernible pattern.

Although we can use respondent characteristics to predict difficulty in completing an interview, these characteristics cannot be used to estimate the attributes of nonrespondents. This is because the number of attempts measures primarily how accessible a person is, while the final nonrespondents are made up primarily of refusals, not inaccessible. The proportion of refusals steadily rises with callbacks as the proportion of not-at-homes falls. After two calls, completed cases

⁵ Interviewers are just as likely, however, to find an urban respondent at home on the first call as a rural respondent.

⁶ In addition, those disagreeing that people shouldn't have children given the state of the world were harder to reach. It is unclear whether this relationship has any substantive meaning.

make up 32.5 percent, refusals 9.9 percent, not-at-homes 50.2 percent, others 6.1 percent, and miscellaneous 1.2 percent. After the final attempt, completed cases have risen to 76.0 percent and refusals to 16.3 percent, while not-at-homes fell to 3.4 percent, others to 4.0 percent, and miscellaneous to 0.2 percent. In brief, repeated callbacks nearly eliminated the not-at-home problem, while both in relative and absolute terms refusals increased. This relationship is also evident in the conversion rates. Only 35 percent of temporary refusals are converted to respondents while 91 percent of not-at-home (excluding those who were both not-at-homes and temporary refusals) were eventually "converted" to respondents. To use a difficulty measure to impute the attributes of final refusals is thus essentially to use the correlates of inaccessibility to predict the correlates of refusal. Given the differences in nonresponse motivations and known demographic profile, this is obviously an improper imputation procedure.

Of course, even if difficulty cannot be used to estimate the values of nonrespondents as a whole, it may be possible that this procedure would improve estimates of the final not-at-homes. We found that for five of seven variables—marital status, community type, region, race, number of adult males—the use of simple difficulty extrapolations did give us improved estimates of not-at-homes. In two cases—number of adults and age of head—the procedure proved inappropriate (for details see Smith, 1981).

Temporary Refusals

Finally, temporary refusals were used as an indicator of final refusals. On the basis of our review of this technique and reasons for refusals in general, we related temporary refusals to (1) mistrust and fear, (2) apathy toward social and political issues (replying DK to questions, not voting, no party identification), (3) negative psychological feelings (unhappiness, dissatisfaction, high anomia), (4) deviant behavior (having been arrested, receiving ticket), (5) attitudes toward science, (6) illness, (7) being too busy (long hours, labor force status), (8) uncooperativeness (interviewer rating and refusal to give family income), (9) place of residence, (10) conservatism, (11) socioeconomic status (education, income, occupational prestige), and (12) standard demographics (age, sex, race).

Only urbanness had a strong relationship to temporary refusals. In the 12 largest central cities 29 percent of the cases were temporary refusals while in rural counties only 11 percent of cases were temporary refusals. Of all other items only refusing to give family income had a significant (prob. = .035) association with reluctance. Of the

other variables only being cooperative and being fearful approached significance (prob. $< .10$). In general, the results show fewer and more modest associations than most previous studies. On the 21 liberalism items (race relations, spending priorities, morality/personal life style, tolerance of Communists and atheists, and abortion) there was not a single significant relationship between conservatism and refusing. This refutes evidence from Hawkins (1975); Schuman and Gruenberg (1970); and Benson et al. (1951), but agrees with Brannon et al. (1973). We find no association with low socioeconomic status as Benson et al. (1951) and O'Neil (1979) found, no tendency to reply "don't know" (Stinchcombe et al., 1981), and no association with race, or number of children (O'Neil, 1979). Our lack of a difference between temporary refusals and age confirms Benson et al., but contradicts O'Neil, who finds the elderly refusing more. In addition our one notable association, between urbanness and refusals, differs from DeMaio (1980), who found no variation between rural and urban. The results do follow Robbins's findings (1963) that there were no significant differences. In general, the discrepancy of results suggests that findings depend on the specialized populations sampled, the survey procedures used, or other variable conditions.

