Where Do We Go from Here? Nonresponse and Social Measurement

By
DOUGLAS S. MASSEY
and
ROGER TOURANGEAU

Surveys undergird government statistical systems and social scientific research throughout the world. Rates of nonresponse are rising in cross-sectional surveys (those conducted during a fixed period of time and not repeated). Although this trend worries those concerned with the validity of survey data, there is no necessary relationship between the rate of nonresponse and the degree of bias. A high rate of nonresponse merely creates the potential for bias, but the degree of bias depends on how factors promoting nonresponse are related to variables of interest. Nonresponse can be reduced by offering financial incentives to respondents and by careful design before entering the field, creating a trade-off between cost and potential bias. When bias is suspected, it can be countered by weighting individual cases by the inverse of their response propensity. Response propensities are typically estimated using a logistic regression equation to predict the dichotomous outcome of survey participation as a function of auxiliary variables. The Multi-level Integrated Database Approach employs multiple databases to collect as much information as possible about the target sample during the initial sampling stage and at all possible levels of aggregation to maximize the accuracy of estimated response propensities.

Keywords: surveys; nonresponse; bias; auxiliary data; adjustment; response rates; paradata

In their systematic assessment of the challenge that survey nonresponse poses to social statistics and social research, the authors of articles in this volume have abundantly highlighted the critical importance of surveys to

Douglas S. Massey is the Henry G. Bryant Professor of Sociology and Public Affairs at Princeton University and the president of the American Academy of Political and Social Science. He is a member of the National Academy of Sciences, the American Academy of Arts and Sciences, and the American Philosophical Society.

Roger Tourangeau is a vice president at Westat. Before coming to Westat, he was research professor at the University of Michigan's Survey Research Center and the director of the Joint Program in Survey Methodology at the University of Maryland. He has been a survey researcher for more than 30 years.

DOI: 10.1177/0002716212464191

contemporary society. Table 1 lists all the surveys mentioned by the authors in their contributions. It is by no means an exhaustive list of surveys in the world today, or even the most important surveys. It simply includes those surveys with which the authors were familiar and to which they turned for purposes of example or illustration. This admittedly ad hoc list includes twenty-two surveys that are sponsored by U.S. federal agencies along with fourteen others conducted in the United States mainly for research purposes, most with funding from federal granting agencies, such as the National Institutes of Health or the National Science Foundation. In addition, the authors mentioned some nineteen surveys fielded outside the United States, mostly, but not entirely, in Europe.

The sheer number and breadth of the surveys listed extemporaneously by the authors indicates the centrality that surveys have assumed in statistical reporting and social science research. Indeed, the twenty-two surveys conducted by federal agencies are likely responsible for the bulk of publicly reported statistics in the United States, from unemployment rates (Current Population Survey) to poverty rates and receipt of welfare (Survey of Income and Program Participation) to the basic social, demographic, and economic composition of the nation (American Community Survey, American Housing Survey). It is for this reason that we asserted in the Introduction to this volume that rising rates of nonresponse constitute a potential threat to national statistics that must be taken seriously.

The list of surveys in Table 1 not only suggests their importance to national statistics but underscores their importance to scientific research, not just in the social and behavioral sciences but in the health and biomedical sciences as well, and not only in the United States but throughout the world. America's General Social Survey and its counterpart across the Atlantic, the European Social Survey, provide the fundamental data to study shifts in social attitudes, political opinions, and societal values and their changing determinants over time. The Panel Study of Income Dynamics, the Health and Retirement Survey, and the Asset and Health Dynamics Survey undergird much economic research on lifetime processes of income, asset accumulation, and intergenerational transfers in the United States. The Survey of Health, Aging, and Retirement plays the same role in Europe, the Survey of Income and Labor Dynamics does so in Canada, as does the Household Income and Labor Dynamics Survey in Australia. In studies of how neighborhood circumstances condition individual well-being, the Los Angeles Family and Neighborhood Study, the geocoded version of the Panel Study of Income Dynamics, and the Project on Human Development in Chicago Neighborhoods provide critical data.

A survey of quantitative articles published in the typical social science journal reveals that a large fraction of the papers, often the majority, draw on survey data. Moreover, the U.S. statistical system as we know it today would cease to exist without surveys. As a result, increases in rates of nonresponse represent a potentially serious threat to the validity of scientific measurement and social research. In closing this volume, we therefore review what we have learned from the authors in this volume and attempt to summarize current knowledge on trends, causes, and consequences of survey nonresponse for statisticians, researchers, and the general public.

