Consequences of Survey Nonresponse

By
ANDY PEYTCHEV

Nonresponse is a prominent problem in sample surveys. At face value, it reduces the trust in survey estimates. Nonresponse undermines the probability-based inferential mechanism and introduces the potential for nonresponse bias. In addition, there are other important consequences. The effort to limit increasing nonresponse has led to higher survey costs—allocation of greater resources to measure and reduce nonresponse. Nonresponse has also led to greater survey complexity in terms of design, implementation, and processing of survey data, such as the use of multiphase and responsive designs. The use of mixed-mode and multiframe designs to address nonresponse increases complexity but also introduces other sources of error. Surveys have to rely to a greater extent on statistical adjustments and auxiliary data. This article describes the major consequences of survey nonresponse, with particular attention to recent years.

Keywords: response rates; nonresponse bias; survey cost; survey design

Response rates in household surveys have been declining in the United States for some time (Curtin, Presser, and Singer 2005; Groves and Couper 1998; Stussman, Dahlhamer, and Simile 2005) as they have been in other developed countries (de Leeuw and de Heer 2002). Current unit nonresponse rates and their seemingly continual ascent have led to multiple consequences for statistical surveys. These consequences include direct changes, such as differing compositions of nonrespondents and increased potential for nonresponse bias, along with more indirect changes, such as

Andy Peytchev is senior survey methodologist at RTI International and instructor at the Odum Institute, University of North Carolina at Chapel Hill. His primary research interests include the measurement of, reduction of, and adjustment for survey nonresponse and the interplay among different sources of survey error.

NOTE: The author thanks Paul Biemer, Emilia Peytcheva, Roger Tourangeau, and Douglas Massey for providing comments on earlier drafts.

DOI: 10.1177/0002716212461748

increased survey costs and greater complexity of survey designs. This article offers a summary of the principal consequences of survey nonresponse, devoting greatest attention to nonresponse bias in estimates.

Response Rates and Nonresponse Bias

Unbiased inference from probability-based surveys relies on the collection of data from all sample members—in other words, a response rate of 100 percent. Response rates reflect the degree to which this goal of preserving the respondents' inclusion probabilities is achieved. The simplicity of the response rate, often a single value for the entire survey, has inarguably led to its widespread adoption as a summary measure of a survey's representativeness. Furthermore, as other major sources of survey error, such as measurement error, are much more complex to measure and quantify using standardized statistics, response rates may be receiving a disproportionate amount of attention.

The adoption of the response rate, along with the subsequent decline of response rates in many cross-sectional surveys, has affected official statistical tabulations and scientific research alike. For example, because of the increased potential for nonresponse bias, the Office of Management and Budget, which regulates federally funded surveys, has prohibited in some instances the labeling of survey-based statistics as prevalence rate estimates. In the research community, publications such as the *Journal of the American Medical Association* reject manuscripts based on low response rates, even while allowing research that is not based on probability sampling. A recent survey of social science journal editors found that almost all editors consider the study's response rate in making a decision on a manuscript (Carley-Baxter et al. 2009).

Using the response rate as the only measure of the representativeness of a survey is erroneous because it is nonresponse bias that is feared, not nonresponse itself. A shift from a response rate of 95 percent to 90 percent, for example, introduces a small potential for nonresponse bias—the additional 5 percent of nonrespondents need to be extremely different from the respondents in the first case to introduce substantial nonresponse bias. The potential is not so small when the response rate is 50 percent (i.e., one-half of eligible sample members are nonrespondents), however. In this case, relatively small differences between respondents and nonrespondents may yield significant biases. Moreover, in two surveys, one with a 50 percent and another with a 40 percent response rate, the response rate merely indicates that there is potential for nonresponse bias in both surveys but not which survey suffers from greater nonresponse biases. Nonetheless a shift from a relatively low likelihood of nonresponse bias (for example, a 95 percent response rate) to a much higher potential for nonresponse bias has occurred during recent years.

Unit nonresponse creates the potential for nonresponse bias in survey estimates. For purposes of this discussion, nonresponse bias is defined for an estimate of the mean. A commonly employed formula presents the expected value of the difference between the respondent and the full sample means; that is, nonresponse bias in the mean prior to any nonresponse adjustments, as the expectation of the product of the nonresponse rate $\left(\frac{m_s}{n_s}\right)$ and the difference between the respondent and nonrespondent means $(\bar{y}_r - \bar{y}_m)$, as presented in Groves (1989):

$$Bias(\bar{y}_r) = E\left[\frac{m_s}{n_s}(\bar{y}_r - \bar{y}_m)\right]. \tag{1}$$

This representation of the bias clearly shows how the nonresponse rate relates to nonresponse bias in a survey estimate. If a 99 percent response rate is achieved, any difference between respondents and nonrespondents will be multiplied by .01, making the resulting bias inconsequential. As the response rate declines, however, the differences between respondents and nonrespondents begin to result in nonresponse bias. The formula can thus be interpreted that increasing the response rate will invariably decrease nonresponse bias. That is, the two terms—the nonresponse rate and the difference between respondents and nonrespondents—are often seen as orthogonal.

A potentially less intuitive formula that underscores the likely dependency of nonresponse rates and differences between respondents and nonrespondents has been proposed by Bethlehem (2002). Nonresponse bias in the mean is defined as the ratio of the covariance between the response propensity (ρ) and the survey variable (y), and the mean response propensity ($\bar{\rho}$). The response propensity is the likelihood of participating in the survey, conditional on the applied survey protocol. Thus,

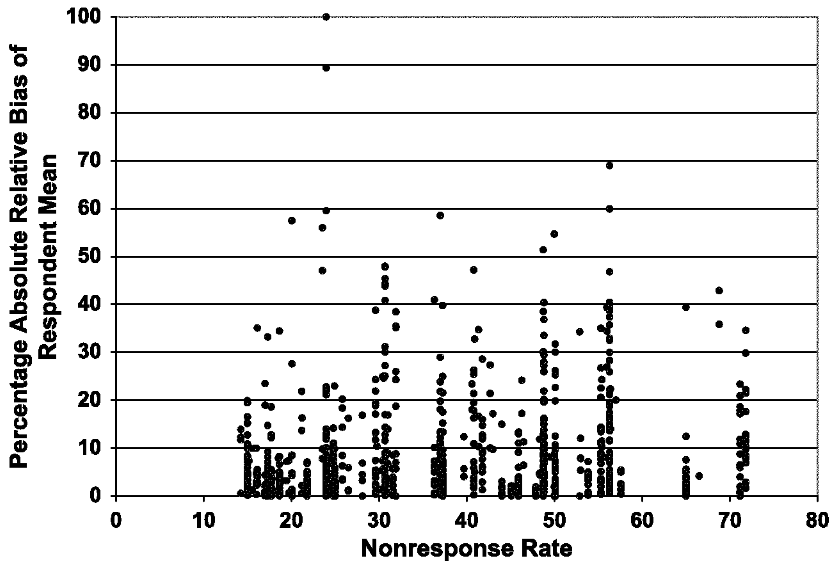
$$Bias(\bar{y}_r) \approx \frac{\sigma_{\rho,y}}{\bar{\rho}}. \tag{2}$$

In this formulation, it is more apparent that the covariance in the numerator (which includes the response propensity) is not independent of the response rate (mean response propensity) in the denominator. Thus, increasing the response rate may not necessarily decrease nonresponse bias, and a lower response rate may not necessarily increase nonresponse bias.