On one hand the lack of associations between the hypothesized variables and temporary refusals is encouraging. If we accept temporary refusals as indicators of final refusals then the lack of significant associations suggests that except for city type, final refusals are not significantly different from completed cases and therefore little bias is introduced. Yet it is somewhat surprising that temporary refusals had such consistently low correlations with variables that might be expected on theoretical grounds to show more substantial relationships. One can hypothesize variables (e.g., willingness to be interviewed) that would have large associations with refusing. In addition it is reasonable to suppose that other variables touching on privacy, misanthropy, paranoia, and fear, and other sociopsychological attitudes that should be closely related to willingness to be interviewed would show substantial associations. The fact that we were largely unable to find these associations may simply mean that we do not have the right variables, that indicators more closely related to refusing are needed before the anticipated relationships can be detected.⁷ It might also be that refusing is really more of a random occurrence, like a transitory mood, and therefore there are no other related variables, but the difficulty of converting temporary refusals and evidence from other studies (Stinchcombe et al., 1981) suggests that this is not the case. Another alternative is that temporary refusals do not adequately indicate attributes of final refusals. Perhaps many of the temporary refus-

⁷ This is being tested on the 1982 General Social Survey.

als, but not the final refusals, really represent transitory states. Unfortunately, we cannot fully test these alternatives.

We were able, however, to carry out a more general, less focused comparison between completed cases without temporary refusals (nonrefusals), temporary refusals, and final refusals on the nine variables for which we had aggregate-level data or interviewer estimates. On the three geographic variables, temporary refusals performed well. Estimates using substitution or extrapolation were closer to the true distribution than was the case when assuming no difference (i.e., completed cases equal all cases). On three of the interviewer estimates (number of males, number of adults, race) there are no significant differences between the three groups. While temporary refusals are not really needed for estimating distribution of final refusals, they correctly predict the characteristics of final refusals. For marital status the temporary refusals are in the right direction and provide a better estimate than assuming no difference between nonrefusals and final refusals. On age, however, temporary refusals point in the wrong direction for the old and in the correct direction for the young. Finally, on income temporary refusals do not differ from nonrefusals, while nonrefusals significantly differ from final refusals and temporary refusals differ from final refusals. This may result from the unreliability of the income estimates (see Smith, 1981).

Overall, the evidence is mixed about the appropriateness of using temporary refusals as indicators of final refusals. The fact that temporary refusals perform well on the geographic variables is encouraging both because there are no complications from missing values in these cases and these variables have the strongest theoretical connection with refusals. The evidence is further mixed on whether temporary refusals can best be substituted for final refusals or used to extrapolate to them. In general, the performance of temporary refusals is satisfactory enough to merit further investigation and selective application, but it is clear that neither substitution nor extrapolation of temporary refusals can be used routinely as a sure adjustment for final refusals.

Conclusion

We come close to the conclusion that nothing works in estimating nonresponse bias. Each of the methods we examined proved to be of limited usefulness. The geographic/aggregate-level approach allows definitive measurement of nonresponse bias, but it is limited to readily observable data or data linked from other sources such as the Census. Usually only a few variables of interest are available, and results from them are not necessarily generalizable to other variables (Lagay, 1969–70). Interviewer estimates help to expand the range of

variables that can be checked, but (1) missing estimates (typically 25 to 35 percent of cases) prevent complete coverage, (2) the estimated portion of nonrespondents may not be representative of all nonrespondents, (3) some estimates probably have low reliability (e.g., income), and (4) the range of checkable variables is limited. Difficulty extrapolation was found to be inappropriate for nonresponse in general because of the high proportion of refusals among nonrespondents. While probably often useful for imputing to the not-at-homes, evidence in the case of age (and labor force participation) indicates that final not-at-homes are not always extensions of the hard-to-get. Temporary refusals also gave mixed results in estimating final refusals. Even when temporary refusals are indicative of final refusals the evidence is unclear whether substitution or extrapolation would be most appropriate. In sum, our analysis of nonresponse on the 1980 GSS suggests that there is no simple, general, accurate way of measuring nonresponse bias.