TABLE 1
Surveys Mentioned in Articles in This Volume of *The Annals*

Name of Survey	Sponsor of Survey
U.S. government surveys	
American Community Survey	Census Bureau
American Housing Survey	Census Bureau
Behavioral Risk Factor Surveillance Survey	Centers for Disease Control
Consumer Expenditure Survey	Bureau of Labor Statistics
Current Population Survey	Census Bureau
National Assessment of Educational Progress	National Center for Educ. Statistics
National Educational Longitudinal Survey	National Center for Educ. Statistics
National Health Interview Survey	Centers for Disease Control
National Health and Nutrition Examination Survey	Centers for Disease Control
National Household Education Survey	National Center for Educ. Statistics
National Immunization Survey	Centers for Disease Control
National Longitudinal Surveys	
National Longitudinal Survey of Youth 1997	Bureau of Labor Statistics
National Longitudinal Survey of Youth 1979	Bureau of Labor Statistics
National Longitudinal Survey of Women	Bureau of Labor Statistics
National Longitudinal Surveys of Men	Bureau of Labor Statistics
National Medical Care Expenditure Survey	Dept. of Health & Human Services
National Survey of Family Growth	Centers for Disease Control
National Survey on Drug Use and Health	Public Health Service
Residential Energy Consumption Survey	Energy Information Administration
Survey of Consumer Finances	Federal Reserve System
Survey of Income and Program Participation	Census Bureau
Survey of Early Child Care and Youth Development	National Institutes of Health–NICHD
Other U.S. surveys	
Asset and Health Dynamics Survey of the Oldest Old	University of Michigan
American National Election Studies	Stanford and University of Michigan
Fragile Families Survey	Princeton University
General Social Survey	National Opinion Research Center
Health and Retirement Survey	University of Michigan
Los Angeles Family and Neighborhood Study	RAND Corporation

(continued)

TABLE 1 (CONTINUED)

Name of Survey	Sponsor of Survey
Multi-City Study of Urban Inequality	Russell Sage Foundation
National Comorbidity Survey	Harvard University
National Longitudinal Survey of Adolescent Health	University of North Carolina
National Survey of American Life	University of Michigan
New Immigrant Survey	RAND Corporation
Panel Study of Income Dynamics	University of Michigan
Project on Human Dev. in Chicago Neighborhoods	Harvard University
Survey of Consumer Attitudes and Behavior	University of Michigan
Non-U.S. surveys	
British Birth Cohort Studies	Center for Longitudinal Studies
British Crime Survey	British Home Office
British Household Panel Study	University of Essex
British Household Survey	University of Essex
British National Survey of Sexual Attitudes and Lifestyles	National Center for Social Research
British Survey of Social Attitudes	UK Economic & Social Data Service
Dutch Integrated Survey on Household Living Conditions	Statistics Netherlands
Dutch Labor Force Survey	Statistics Netherlands
English Longitudinal Study of Aging	Institute for Fiscal Studies
European Social Survey	European Commission
Family Resources Survey	UK Dept. for Work and Pensions
German Socioeconomic Panel	German Inst. for Economic Research
Household, Income and Labor Dynamics Survey	University of Melbourne
International Assessment of Adult Competencies	OECD
Lifelong Learning Surveys	European Commission
National Health Interview Surveys in Europe	Eurostat
Survey of Health, Aging and Retirement in Europe	Max Planck Institute
Survey of Labor and Income Dynamics	Statistics Canada
Youth Cohort Study of England and Wales	Economic and Social Data Service

Dimensions of the Problem

As Brick and Williams note in their contribution, the consensus view among social scientists is that survey nonresponse rates are indeed rising to an alarming extent. As recently as 1979, for example, the Survey of Consumer Attitudes of the University of Michigan boasted a response rate above 70 percent, whereas the current rate is below 40 percent and falling (Curtin, Presser, and Singer 2005). In the four national surveys examined by Brick and Williams, the increase in nonresponse averaged roughly 0.5 percentage points per year since the mid-1990s, and this occurred despite greater efforts and more resources aimed at securing cooperation. Nonresponse rates above 30 percent are quite common in major national surveys today, and rates above 60 percent are no longer rare.