Groves (2006) and Groves and Peytcheva (2008) conducted a meta-analysis of nonresponse bias studies, examining the association between response rates and nonresponse bias in survey estimates. Figure 1 presents a scatter plot from their findings, showing virtually no relationship between rates and bias across studies. More important, most of the variation in nonresponse bias in survey estimates is within studies. To reiterate two key conclusions from their work, response rates are a poor indicator of nonresponse bias, and nonresponse bias is estimate- rather than survey-specific.

The lack of association between nonresponse rates and nonresponse bias among studies does not necessarily imply that increasing the response rate within a study will not decrease nonresponse bias in survey estimates. What it

FIGURE 1
Percentage Absolute Relative Nonresponse Bias of 959 Respondent Means, by Nonresponse Rate of the 59 Surveys in Which They Were Estimated

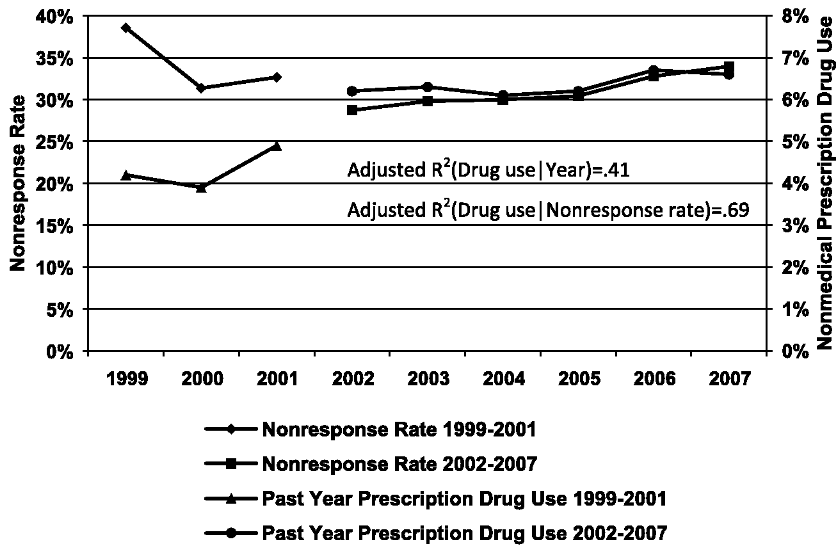


SOURCE: Groves and Peytcheva (2008).

does mean is that comparison of studies based on response rates may not be informative of differences in nonresponse bias. Figure 2 shows the relationship between a key survey estimate in the National Survey on Drug Use and Health (NSDUH)—the nonmedical use of prescription psychotherapeutics during the year prior to the interview—and the weighted survey nonresponse rate. Note that the relationship between the two changes in 2002. In that year, the NSDUH introduced an incentive for participation, and its name was changed from the National Household Survey on Drug Abuse to the NSDUH. These protocol changes had the potential to affect both nonresponse and nonresponse bias (omission of "drug abuse" from the survey's name and the use of a monetary incentive led to higher estimates of drug use, which is generally underestimated).

Figure 2 shows how closely the pattern in estimated prescription drug abuse follows the pattern of the nonresponse rate. When one regresses the drug abuse rate on the nonresponse rate for 2002 through 2007, approximately 70 percent of the variation in drug abuse is explained (R^2 = .69). If year is used instead of the nonresponse rate, however, the proportion of variation drops to about 40 percent (R^2 = .41). Furthermore, while year is a significant predictor of drug abuse, it is no longer significant after controlling for the nonresponse rate,

FIGURE 2
Nonmedical Use of Prescription Psychotherapeutics in Past Year and Nonresponse Rates in NHSDA and NSDUH, 1999–2007



indicating that the trend in increasing drug abuse may be merely an artifact of nonresponse. This observation provides a strong suggestion that there is a relationship between nonresponse bias and response rate within a study and that this relationship can be tempered by changes to survey design that reduce fluctuation in response rates.

In fact, increasing the response rate for a particular survey can also lead to increased nonresponse bias in survey estimates, as was found in the 1998 Dutch Integrated Survey on Household Living Conditions when the response rate was increased from 47 percent to 60 percent over a month (Schouten and Cobben 2007). Indeed, when a regression line is fit to the substantive (nondemographic) estimates in Figure 1, the coefficient is actually negative (Peytcheva and Groves 2009).

How a response rate is achieved can influence how different the respondents are from the remaining nonrespondents and determine the level of nonresponse bias in a survey estimate. In a split ballot experiment, Merkle and Edelman (2009) demonstrated that a survey design producing a higher response rate can lead to greater nonresponse bias in the key survey estimate. When the features in the survey protocol that help to elicit higher survey participation appeal to groups that are already overrepresented in the respondent pool, the difference between respondents and nonrespondents can increase, offsetting any benefit from nonresponse reduction and resulting in higher nonresponse bias. In Merkle and Edelman's study, providing pens (among other similar design features) during an election exit poll led to even higher participation among Democrats, who already participated at higher rates than Republicans.

There seems to be a disjunction between theory and practice in addressing nonresponse. In practice, response rates are often increased by directing more effort, such as more call attempts in telephone surveys and more in-person visits in face-to-face surveys. Evaluations of these greater efforts have shown an inability to obtain interviews from sample members who are different from prior respondents (Curtin, Presser, and Singer 2000; Keeter et al. 2000). Yet from a theoretical perspective, the survey protocol needs to be changed in such a manner that the response propensities (likelihood of participation) for the remaining nonrespondents are altered, particularly for sample members who are different from the current respondents. Examples of such changes might be reducing the size of the request or providing incentives to motivate those who may be less interested in the survey topic and thus less motivated to complete the survey. Experiments have demonstrated that while individuals who are less interested or have lower topic involvement participate at lower rates, they are also influenced to a greater extent by incentives, suggesting that the provision of incentives can reduce nonresponse bias (Groves et al. 2006; Groves, Presser, and Dipko 2004; Groves, Singer, and Corning 2000).