References

- Andersen, Ronald, Judith Kasper, and Martin R. Frankel
1979 *Total Survey Error: Applications to Improve Health Surveys*. NORC Series in Social Research. San Francisco: Jossey-Bass.
- Armstrong, J. Scott, and Terry S. Overton
1977 "Estimating nonresponse bias in mail surveys." *Journal of Marketing Research* 14:396-402.
- Bartholomew, D. J.
1961 "A method of allowing for 'not-at-home' bias in sample surveys," *Applied Statistics* 10:52-59.
- Bebbington, A. C.
1970 "The effect of nonresponse in the sample survey with an example." *Human Relations* 23:169-80.
- Benson, Sherwood, Wesley P. Booman, and Kenneth E. Clark
1951 "A study of interview refusals." *Journal of Applied Psychology* 35:116-19.
- Brannon, Robert, et al.
1973 "Attitude and action: a field experiment joined to a general population survey." *American Sociological Review* 38:625-36.
- Bureau of Social Science Research
1980 "Long interviews are not main cause of refusals." *BSSR* 14 (Fall):1-2.
- Cranberg, Gilbert
1975 "Mail survey respondents and nonrespondents." *Journalism Quarterly* 52:542-43.
- Crossley, Helen M., and Raymond Fink
1951 "Response and nonresponse in a probability sample." *International Journal of Opinion and Attitude Research* 5:1-19.
- Daniel, Wayne W.
1975 "Nonresponse in sociological surveys: a review of some methods for handling the problem." *Sociological Methods and Research* 3:291-307.
- Davis, James A.
1975 "Analyzing contingency tables with linear flow graphs: d-systems." In David Heise (ed.), *Sociological Methodology*. San Francisco: Jossey-Bass.
- Davis, James A., Tom W. Smith, and C. Bruce Stephenson
1980 *General Social Surveys 1972-1980: Cumulative Codebook*. Chicago: National Opinion Research Center.

- De Maio, Theresa
1980 "Refusals: who, where, and why." *Public Opinion Quarterly* 44:223-33.
- Dunkelberg, William C., and George S. Day
1973 "Nonresponse bias and callbacks in sample surveys." *Journal of Marketing Research* 10:160-68.
- Durbin, J., and A. Stuart
1954 "Callbacks and clustering in sample surveys." *Journal of the Royal Statistical Society (A)* 117:387-428.
- Ellis, Robert A., Calvin M. Endo, J. Michael Armer
1970 "The use of potential nonrespondents for studying nonresponse bias." *Pacific Sociological Review* 13:103-9.
- Filion, F. L.
1975-76 "Estimating bias due to nonresponse in mail surveys." *Public Opinion Quarterly* 39:482-92.
1976 "Exploring and correcting for nonresponse bias using follow-ups of nonrespondents." *Pacific Sociological Review* 19:401-8.
- Gaertner, Karen Newman
1976 "The use of AIPO surveys: to weight or not to weight." In James A. Davis (ed.), *Studies of Social Change Since 1948*, NORC Report 127A. Chicago: National Opinion Research Center.
- Groves, Robert A., Robert L. Kahn
1979 *Surveys by Telephone: A National Comparison with Personal Interviews*. New York: Academic Press.
- GSS
1980 "Interviewer procedural manual." Unpublished NORC booklet.
- Hansen, M. H., and W. N. Hurwitz
1946 "The problem of nonresponse in sample surveys." *Journal of the American Statistical Association* 41:517-29.
- Hawkins, Darnell F.
1975 "Estimation of nonresponse bias." *Sociological Methods and Research* 3:461-88.
- Hendricks, Walter A.
1956 *The Mathematical Theory of Sampling*. New Brunswick: Scarecrow Press.
- Hilgard, Ernest R., and Stanley L. Payne
1944 "Those not at home: riddle for pollsters." *Public Opinion Quarterly* 8:254-61.
- House, James S., and Sharon Wolf
1978 "Effects of urban residence on interpersonal trust and helping behavior." *Journal of Personality and Social Psychology* 36:1029-43.
- Houseman, Earl E.
1953 "Statistical treatment of the nonresponse problem." *Agricultural Economics Research* 5:12-18.
- King, Benjamin F., and Carol Richards
1972 "The 1972 NORC national probability sample." Unpublished NORC memo.
- Kish, Leslie
1965 *Survey Sampling*. New York: John Wiley.
- Kish, Leslie, and Irene Hess
1959 "A 'replacement' procedure for reducing the bias of nonresponse." *American Statistician* 8 (October):17-19.
- Lagay, Bruce W.
1969-70 "Assessing bias: a comparison of two methods." *Public Opinion Quarterly* 33:615-18.
- Lansing, John B., et al.
1971 *Working Papers on Survey Research in Poverty Areas*. Ann Arbor: Survey Research Center.
- Love, Lawrence T., and Anthony G. Turner
1976 "The census bureau's experience: respondent availability and response rates." *Proceedings of the Business and Economics Section, American Statistical Association, Washington, D.C.:* ASA.