Although the decline in response rates has occurred for both telephone surveys and face-to-face interviews, the drop appears to be steeper for phone surveys (Atrostic et al. 2001; de Leeuw and de Heer 2002). Moreover, whatever forces are driving response rates downward, they are not peculiar to the United States. Similar, and in some cases worse, declines in response rates are apparent in Europe (de Leeuw and de Heer 2002; Stoop 2005). It is not clear, however, whether mail surveys are subject to the same trend. The 2010 census had a somewhat higher mail return rate than did the 2000 census, although this may simply reflect the greater use of advertising to encourage mail returns or the fact that all households received a short, rather than a long, form. Whereas mail surveys historically displayed lower response rates than telephone surveys, today they often evince higher response and are increasingly accepted as a viable alternative to phone surveys (Link et al. 2008).

The problem of nonresponse appears to be more serious for cross-sectional than longitudinal surveys. In the six panel surveys examined by Schoeni et al. in their article, none displayed consistent declines. In general, response rates either remained stable at high rates or actually increased. This pattern may reflect the fact that longitudinal surveys rely extensively on financial incentives. Schoeni et al. report that even small increases in payments to respondents yield significant improvements in response rates, and most panels have increased the size of their incentives over time. In the Panel Study of Income Dynamics, for example, the reward for participation rose from $20 in 1999 to $60 in 2009, in constant 1999 dollars. Although the length of the survey also increased from 35 to 75 minutes over the same period, the rise in incentives was faster, yielding an increase from 58 cents to 67 cents per minute of interview time. In addition, respondents were offered further incentives for assistance in locating other sample members.

Causes of Nonresponse

Survey researchers classify the reasons for nonresponse into three basic categories: noncontact, meaning that interviewers or screeners were unable to communicate with a targeted respondent; refusals, in which contact is established but the respondent declines to participate in the survey; and a residual "other" category (too infirm, inability to schedule a time, interviewer problems, etc.). Refusals generally constitute the largest category of nonresponse, followed by noncontact, and finally other reasons. Moreover, according to Brick and Williams, this ordering has changed little over time, though noncontact may be starting to rise a bit in the United States (Tourangeau 2004). Nonetheless, when they examined the National Health Interview Survey from 1997 through 2007, they found that 60 to 65 percent of nonresponse was attributable to refusals, 25 to 30 percent to noncontact, and 5 to 10 percent to other reasons. Likewise, on the National Household Education Survey the respective shares were 75 to 85 percent, 10 to 20 percent, and 5 to 10 percent. When nonrespondents were asked for the specific reason for declining to participate, the top three explanations were not interested, too busy, and takes too much time. Refusals likely account for the lion's share of nonresponse because most surveys make use of repeated callbacks until contact is achieved.

In terms of specific respondent characteristics that predict nonresponse, Schoeni et al. report that young people, minorities, males, renters, urban residents, single persons, the poor, and people with fewer social ties and attachments tend consistently to display lower probabilities of participation. The likelihood of responding also varies by context, with Brick and Williams finding that persons living in areas with greater concentrations of young children, higher crime rates, more family households, and shorter average commuting times generally display higher response rates. In the end, they conclude that the increase in nonresponse is driven by powerful generational changes, as younger cohorts increasingly display characteristics associated with low response likelihoods and are less comfortable with communication channels favored by survey researchers—telephone surveys and face-to-face interviews—than with emerging Internet and social networking technologies.

Effects of Nonresponse

Whatever its causes, whether nonresponse biases sample estimates depends not only on the response rate itself but also how different respondents are from nonrespondents. When the rate of nonresponse is low, the second factor is largely irrelevant. And even if differences between respondents and nonrespondents are great, if very few fail to respond, their absence from the sample will have little effect on estimates. As the rate of nonresponse rises, however, the potential for bias increases, and differences between those who do and do not respond become increasingly relevant in determining the degree and nature of any bias.