Making additional calls under a different protocol that reduces the size of the request and increases incentive amounts has been shown to obtain participation from sample members who are different on key survey estimates, compared to making more calls under the same protocol (Peytchev, Baxter, and Carley-Baxter 2009). While the decline in response rates has been taxing for surveys, in the NSDUH the inclusion of incentives has been shown to decrease overall costs while increasing response rates (Kennet et al. 2005). The implications of nonresponse for survey costs are discussed in more detail below, as are other complexities imposed on studies such as the imputation of items in shortened questionnaires and the need for multiple protocols. In sum, the consequences of nonresponse on nonresponse bias are dependent on the survey design.

Response rates are clearly deficient in reflecting survey representativeness. Although they do carry some information, they can be misleading when used in isolation. Alternative or additional measures of survey representativeness are needed. This need is a direct consequence of rising nonresponse rates—no such need for additional or alternative measures arose when survey data users were content with response rates and were less concerned that nonresponse affected the properties of survey estimates.

A summit was convened in 2007 to compile and organize existing measures of survey representativeness (Groves et al. 2008). These measures vary on several dimensions, including whether they are survey-specific or estimate-specific, the level of complexity they incorporate, the model assumptions they make, and the sources of information they utilize. Although a desirable outcome in the future is to see multiple measures of survey representativeness that help to inform about the potential for nonresponse bias in addition to response rates, they will likely require greater sophistication on the part of those who prepare the data for dissemination and greater knowledge by consumers of survey data.

Nonresponse Bias for Different Estimators

The real threat of nonresponse is bias in survey estimates—commonly, in estimated means, percentages, and totals. Nonresponse, however, can also lead to bias in other important statistics, such as estimates of variances, associations (including regression coefficients), and change in parameter values over time, all of which are reviewed below.

Means and proportions

Undoubtedly, bias in estimates of means and proportions has received the greatest attention, as these statistics are of prime interest in government-sponsored surveys. The threat of nonresponse bias in descriptive statistics is certainly not a straw man in surveys. For example, one evaluation found that nonrespondents to a component of the National Health and Nutrition Examination Survey III were 59 percent more likely to report being in poor or fair health than the respondents (Khare et al. 1994). Similarly, the Belgian National Health Interview Survey (response rate of 61.4 percent) obtained a 19 percent lower estimate for reporting poor health compared to the Belgian census (response rate of 96.5 percent; see Lorant et al. 2007). Furthermore, nonresponse biases vary across subgroups (see for example Groves and Peytcheva 2008) in ways that can lead to poor policy decisions, such as when nonresponse bias is greater in particular geographic areas or in particular minority groups.

Nonresponse may lead to substantial bias in some survey estimates and to very minimal bias in others. While the formulas for nonresponse bias are estimate-specific, and theories for nonresponse bias reflect causal relationships particular to a variable, the empirical evidence from the recent meta-analyses (Groves 2006; Groves and Peytcheva 2008) underscores this point; nonresponse bias varied more across estimates within a study than between studies, even when different methods to measure bias were used across studies.

Variance

Nonresponse bias can also lead to bias in variance estimates. To the extent that respondents and nonrespondents have different means on a variable, a variance estimate based solely on respondents will omit sample members at one end of the distribution and be downward biased. A second source of bias in variance estimates is that nonrespondents may vary to a different degree in values of a variable compared to respondents—a condition that may lead to either underestimation or overestimation. It is possible, for example, for a variance estimate to be upward biased if the variance for respondents is substantially larger than the variance for nonrespondents.

Finally, even when nonresponse bias is not present in means or proportions, estimates of variances can still be biased. Although the use of replicate weights and replicate methods of variance estimation attempts to reflect the uncertainty in estimates due to nonresponse (see Wolter 2007, 185–87), most surveys offer a single weight that omits uncertainty in the nonresponse weights (the certainty gained through the use of known population totals in post-stratification). This area has received little attention in the research literature, yet it can lead to underestimated variances; the auxiliary data used in adjustments could have not only sampling variability themselves, but could be subject to measurement error.

In sum, although variance estimates can be biased in either direction, they are most likely to be underestimated in surveys. Owing to nonresponse, therefore, descriptive statistics may have unjustly small confidence intervals. The degree of the bias in variance estimates will depend on the response rate, the amount and properties of auxiliary data, the remaining nonresponse bias in adjusted estimates, differences in variability among nonrespondents, and the weighting and variance estimation methods employed.

Associations

The effect of nonresponse on associations is often overlooked, yet associations are invariably estimated in surveys and are sometimes of more interest than overall means and proportions. Nonresponse can affect not only estimates of bivariate relationships, but also multivariate relationships, such as coefficients in regression models. Evidence in the research literature is lacking. Lepkowski and Couper (2002) found bias in associations to be relatively small in magnitude across selected estimates from a face-to-face survey, while Peytchev, Carley-Baxter, and Black (2011) found it could be relatively large for some estimates and subpopulations in a telephone survey. In a mail survey, Martikainen et al. (2007) found substantial nonresponse bias in two variables—occupational social class and sickness absence spells—yet no significant bias in the association between the two variables.

Two major factors have hindered the study of the impact of nonresponse on associations. First, methods to evaluate nonresponse bias that rely on information for both respondents and nonrespondents require information for multiple variables of interest. When data come from administrative records or from the sampling frame, they are seldom available for many survey variables. This lack of multivariate data on respondents and nonrespondents limits the routine evaluation of nonresponse bias in bivariate and multivariate relationships. Second, although theories have been proposed for nonresponse bias in univariate statistics that help to identify specific biasing mechanisms, such as lack of interest in the survey topic, the theories for bias in associations need to be far more complex, and the identification of mechanisms are exceedingly challenging.

Estimates of change

For many surveys, particularly those with a panel design, that track the same respondents over time, estimates of change represent key survey statistics. Nonresponse can affect estimates of change through the failure to initially recruit sample members into the panel and through the attrition that occurs at each subsequent wave of data collection. For example, in a survey of smoking habits in Florence, Italy, nonrespondents to the initial mail survey were more likely to have died a decade later compared to respondents (Barchielli and Balzi 2002). That is, the nonresponse to a cross-sectional survey was related to a key estimate of change over time, the mortality rate. Attrition due to nonresponse in subsequent surveys can also bias estimates of change. In another study, Vestbo and Rasmussen (1992) found that nonrespondents to a 10-year follow-up survey of a sample of cement factory workers had rates of hospital admission due to respiratory diseases that were twice as high as those of respondents. This difference persisted after controlling for reported smoking habits in the first survey.

The survey literature on nonresponse and attrition bias in estimates of change is quite scarce. The topic has received more attention in case-control studies in which individuals are assigned to different treatment conditions and outcomes are often measured with the use of surveys. The concern in these studies is that differential nonresponse bias across treatment conditions can bias treatment effect estimates, the estimates of change. More attention is needed for large-scale panel surveys (at least with a longitudinal component), where studies have predominantly focused on bias in demographic characteristics over time. Such examinations will increase the need for theories that could link factors that occur before and after the survey measurement to survey participation—both the likelihood of contact and the likelihood of cooperation.