- Lowe, Francis E., and Thomas C. McCormick
1955 "Some survey sampling biases." *Public Opinion Quarterly*:303-15.
- Lundberg, George A., and Otto N. Larsen
1949 "Characteristics of hard-to-reach individuals in field surveys." *Public Opinion Quarterly* 13:487-94.
- Mayer, Charles S.
1964 "The interviewer and his environment." *Journal of Marketing Research* 1 (November):24-31.
- Moser, C. A., and G. Kalton
1972 *Survey Methods in Social Investigation*, 2nd ed., New York: Basic Books.
- Ognibene, Peter
1971 "Correcting nonresponse bias in mail questionnaires." *Journal of Marketing Research* 8:233-35.
- O'Neil, Michael J.
1979 "Estimating the nonresponse bias due to refusals in telephone surveys." *Public Opinion Quarterly* 43:218-32.
- Politz, A., and W. R. Simmons
1949 "An attempt to get the 'not-at-homes' into the sample without callbacks." *Journal of the American Statistical Association* 44:9-31.
- Presser, Stanley
1981 "A comparison of Harris sample composition with census, IRS, and NORC." In Elizabeth Martin, Diana McDuffee, and Stanley Presser (eds.), *Sourcebook of Harris National Surveys: Questions 1963-76*. Chapel Hill, North Carolina: Institute for Research in Social Science.
- Robins, Lee N.
1963 "The reluctant respondent." *Public Opinion Quarterly* 27:276-86.
- Schuman, Howard, and Barry Gruenberg
1970 "The impact of city on racial attitudes." *American Journal of Sociology* 76:213-61.
- Simmons, W. R.
1954 "A plan to account for 'not-at-homes' by combining weighting and call-back." *Journal of Marketing* 19:42-53.
- Smith, Tom W.
1978 "Response rates on the 1975-78 General Social Surveys with comparisons to the Omnibus surveys of the Survey Research Center, 1972-1976." GSS Technical Report No. 7. Chicago: National Opinion Research Center.
1979 "Sex and the GSS: nonresponse differences." GSS Technical Report No. 17. Chicago: National Opinion Research Center.
1981 "The hidden 25 percent: an analysis of nonresponse on the 1980 General Social Survey." GSS Technical Report No. 25. Chicago: National Opinion Research Center.
- Stephan, Frederick F., and Philip J. McCarthy
1958 *Sampling Opinion: An Analysis of Survey Procedure*. New York: John Wiley.
- Stephenson, C. Bruce
1979 "Probability sampling with quotas: an experiment." *Public Opinion Quarterly* 43:477-96.
- Stinchcombe, Arthur L., Calvin Jones, and Paul B. Sheatsley
1981 "Nonresponse bias for attitude questions." *Public Opinion Quarterly* 45:359-75.
- Weaver, Charles N., Sandra L. Holmes, and Norval D. Glenn
1975 "Some characteristics of inaccessible respondents in a telephone survey." *Journal of Applied Psychology* 60:260-62.
- Wilcox, James B.
1977 "The interaction of refusal and not-at-home sources of nonresponse bias." *Journal of Marketing Research* 14:592-97.