Historically, researchers have mostly relied on the size of the response rate itself as a rough gauge of the risk of bias, to the point where many scientific journals and statistical agencies required the reporting of response rates and sometimes specified a minimum acceptable value. More recently, however, a statistical formula derived by Bethlehem (2002) revealed that there is no necessary relationship between response rate and degree of bias. Consistent with this theoretical insight, in his article for this volume Peytchev reports that a meta-analysis of studies done on nonresponse bias yields scatterplots that show virtually no relationship between response rates and degree of bias across studies. As Olson explains in her contribution, the Bethlehem formula reveals that bias is inversely related to the response rate (the higher the response rate the lower the bias) but directly related to the covariance or degree of association between the response rate and the survey variable of interest. Even if nonresponse is high, no bias will result unless the likelihood of response is somehow related to the variable under consideration (see also Lessler and Kalsbeek 1992). That correlation is likely to vary from one survey variable to the next, so there will be considerable variation across estimates within the same survey in the level of nonresponse error. In general, such a relationship will exist if the variable is causally related to survey participation or if a third variable is a common cause of both the variable and participation.

The Bethlehem equation has several important implications. First, it is impossible to know a priori whether a high rate of nonresponse will produce bias. Second, to the extent that bias results from nonresponse, it will vary from item to item, and in fact, most of the variation in bias is not across surveys but between items within surveys. Last, raising response rates does not guarantee a reduction in bias, and it is even possible to increase the degree of bias by reducing the rate of nonresponse. Given the emphasis historically placed on achieving high response rates, the latter may seem counterintuitive, but in their contribution Singer and Ye offer a simple example using financial incentives. They point out that if the incentives affect all sample members equally, they will have no effect on bias; but if they bring into the sample more of those who are already overrepresented, they will increase bias. Only if incentives have the effect of bringing into the sample those who are underrepresented will they have the desired effect of reducing bias and improving the accuracy of estimates.

Drawing on the foregoing insights, investigators have sought to develop better indicators of potential bias that go beyond simple nonresponse rates. In her article, Kreuter describes two such indicators: the R-indicator, which captures imbalances in response propensities between respondents and nonrespondents and measures the similarity between the original sample and the sample respondents, and the FMI (fraction of missing information) indicator, which measures uncertainty about values imputed for nonresponding cases. Both of these indicators require auxiliary information on both the respondents and nonrespondents from outside the survey. To assess bias, therefore, one needs additional information from sources such as sampling frames, contact forms, process logs, interviewer observations, administrative files, or geocoded data files.

Limiting Nonresponse

The most effective way of minimizing bias is to maximize the response rate to a survey through improvements in design and implementation. As Schoeni and his colleagues point out in their article, these improvements begin before any contact with respondents is attempted by making sure that address lists are accurate and up to date using the U.S. Postal Service's National Change-of-Address System and, for federal surveys at least, relying on the Census Bureau's Master Address File. In panel surveys, mailings sent to all respondents should include a prepaid postcard with a request for address verification and correction. Special respondent websites are increasingly being used to share information of value to survey participants, maintain their interest, and permit online updating of contact information.

Schoeni et al. report on strategies that have proved effective in boosting participation during fieldwork itself. The most important is establishing contact with targeted respondents prior to attempting the survey, normally through a letter or email that, to the extent possible, is tailored to the interests and circumstances of the respondent, explaining why the survey is of interest or relevance to that person or addressing the stated reasons for an earlier refusal. Most surveys require interviewers to attempt repeated callbacks before giving up on a targeted respondent. In panel studies, interviewers also collect names and contact information of friends and relatives of the respondent, people he or she is likely to stay in contact with over time. Many surveys increasingly employ directory assistance and other online databases to update address lists and refresh the sampling frame. In her article, Kreuter advocates the use of management techniques developed in the quality control literature—such as flow diagrams, scatterplots, Pareto charts, and control charts—to achieve greater efficiency and higher rates of participation.

Researchers can also take steps to maximize response in the design of the questionnaire itself. In general, anything that can be done to reduce the burden on respondents can be expected to improve participation. In addition to the obvious strategy of limiting the number of questions asked, whenever possible investigators should make use of administrative data to fill in certain variable fields prior to the interview and offer respondents a choice of survey mode—telephone interview, personal interview, online survey, or mail survey. The number of items on a survey can also be limited through the use of "matrix" sampling whereby different sets of questions are created and administered to random subsets of respondents so that information is gathered on a larger number of items than would be included in any single respondent's survey. In the absence of self-reported information, proxy reports by knowledgeable others may also be used. In panel studies, the burden may be reduced by reducing the frequency of interviews by spacing out the survey waves. Panel studies also offer the opportunity to attempt reinterviews with respondents who dropped out of earlier waves.