Components of Nonresponse

In addition to being relatively uninformative about bias in estimates, the overall nonresponse rate masks changes in the components of nonresponse. For interviewer-administered surveys, the largest components of nonresponse are refusals and noncontacts. The reasons for a sample member to refuse to participate in a survey can be very different from the reasons not to be contacted, leading to different magnitudes of bias from refusals and noncontacts across different estimates (for example, see Lin and Schaeffer 1995; Lynn and Clarke 2002). For example, refusals may be the dominant source of nonresponse bias in estimates of drug use in a survey on drug abuse, while noncontacts may be the dominant source in estimates of time use.

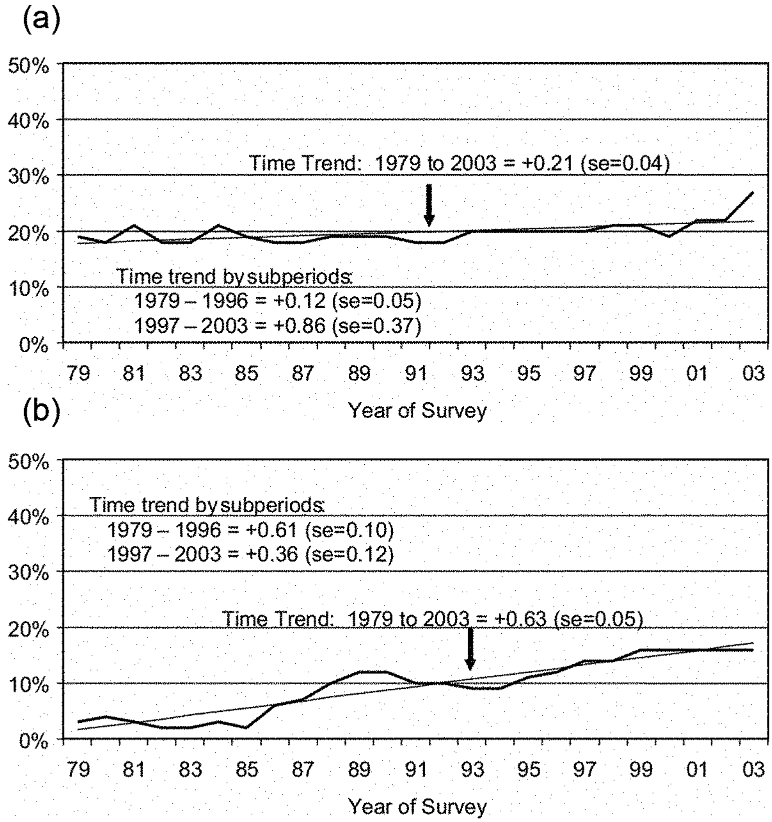
The proportions of refusals and noncontacts have been changing over the past several years, largely driven by reluctance to take part in surveys but also because of increasing difficulty in contacting sample members on the telephone. Thus, the balance between the two main forms of nonresponse is changing in telephone and face-to-face surveys. This difference may reflect multiple reasons, including the growing proportion of nonworking numbers in random-digit-dialing (RDD) frames, the timing of call attempts (in centralized computer-assisted telephone interviewing [CATI] there is little prior information to guide the timing of calls, but field interviewers can make observations that may suggest the best time for a callback), the adoption of new technologies (for example, caller ID), and differences in respondent-interviewer interactions (brief and fast-paced over the phone vs. stronger social presence at the doorstep with visual cues). In addition, every survey implements its own measures to combat each component of nonresponse over time, making generalizations difficult.

Figure 3a shows that refusal rates in the Survey of Consumer Attitudes (SCA) remained quite constant in the period 1979 to 2002, with a sharp increase in 2003. Furthermore, noncontact rates increased between 1985 and 1999 despite more call attempts, and then remained constant through 2003 (Figure 3b). The meaning of noncontact in telephone surveys is also somewhat different from that in face-to-face surveys—many of the noncontacts in the former may be implicit refusals, such as screening calls with answering machines or caller IDs. Nonetheless, in terms of nonresponse bias, it is somewhat consoling that nonresponse is largely driven by participation decisions that are not informed by the study topic and content (since noncontacts are generally unaware of who the caller is and cannot link the number to a prenotification letter).

The components of nonresponse are clearer in face-to-face surveys. Figure 4 shows that refusals in the National Health Interview Survey (NHIS) between 1990 and 2009 increased by 8 percentage points while the noncontact rate increased by only 3 percentage points. If naïve linear regression lines are fit to these rates, the refusal rate increases by 0.42 percentage points per year while the noncontact rate increases by 0.17 percent, which, when combined with other types of unit nonresponse, leads to an average of 0.68 percentage points annual increase in the overall nonresponse rate, or 13 percent over the 20-year period. Although the response rates are far higher than those in telephone surveys, the increasing nonresponse primarily due to refusals is worrisome to the extent that a large proportion of the refusals may be aware of the study topic, sponsor, or other information that can lead to a higher correlation between the likelihood of responding and key survey variables, and thus to greater nonresponse bias.

Experiments have shown that when respondents are aware of the topic and sponsor of the survey, estimates can be biased by nonresponse. In particular, when the topic is made more salient, exaggerated population estimates of interests and activities related to the topic can be expected (Groves et al. 2006). Conversely, less bias can be expected when participation is not related to survey topic or to the sponsor, and noncontacts are generally not exposed to these survey features. Even when salient, however, the biasing effects of these factors can be reduced—incentives have been shown to reduce the link between nonresponse and nonresponse bias by disproportionately increasing participation among those who tend to be underrepresented.

FIGURE 3
Refusal (a) and Noncontact (b) Rates in the Survey of Consumer Attitudes, 1979–2003

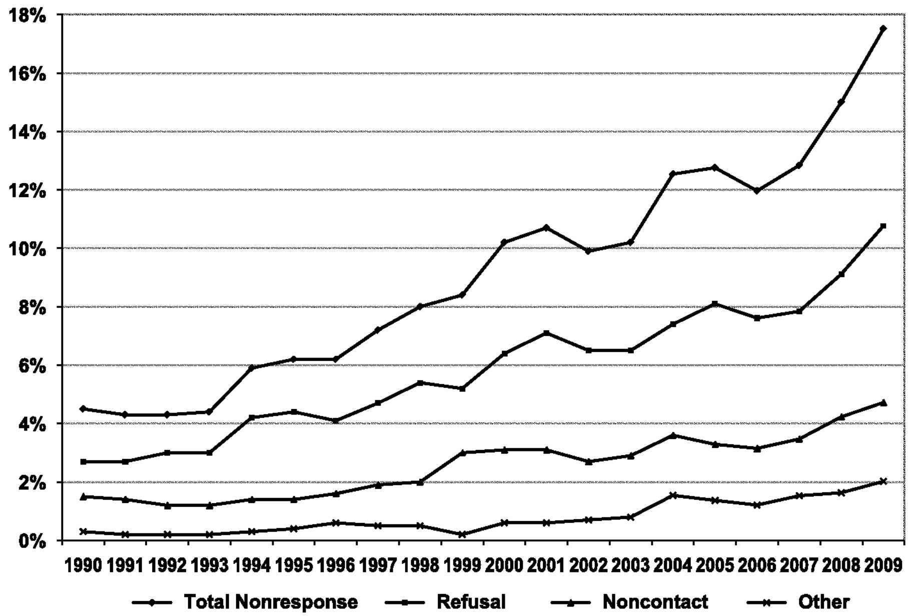


SOURCE: Curtin, Presser, and Singer (2005).