As a result of increased efforts to minimize nonresponse, survey costs have risen even as response rates have fallen. For example, the cost per household of the decennial census rose from about \$77 in 2000 to about \$111 in 2010; given that these figures are in constant dollars, they represent a cost increase of nearly 50 percent. The Census Bureau does not pay respondents, of course, since participation in the census is required by law. Other surveys, however, increasingly rely on financial incentives to induce participation; and in their contribution Singer and Ye draw on the leverage-salience theory of respondent decision-making to argue that payments are indeed quite useful in recruiting into the sample people who otherwise would not be motivated to respond, in particular those who have little interest in the topic, those who lack altruistic motives for responding, and those with alternative obligations and competing uses for their time.

Their careful review of the literature clearly indicates that incentives increase response rates across all modes of implementation (telephone, face-to-face, mail, and Internet), that monetary incentives work better than gifts to promote participation, and that prepaid incentives increase the likelihood of participation more than promised incentives or lotteries. However, there is no good evidence about how large an incentive should be, only that as payments rise the effect on response rates declines. Although studies are limited, they generally indicate that incentives have few effects on the quality of responses or the composition of the sample.

In sum, it is clear that response rates can be increased by spending more money, either indirectly by improving the design and implementation of the survey or directly by incentivizing respondents with monetary payments. In this sense, the issue of nonresponse sometimes boils down to a trade-off between cost and bias and ultimately depends on how much researchers are willing or able to spend to minimize the potential for bias inherent in a high rate of nonresponse. Striking this balance is a guessing game, however, for, absent auxiliary information, it is impossible to know the degree to which nonresponse will, in fact, bias estimates of means, variances, covariances, or other parameters such as regression coefficients. While access to such additional information allows an assessment of potential bias, the assessment must be carried out on an item-by-item basis. Moreover, auxiliary data also open up the possibility of adjusting sample estimates to correct for nonresponse bias, which may be a less expensive alternative than boosting response rates.

Adjusting for Nonresponse

The simplest and most common means of adjusting for bias created by nonresponse is to weight individual cases by the inverse of the response propensity—that is, the likelihood of survey participation. In this way, cases from subgroups that are less likely to be included in the sample are given relatively more weight in computing the estimated parameter, and the inverse of the response propensity is known as the response propensity weight. The response propensity for any sampled element depends on its characteristics, which of course are not observed for nonrespondents. Therefore, response propensities must be estimated, typically using a logistic regression equation to predict the dichotomous outcome of survey participation as a function of auxiliary variables. In her contribution here, Olson lists the characteristics of the ideal auxiliary variables for use in estimating response propensities: nonmissing values are available for respondents and nonrespondents, values are measured completely and without error for all cases, and the variables are strongly associated with survey variables of interest as well as the response propensity. Even when such data are available, adjustment is always complicated by the fact that investigators typically have many variables of interest, each of which may be biased to a different degree (or not at all) by nonresponse, and by the fact that both item and unit nonresponse are relevant.

In his appraisal of the suitability of administrative data for nonresponse adjustment, Czajka notes that several sources provide excellent coverage of the U.S. population: the IRS, the Social Security Administration (SSA), and files compiled by the Census Bureau. The coverage rate for tax return files and associated information documents (e.g., W-2s and 1099 Forms) from the IRS are estimated at 95 to 96 percent, while the SSA's application files provide basic data on age, race, gender, and Hispanic origin. The Census Bureau's Statistical Administrative Records System seeks to create a census-style database from a variety of federal administration systems, including the SSA and the IRS files as well as data from Medicare, the Department of Housing and Urban Development, the Selective Service System, and the Indian Health Service.

The most important limitation on the potential use of these administrative systems as data for nonresponse adjustment is legal: federal legislation prohibits the public release of the data (Title 13 of the U.S. Code for the Census Bureau and Title 26 for the IRS). Thus, they are available only for adjustment of data collected by the federal government and can only be done in-house by the Census Bureau or some other authorized agency. Even then, files from the IRS, the SSA, and the Census Bureau have weaknesses; coverage is not perfect, as 4 to 5 percent are typically missing; administrative data contain significant measurement error, especially on items that are not central to the agency's mission; and information is generally not available in a timely fashion but at a lag (e.g., tax returns are filed only once a year).