For surveys of persons in households, nonresponse occurs at two stages in the design: the screening interview that establishes eligibility and selects respondents and the main interview used to collect the substantive data from the selected respondents. Figure 5 shows both household screening and main interview components of nonresponse in another ongoing face-to-face survey, NSDUH, over the same 20 years. As can be seen, despite notable design changes between 1990 and 2009, overall nonresponse increased at a similar annual rate of 0.63 percentage points. Two other observations are noteworthy, however: the screening interview is a much smaller component of overall nonresponse, but its rate is increasing at a faster pace than the rate for main interview nonresponse.

Unlike the NHIS, the NSDUH implemented several major changes to the survey design during this period, some of which were aimed at addressing

FIGURE 4
Unweighted Household Nonresponse in the National Health Interview Survey between 1990 and 2009, by Type of Nonresponse



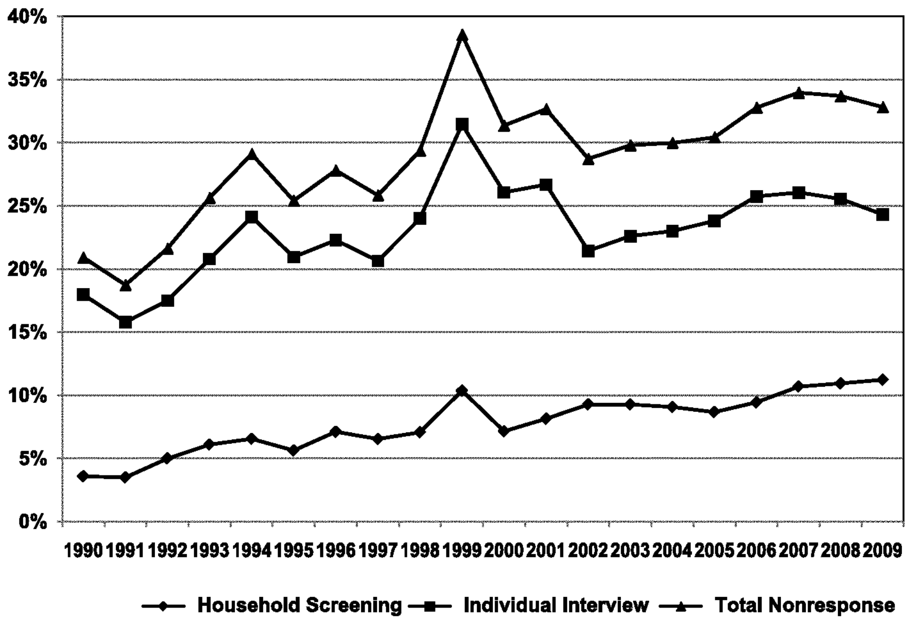
SOURCE: Stussman, Dahlhamer, and Simile (2005).
NOTE: The author is grateful to James Dahlhamer for providing an update through 2009.

nonresponse. In 1995 the questionnaire was changed, reducing respondent burden, and in 2002 incentives were introduced. These changes had a positive impact mostly on the main interview. In 1999, however, NSDUH was redesigned to produce state-level estimates instead of only national estimates, substantially increasing the sample size. The hiring of so many new interviewers and supervisors who lacked prior experience on the survey, among other related factors, led to an increase in both screening and interview nonresponse.

These changes to the study design led to substantial shocks to trend estimates. Between 2002 and 2009 when no new changes were made to the NSDUH design, nonresponse increased by 0.58 percentage points annually. Despite efforts to reduce interview nonresponse, which is particularly evident in 2007 to 2009, the annual increase in the main interview nonresponse rate was higher than the rate of increase in the screening nonresponse rate (0.41 vs. 0.28 percentage points, respectively).

Consistent with the evidence from NHIS, refusals were a key driver for both screening and main interview nonresponse rates. To understand whether these

FIGURE 5
Weighted Household, Individual, and Total Nonresponse in the National Survey of Drug Use and Health between 1990 and 2009, by Study Component



increases in refusals lead to greater bias in survey estimates, an intriguing line of research could examine how the reasons for refusals have changed over time—whether it is unwillingness to do a long interview or unwillingness to do an interview on this topic. As is the case with the SCA and other surveys, these increases are despite greater efforts to tackle nonresponse through fieldwork.

Cost

Combating declining response rates usually requires greater resources.[1] In the SCA, a telephone survey where a large part of the cost is interviewer time, the number of call attempts to complete an interview doubled between 1979 and 1996 from 3.9 to 7.9 (Curtin, Presser, and Singer 2005); the noncontact rate more than tripled during this period (see Figure 3). It is more challenging to link changes in field data collection costs to nonresponse in face-to-face surveys, yet higher proportions of noncontacts and refusals undoubtedly require more visits and more refusal conversion attempts.

The cost stemming from nonresponse is not only monetary. To address nonresponse due to gated communities (or controlled-access buildings), a subsample of such structures was selected and greater effort expended to obtain access to them in the National Comorbidity Survey-Replication (NCS-R). A subsample of all nonrespondents (Deming 1953) is also used in studies such as the National Survey of Family Growth (NSFG) cycles 6 and 7. When properly accounting for this additional selection probability, the effective sample size is usually reduced[2] through an increased variability in survey weights, particularly when only a small fraction of nonrespondents is selected. Even if a double sample approach to nonresponse is not used, nonresponse leads to greater potential for increased variability in survey weights due to nonresponse and poststratification adjustments. The resulting larger design effects on survey estimates reduce the utility of the collected data.

Nonresponse and Other Sources of Survey Error

Even more challenging than quantifying the impact of increasing nonresponse on survey costs is gauging the impact of nonresponse on other sources of error. Nonresponse may lead to suboptimal study designs that produce estimates with more overall bias than could be achieved with the same budget. This could occur because resources allocated to reducing nonresponse could be more effectively used to reduce other sources of error. Nonresponse rates are easily measured, so nonresponse may draw greater attention than other error sources. Nonresponse (unit and item) is also one of only two sources of error that have quantifiable criteria in the Office of Management and Budget guidelines (2006).[3] Requiring nonresponse bias analyses for studies with expected response rates below 80 percent may be beneficial, but not having such criteria for all other sources of error may reduce the relative importance of the other error sources. While for some estimates, nonresponse was a dominant source of error in a study that decomposed total survey error (Groves and Magilavy 1984), it was measurement (or "response") error that was dominant for other estimates; study designs should adequately protect against and measure errors from both sources.