Given restricted access to federal administrative data, survey researchers have turned to the Multi-level Integrated Database Approach (MIDA). As Smith and Kim explain in their article, MIDA employs multiple databases to collect as much information as possible about the target sample during the initial sampling stage at all levels of aggregation: individual, household, block, tract, zip code, county, and even state. The first step is to extract all publicly available information at several levels of aggregation and link it to units on the sampling frame. The Census Bureau, for example, routinely releases public use data on block groups, census tracts, counties, and states. The second step is to enhance the resulting file by linking units on the sampling frame to other available sources of data, such as phone directories, credit reports, property records, and voter registration lists as well as information from private vendors such as the Claritas Corporation. The third step is to use the combined information from steps one and two to make matches to units that were not possible with information from a single source. For example, a number from a telephone directory can lead to households in databases where a simple address did not enable a match.

The final step is to process, clean, and update the now large amount of paradata on each case and to make the resulting file available routinely to all survey users. If MIDA were to become the norm in the survey research industry, it might produce multiple benefits. It would, of course, facilitate measurement and adjustment for nonresponse. Having access to comprehensive multilevel data would enable researchers to test representativeness of the respondents across a wide range of variables and facilitate the specification of models to estimate response propensities for use in weighting whenever bias is detected. As Smith and Kim point out, MIDA would also make data collection more efficient and effective, improve interview validation, facilitate the use of GIS technology, and create a much richer database for substantive analysis. The principal concern is the confidentiality of respondents, which would need to be assured through protocols similar to those currently used for the release of geocoded datasets such as the Panel Study of Income Dynamics or the Adolescent Health Survey. An additional worry is the accuracy of the data, which may be inaccurate at its source or simply not characterize the sample person because the data are so highly aggregated. It is important to question whether MIDA or auxiliary data more generally will prove to be valuable tools in countering the effects of survey nonresponse.

Best Practices for Managing Nonresponse

When all is said and done, the good news is that methodological studies generally show that even relatively large changes in response rates do not have much effect on final sample estimates (Keeter et al. 2006, 2000), again demonstrating that nonresponse bias does not depend in any simple way on the nonresponse rate. Despite declining response rates in telephone surveys, election polls still seem to yield accurate predictions of the outcomes of elections. As Schoeni et al. report in their contribution, even cumulative nonresponse across waves of the Panel Study of Income Dynamics failed to have a significant effect on estimates of income, health, consumer spending, or wealth. Although nonresponse biases can at times be large, their magnitude does not vary closely with nonresponse rates (Groves and Peytcheva 2008; Tourangeau, Groves, and Redline 2010).

Still, almost all we know about nonresponse bias involves bias in means and proportions; and while response rates may be flawed as measures of nonresponse bias, they are still widely seen as important indicators of overall survey quality. In addition, high response rates impose an upper bound on the possible impact of nonresponse, most clearly for estimated means and proportions. Thus, for the foreseeable future, survey researchers are likely to use various techniques to boost response rates (or at least stem their decline). The proven methods for increasing response rates include

—sending sample members an advance letter describing the purpose of the survey and explaining why it is important,
—making multiple callbacks or other follow-up efforts to contact sample members,

—carrying out persuasion (“refusal conversion” in the parlance of survey research) attempts with reluctant members of the sample, and
—offering small prepaid incentives to encourage participation.

The effectiveness of most of these methods has been clear for at least 30 years (Heberlein and Baumgartner 1978), and recent evidence indicates that they continue to work. Financial incentives continue to be effective, especially prepaid cash incentives, for both mail and interviewer-mediated surveys. Advance letters are likely to be an increasingly important tool for survey researchers as it becomes harder to reach people by telephone or by knocking on doors. Attrition rates in longitudinal surveys do not seem to have dropped in the same way as response rates to cross-sectional surveys for this reason. Panel surveys can and often do deploy an extensive set of tools for minimizing attrition over the life of a panel; and to date, these efforts often seem encouragingly successful.

An additional tactic used in many surveys involves offering sample members more than one method to respond. Some surveys offer respondents a choice of methods for completing the survey from the outset. Sample members may, for example, be given the option of completing a mail questionnaire or of responding to the survey online. Other surveys use a sequence of modes, beginning with the mail, say, and then following up with mail nonrespondents via telephone or face-to-face contacts. The American Community Survey (ACS) follows this sequential strategy, starting with mail and ending with face-to-face interviews.