Reducing nonresponse may increase other sources of survey error in some instances, depending on common causes and the methods used to gain participation from likely nonrespondents. While one study found that including likely nonrespondents led to less total bias (Olson 2006), another study found that it increased total bias due to measurement error in responses from likely nonrespondents (Peytchev, Peytcheva, and Groves 2010). In this latter case, focusing only on nonresponse may lead to more biased estimates.

Implications for Survey Design Selection

The idea of a trade-off between error sources can be extended to the selection among fundamentally different study designs. When design decisions are based primarily on nonresponse, a design may be selected that leads to the highest response rate but possibly more biased estimates than if a different design were chosen. Consider the choice between a landline RDD survey and a mail survey to a sample of postal addresses (Link and Mokdad 2006; Peytchev, Ridenhour, and Krotki 2010). The mail survey may achieve a higher response rate, but will face challenges such as within-household selection, low literacy, and potential nonresponse bias (since sample members can see the questions prior to deciding whether to start the survey). The optimal design is unknown a priori for any given study, but these are the types of questions facing surveys, such as the Behavioral Risk Factor Surveillance System (BRFSS) or the Health Information National Trends Survey (HINTS).

Another example is the use of multiple modes to reduce nonresponse. Introducing another mode of data collection can reduce nonresponse but can produce data with different measurement properties for a nonrandom part of the sample, which affects bias (Tourangeau and Smith 1996) and variance (Dillman et al. 2009) of an estimate. The effect of a mixed mode design on survey estimates needs to be evaluated for each particular survey, but, more important, more research is needed on when to expect error trade-offs. Thus, depending on the design, reducing nonresponse may not only lead to an increase in nonresponse bias as described earlier (Merkle and Edelman 2009; Schouten and Cobben 2007) but may also lead to an increase in other sources of error, such as measurement, that result in greater total error (Peytchev, Peytcheva, and Groves 2010).

Increasing Study Complexity

Increasing nonresponse and survey costs are creating the need for more complex study designs that address nonresponse bias in a cost-efficient manner. Over the past several years, survey designs have progressively reflected this need through multiple phases, survey protocol changes informed by data collection, incorporation of multiple methods of data collection, changes in frame construction, use of multiple frames, and so on. These changes, influenced by nonresponse, have direct consequences for survey organizations. For example, survey organizations need systems that can better monitor data collection in real time, as well as staff who are well versed as managers and as survey statisticians. Although fundamental theories and procedures for some of these complex designs have been proposed in the past, there are large voids that remain; double sampling was proposed over half a century ago, yet practitioners lack guidance on when to stop the initial phase of data collection and how to design the sample of nonrespondents. Greater study complexity also affects data users, requiring them to have more technical knowledge about the study design and implementation and a higher level of familiarity with analytic procedures. Several areas in study design that have been influenced by nonresponse are outlined below, focusing on those being actively investigated at present.

Multiphase designs

Neither the idea nor the use of multiple phases in a study is new, but the popularity of multiphase designs has increased due to nonresponse. The notion of response propensities is conditional on the survey protocol (for example, the mode of data collection) and its implementation (e.g., time of call attempt and interviewer-respondent interaction). A change in these essential survey conditions (Hansen, Hurwitz, and Bershad 1961) should result in different response propensities for the sample members. There are ample examples of studies that use multiple phases to reduce nonresponse (and the potential for nonresponse bias) or to obtain measures of nonresponse bias. The effectiveness of this method to address nonresponse bias depends on the protocol used in the second (or later) phase, but, in general, it has been found to produce the largest estimates of nonresponse bias when compared to other methods for assessing nonresponse bias, such as using screener, frame, or supplemental data to compare respondents and nonrespondents (Groves and Peytcheva 2008).

There are several dimensions on which multiphase designs vary. These designs can be ad hoc (for example, implemented when response rates toward the end of data collection are deemed deficient) or reflect a prior plan for the overall design. They can be designed so that the data collected in a second phase are combined with data from an initial phase of data collection, thus limiting the number of changes made to the protocol used in the second phase; or they can be used only to measure nonresponse bias in key variables. Multiphase designs may also include all active sample cases in each subsequent phase, or they may involve taking a double sample of remaining cases to reduce cost and allow for a more costly protocol in the later phase.

These dimensions can affect the utility of incorporating multiple phases. For example, taking a subsample of nonrespondents and using a more effective protocol with them can reduce bias and substantially increase the weighted response rate. However, this may also lead to larger variances and a larger mean square error (MSE) than if the second phase had not been implemented.

Responsive designs

Another development in the face of increasing nonresponse rates is the active management of data collection to reduce bias in estimates and minimize cost of data collection—termed responsive design[4] (Groves and Heeringa 2006). Responsive design requires the monitoring of key indicators, such as survey estimates, variables available on the sampling frame, paradata (information about the process of data collection), and cost. This information is then used to determine when the usefulness of a phase has been exhausted and a new protocol should be launched.

Such active management during data collection places greater burden on those involved in data collection. Numerous models may need to be fit to the various data elements and continuously evaluated, and complex decisions need to be made about cost-error trade-offs. As is the case for studies with predetermined multiple phases, data from the phases in a responsive design also need to be combined, which is not always straightforward and affects other parts of the study.

Not having a fixed predetermined study protocol increases complexity but offers the potential for addressing the threat of nonresponse bias in a cost-efficient manner. Many studies have incorporated responsive design features, such as the use of sample replicates to control the final number of respondents in the face of nonresponse, monitoring and directing effort to strata with lower response rates, changing within-household selection probabilities during data collection to achieve desired sample composition among respondents, and determining the number of adults to interview within a household. Responsive design will undoubtedly continue to receive attention.

Mixed- and multiple-mode designs

The use of multiple modes can help to reduce nonresponse through two avenues. Different modes often provide different ways to contact sample members, and once contacted, people can exhibit different response propensities in each mode. The different costs associated with each mode provide additional impetus for the use of multiple modes, particularly when the modes are offered sequentially, as in the American Community Survey. Perhaps the influence of nonresponse on the use of multiple modes is more evident in simultaneous mixed-mode designs, in which the sample member is given a choice of mode.

Despite the large literature on differences between modes, more research is needed to help to inform study designs. For example, some recent unpublished studies found that offering mail and Web survey options simultaneously lowered the response rate for reasons that remain unclear. The effect of the order in which modes are presented is also not understood. Attempting a survey by mail first may increase the likelihood of participation by telephone by establishing the survey's legitimacy; attempting the telephone first may decrease the likelihood of completion of the mail survey if sample members are consistent with their initial nonparticipation decision.

Far more attention has been given to measurement differences between modes, but causes for differences are not always known, and far more attention is needed on how to design instruments that produce data with equivalent, or even improved, measurement properties across modes. Many survey instruments are designed for a particular mode and then modified for another mode, rather than incorporating considerations about measurement differences across modes during the initial instrument design. An illustration of these divergent approaches can be seen in the selection of response scales. Aural modes have been found to elicit more extreme responses compared to visually administered modes; reducing the number of scale points may be one solution. Similar considerations are needed to address recency, the tendency to select the last response options in aural modes; and primacy, the tendency to select the first response options in visual modes. Multiple-mode designs are needed in part to address nonresponse but can increase error in survey estimates. In one empirical study, measurement error bias due to mode was larger than nonresponse bias in estimates from the NHIS (Biemer 2001). A related line of research would investigate multiple-mode designs that reduce measurement error while also increasing survey participation; for example, some respondents may provide more accurate responses in one mode, while other respondents may do so in another.

Multiple frames

Multiple frames are often used to reduce coverage error in surveys but can also be used to reduce nonresponse and cost. In landline RDD studies, part of the sample is sometimes drawn from listed numbers in addition to the RDD frame. Households with listed numbers have been found to participate in surveys at higher rates, and a higher proportion of telephone numbers are working. This is the dual-frame design being considered for the largest telephone survey in the United States, the National Immunization Survey (Shin, Molinari, and Wolter 2008).

Response rates and cost were also driving factors that led to the dual-frame design of the 2007 HINTS, which obtained half of the interviews from a list-assisted RDD sample and the other half from an address-based sample (ABS), where an alternative with a possibly lower response rate would have been an RDD sampling design with both landline and cell phone numbers. Current examples abound, but the statistical theory on the optimization of multiple frame designs has seen little progress since the work by Hartley (1962). Optimization of dual- and multiframe designs is challenging, particularly in the presence of nonsampling errors that include differential nonresponse and nonresponse bias.

Increased need for additional data

Higher nonresponse increases the need for auxiliary information to evaluate nonresponse bias, to direct efforts during data collection to reduce it, and to construct effective nonresponse adjustments. Demographic variables are clearly insufficient; analysis of nonresponse bias studies shows lack of association between bias in substantive variables and bias in demographic variables (Peytcheva and Groves 2009).

There are at least two main sources of ancillary data that show promise for these purposes. First, data on individuals may reside with government agencies or commercial vendors, or even be in the public domain. Variables such as voting in the previous election tend to be predictive of survey participation and may also be related to key variables in a survey. Since they have not been collected for survey purposes, their usefulness and quality is usually unknown. Further problems of data linkage and missing data also need to be addressed in future research.

Second, survey organizations can collect data about the survey process, such as interviewer-respondent interactions, interviewer performance measures, and timing of call attempts. A broad definition of such paradata will also include the deliberate collection of interviewer observations that are related to survey participation and key variables, as in the Los Angeles Family and Neighborhood Survey (LA FANS), which collected household and neighborhood observations; and NSFG cycles 6 and 7, which collected household observations and direct interviewer guesses about responses to key survey variables. To date, these data have shown limited promise for nonresponse adjustment (Kreuter et al. 2010). Paradata, nonetheless, are being collected in an increasing number of studies.

Greater reliance on properties of auxiliary data

Consideration needs to be given to what elements of paradata are collected and how they are collected; high levels of error in paradata reduce their utility in addressing nonresponse bias. Measurement properties of linked external data are even more suspect, as they are often collected and maintained for other purposes that do not meet the level of rigor needed to produce population estimates. Studies need to evaluate the measurement properties of various paradata.

Random error in auxiliary data can lead to diminished utility, attenuating the relationship between them and response propensities and between them and key survey variables. Random error in paradata can lead to inefficient allocation of resources, such as failure to target cases that are different from earlier respondents, in a responsive design. The attenuation of associations with auxiliary data due to random error will also lead to less effective nonresponse adjustments.

Increased reliance on modeling

Nonresponse has led to an increased reliance on models during and after data collection to capture sample members' probabilities of inclusion in the respondent pool. A few surveys use such models during data collection, directing effort to strata with low response rates or reassigning cases from interviewers with low cooperation rates. Interviewers may induce bias by targeting households that they think will participate based on their own neighborhood and household observations. Through the use of statistical models, data collection managers can undo such influences by directing interviewers to particular sample households.

Even more common, however, is the use of models in nonresponse adjustments. Due to growing nonresponse, weighting class and response propensity models in particular are increasingly being relied on to preserve the probability-based inferential paradigm. There have been at least four interrelated areas of development over the past several years, mentioned only in brief here. First, theoretical work has examined the criteria for selecting variables for nonresponse adjustments. Little and Vartivarian (2005) show that variables associated with the likelihood of responding but not associated with the survey variables contribute to increased variance without a corresponding reduction in bias. Second, there has been incorporation of the uncertainty in nonresponse adjustments into variance estimates. Both WesVar and Stata now provide the ability to create replicate weights to address this issue, and other statistical packages such as SUDAAN can accommodate replicate weights in analysis.[5] Third, developments have been made in model estimation procedures. Response propensity models are increasingly being used for weight adjustment, and these procedures are being incorporated into standard statistical packages. Fourth, and possibly most important, has been work on the identification variables to use in adjustments that impact survey estimates. While this area is also related to the research by Little and Vartivarian (2005), it is more applied in nature—whether associations of auxiliary variables with response propensities and with survey variables are of sufficient magnitude to reduce nonresponse bias in survey estimates (Groves, Wagner, and Peytcheva 2007; Kreuter, Lemay, and Casas-Cordero 2007; Peytchev and Olson 2007; Yan and Raghunathan 2007). Such research could help to guide the collection of new auxiliary data to construct more effective adjustments.

Summary and Conclusion

Nonresponse has clearly had an effect on survey costs, how surveys are designed, how they are conducted, and how they are processed and analyzed; all these changes, undoubtedly, have affected the properties of survey estimates. Substantial nonresponse produces the potential for nonresponse bias in survey estimates. However, higher rates of nonresponse do not necessarily lead to greater nonresponse bias, and response rates are in fact a poor indicator of nonresponse bias.

Nonresponse bias in estimates of means and proportions may also offer an insufficient indication of the overall effects of nonresponse on survey estimates. Nonresponse can also affect the variance of any statistic, providing a misleading level of confidence in univariate statistics, and can also bias estimates of bivariate and multivariate associations. Furthermore, through nonresponse to the initial survey request and through nonresponse in subsequent waves (attrition), estimates of change can be affected by bias.