Advance letters, multiple callbacks, refusal conversion, incentives, and multiple modes of data collection are all attempts to raise response rates. Other methods for coping with nonresponse represent attempts to reduce nonresponse bias. One method that has been used (although infrequently) for several decades is two-phase sampling. With a two-phase sample, a subsample of the initial cases is selected during the field period, and only those selected for the subsample are retained for further follow-up efforts. The subsampling allows a greater concentration of effort and resources on the remaining sample members; they may be offered larger incentives than the initial sample was, approached by more persuasive interviewers, asked to complete a shorter questionnaire, or approached with some combination of these enticements to get them to respond.

Depending on how effective these second-phase efforts are, the researchers may estimate the level of bias using the data from the respondents to the second phase effort (comparing the first phase respondents with the second phase respondents), or they may just combine the data collected during the two phases (giving greater weight to the second phase respondents to compensate for the additional stage of sampling that these cases underwent). The ACS uses a two-phase sampling strategy, selecting only one in three of the remaining nonrespondents for the final face-to-face follow-up efforts.

As noted earlier, nonresponse introduces bias into means and proportions only when the response propensities of the sample members are related to their values on survey variables. A consequence of this fact is that, if everyone has the same response propensities, there cannot be any relation between the response propensities and the survey variables, and bias is no longer an issue. When the overall response propensities are equal, nonresponse simply functions as an extra stage of random sampling. Thus, one goal of the field effort is (or should be) to equalize response propensities across the sample members. Two-phase sampling is one means of accomplishing this (or at least of reducing the variability in response propensities). The change in protocol during the second phase of the field period raises the response propensities among the initial nonrespondents selected for further follow up; these initial nonrespondents presumably had low response propensities during the first phase of data collection.

"Responsive design" (Groves and Heeringa 2006) is a more general strategy for retooling the field effort to raise the response propensities among those with low initial propensities. Survey field managers have long attempted to tailor the approach taken with specific members of the sample to reduce the chances that they would become nonrespondents. Responsive design attempts to use paradata (process data on the prior attempts to complete the interview) in place of field manager judgment to tailor the follow-up efforts to specific cases, with the goal of reducing variation across individual cases or subgroups in overall response propensities. Furthermore, response propensities estimated through statistical models can be used in place of interviewer judgment to direct how much effort a given case receives.

After the data have been collected, many surveys adjust the weights for the respondents to compensate for any remaining differences in response propensities across the responding cases. Various weighting schemes can be used (for example, cell-based weighting adjustments or adjustments based on logistic regression models), but almost all of them are based on the principle of increasing the weight given to a respondent by the inverse of his or her estimated response propensity (Kalton and Flores-Cervantes 2003). Such nonresponse adjustments can reduce nonresponse bias, but, as Little and Vartivarian (2005) demonstrate, they do not always work; the adjustments can also increase the variance in survey estimates. Still, most survey researchers regard weighting adjustments as a useful, even essential, step to take in addressing nonresponse.

All these tactics aim either to prevent nonresponse bias by increasing the response rates (or equalizing the response propensities) or to reduce the effects of nonresponse on the estimates by adjusting the case weights. As a final move, researchers may attempt to estimate the level of nonresponse bias that remains. The usual strategies are to compare the respondents and nonrespondents using frame data or data from administrative records; in addition, in some surveys, there are data from preliminary screening interviews that can be used to compare respondents and nonrespondents (Groves and Peytcheva 2008). A final option is to compare the characteristics of the respondents to some external population benchmark, such as figures from the ACS. The Office of Management and Budget's Standards and Guidelines for Statistical Surveys, issued in 2006, recommends that investigators carry out studies like these to estimate the level of nonresponse bias whenever the response rate for a survey falls below 80 percent.[1] Clearly, these efforts are useful whenever nonresponse bias is a threat to a survey's conclusions.

Conclusion

Perhaps the greatest threat to the future of survey data is ultimately the lack of public recognition for the importance of statistical and scientific surveys in the world today. In our introduction to this volume, we suggested that it might be time for an industry-wide effort to improve the image of survey research and to differentiate legitimate social scientific surveys from the onslaught of unwanted solicitations that masquerade as surveys. This could be achieved through a comprehensive advertising, education, and outreach campaign aimed at the general public through the mass media and sponsored by survey research firms, legitimate polling agencies, and the federal government.

Protecting the integrity of survey research also requires better efforts to educate our political leaders. We appear to be in the midst of a remarkable wave of antiscientific, anti-intellectual sentiment in the United States, one in which data, statistics, and the instruments that produce them are viewed with suspicion and sometimes outright hostility in many quarters of government and the public. The most recent manifestations of such unproductive sentiments were proposals by some members of Congress to repeal mandatory participation in the ACS, the instrument that has replaced the old long form of the decennial census, and the introduction of an amendment (H.R. 5326) to prohibit the use of federal funds to conduct the ACS altogether, both actions that would cripple the federal statistical system.

Somehow our public leaders must be better educated to appreciate and understand the critical importance of social surveys in contemporary society, not only because they provide basic information to inform citizen voters in a democratic society but also because they generate key inputs needed by American businesses to compete effectively in a global, knowledge-based economy and because they give policy-makers the facts they need to make competent, evidence-based decisions. We offer this volume as a first step in a broader and more concerted effort to educate policy-makers and the public about the key importance of social surveys to the healthy functioning of postindustrial society.

Note

1. See http://www.whitehouse.gov/sites/default/files/omb/assets/omb/inforeg/statpolicy/standards_stat_surveys.pdf.

References

Atrostic, B. K., Nancy Bates, Geraldine Burt, and Adriana Silberstein. 2001. Nonresponse in U.S. government household surveys: Consistent measures, recent trends, and new insights. *Journal of Official Statistics* 17 (2): 209–26.

Bethlehem, Jelke G. 2002. Weighting nonresponse adjustments based on auxiliary information. In *Survey nonresponse*, eds. Robert M. Groves, Don A. Dillman, John L. Eltinge, and Roderick J. A. Little, 275–88. New York, NY: Wiley.

Curtin, Richard, Stanley Presser, and Eleanor Singer. 2005. Changes in telephone survey nonresponse over the past quarter century. *Public Opinion Quarterly* 69 (1): 87–98.

de Leeuw, Edith, and Wim de Heer. 2002. Trends in household survey nonresponse: A longitudinal and international perspective. In *Survey nonresponse*, eds. Robert M. Groves, Don A. Dillman, John L. Eltinge, and Roderick J. A. Little, 41–54. New York, NY: John Wiley.

Groves, Robert M., and Steven G. Heeringa. 2006. Responsive design for household surveys: Tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society, Series A* 169:439–53.

Groves, Robert M., and Emilia Peytcheva. 2008. The impact of nonresponse rates on nonresponse bias: A meta-analysis. *Public Opinion Quarterly* 72 (2): 167–89.

Heberlein, Thomas A., and Robert Baumgartner. 1978. Factors affecting response rates to mailed questionnaires: A quantitative analysis of the published literature. *American Sociological Review* 43 (4): 447–62.

Kalton, Graham, and Ismael Flores-Cervantes. 2003. Weighting methods. *Journal of Official Statistics* 19 (2): 81–97.

Keeter, Scott, Courtney Kennedy, Michael Dimock, Jonathan Best, and Peyton Craighill. 2006. Gauging the impact of growing nonresponse on estimates from a national RDD telephone survey. *Public Opinion Quarterly* 70 (5): 759–79.

Keeter, Scott, Caroline Miller, Andrew Kohut, Robert Groves, and Stanley Presser. 2000. Consequences of reducing nonresponse in a national telephone survey. *Public Opinion Quarterly* 64 (2): 125–48.

Lessler, Judith, and William Kalsbeek. 1992. *Nonsampling errors in surveys*. New York, NY: John Wiley.

Link, Michael W., Michael P. Battaglia, Martin R. Frankel, Larry Osborn, and Ali H. Mokdad. 2008. A comparison of address-based sampling (ABS) versus random digit dialing (RDD) for general population surveys. *Public Opinion Quarterly* 72 (1): 6–27.

Little, Roderick J. A., and Sonya Vartivarian. 2005. Does weighting for nonresponse increase the variance of survey means? *Survey Methodology* 31 (2): 161–68.

Stoop, Ineke A. L. 2005. *The hunt for the last respondent: Nonresponse in sample surveys*. Amsterdam: Aksant Academic.

Tourangeau, Roger. 2004. Survey research and societal change. *Annual Review of Psychology* 55:775–801.

Tourangeau, Roger, Robert M. Groves, and Cleo D. Redline. 2010. Sensitive topics and reluctant respondents: Demonstrating a link between nonresponse bias and measurement error. *Public Opinion Quarterly* 74:413–32.