While response rates across surveys may not be informative about nonresponse bias, declining response rates have led to changes in the components of nonresponse. The currently low response rates in face-to-face surveys are primarily the result of refusals. To the extent that refusals to survey requests are informed by study features, the threat of nonresponse bias becomes more substantial.

Survey costs related to nonresponse have been increasing, both monetarily and in terms of data utility. Addressing nonresponse through a double sample, for example, can increase the design effect in weighted estimates, likely achieving a lower effective sample size than a single-phase design.

Survey designs are also increasing in complexity. Including a nonresponse phase is becoming essential for the measurement and reduction of nonresponse bias. More elaborate and dynamic responsive designs are also being implemented, in which design decisions are made in real time during data collection. Some designs attempting to reduce nonresponse and keep costs under control are also incorporating multiple modes, since sample members exhibit different response propensities for different modes, and modes vary in cost structure. Some modes require separate sampling frames. Since combinations of sampling frames and modes of data collection yield higher response rates than single-frame or single-mode designs, the use of multiple sampling frames is also increasingly being considered.

Nonresponse is also increasing the reliance of probability-based survey estimates on modeling. Due to nonresponse, inclusion probabilities of interviewed respondents are no longer known. Nonresponse adjustments aim to remedy this potential threat to probability-based inference.

To inform decisions during data collection and to inform adjustments for nonresponse, auxiliary data need to be identified and collected. Some external data already exist on individuals in the population but remain underutilized in surveys; they can be found in administrative databases, compiled by commercial vendors, and in the public domain. To facilitate their use, problems need to be addressed, such as methods to link data from multiple sources. Other information, or paradata, needs to be obtained during data collection. Such measures can be collected by interviewers or directly by survey systems. Current challenges are the identification of what data elements to collect, and how to organize such data structures to aid data collection and creation of postsurvey adjustments.

Notes

1. There are exceptions. For example, the introduction of incentives in NSDUH led to higher response rates and lower data collection costs per completed interview (Kennet et al. 2005).

2. The sample of nonrespondents can be designed in a manner that minimizes the increase in variability of the weights, as done in NSFG cycle 6, which in some studies may even reduce the design effect due to the double sample.

3. Coverage error indicated by the coverage rate is the other source of error.

4. This is also referred to as adaptive design in clinical trials.

5. Although replication methods for variance estimation could be used without standard procedures in analysis software, such procedures are needed to help their implementation.

References

Barchielli, Alessandro, and Daniela Balzi. 2002. Nine-year follow-up of a survey on smoking habits in Florence (Italy): Higher mortality among non-responders. *International Journal of Epidemiology* 31:1038–42.

Bethlehem, Jelke. 2002. Weighting nonresponse adjustments based on auxiliary information. In *Survey nonresponse*, eds. R. M. Groves, D. A. Dillman, J. L. Eltinge, and R. J. A. Little. New York, NY: Wiley.

Biemer, Paul P. 2001. Nonresponse bias and measurement bias in a comparison of face to face and telephone interviewing. *Journal of Official Statistics* 17:295–320.

Carley-Baxter, Lisa R., Craig A. Hill, David J. Roe, Susan E. Twiddy, Rodney K. Baxter, and Jill Ruppenkamp. 2009. Does response rate matter? Journal editors' use of survey quality measures in manuscript publication decisions. *Survey Practice*, November.

Curtin, Richard, Stanley Presser, and Eleanor Singer. 2000. The effects of response rate changes on the index of consumer sentiment. *Public Opinion Quarterly* 64 (4): 413–28.

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 comparison. In *Survey nonresponse*, eds. R. M. Groves, D. A. Dillman, J. L. Eltinge, and R. J. A. Little. New York, NY: Wiley.

Deming, W. Edwards. 1953. On a probability mechanism to attain an economic balance between the resultant error of nonresponse and the bias of nonresponse. *Journal of the American Statistical Association* 48:743–72.

Dillman, Don A., Glenn Phelps, Robert Tortora, Karen Swift, Julie Kohrell, Jodi Berck, and Benjamin L. Messer. 2009. Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the Internet. *Social Science Research* 38 (1): 1–18.

Groves, Robert M. 1989. *Survey errors and survey costs*. Wiley Series in Probability and Mathematical Statistics. New York, NY: Wiley.

Groves, Robert M. 2006. Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly* 70 (5): 646–75.

Groves, Robert M., J. Michael Brick, Mick P. Couper, William Kalsbeek, Brian Harris-Kojetin, Frauke Kreuter, Beth-Ellen Pennell, Trivellore Raghunathan, Barry Schouten, Tom Smith, Roger Tourangeau, Ashley Bowers, Mathew Jans, Courtney Kennedy, Rachel Levenstein, Kristen Olson, Emilia Peytcheva, Sonya Ziniel, and James Wagner. 2008. Issues facing the field: Alternative practical measures of representativeness of survey respondent pools. *Survey Practice*, October.

Groves, Robert M., and Mick P. Couper. 1998. *Nonresponse in household interview surveys*. New York, NY: Wiley.

Groves, Robert M., Mick P. Couper, Stanley Presser, Eleanor Singer, Roger Tourangeau, Giorgina Piani Acosta, and Lindsay Nelson. 2006. Experiments in producing nonresponse bias. *Public Opinion Quarterly* 70 (5): 720–36.

Groves, Robert M., and Steven Heeringa. 2006. Responsive design for household surveys: Tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society Series A: Statistics in Society* 169 (Part 3): 439–57.

Groves, Robert M., and Lou Magilavy. 1984. An experimental measurement of total survey error. *Proceedings of the American Statistical Association*.

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.

Groves, Robert M., Stanley Presser, and Sarah Dipko. 2004. The role of topic interest in survey participation decisions. *Public Opinion Quarterly* 68 (1): 2–31.

Groves, Robert M., Eleanor Singer, and Amy Corning. 2000. Leverage-saliency theory of survey participation—Description and an illustration. *Public Opinion Quarterly* 64 (3): 299–308.

Groves, Robert M., James Wagner, and Emilia Peytcheva. 2007. Use of interviewer judgments about attributes of selected respondents in post-survey adjustment for unit nonresponse: An illustration with the National Survey of Family Growth. *Proceedings of the Joint Statistical Meetings, Survey Research Methods Section, ASA, at Salt Lake City, UT*.

Hansen, Morris H., William N. Hurwitz, and Max A. Bershad. 1961. Measurement errors in censuses and surveys. *Bulletin of the ISI* 38:351–74.

Hartley, Herman O. 1962. Multiple frame surveys. *Proceedings of American Statistical Association, Section on Social Statistics*.

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

Kennet, Joel, Joseph Gfroerer, Katherine R. Bowman, Peilan C. Martin, and David Cunningham. 2005. Introduction of an incentive and its effects on response rates and costs in NSDUH. In *Evaluating and improving methods used in the National Survey on Drug Use and Health*, eds. Joel Kennet and Joseph Gfroerer. Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

Khare, Meena, Leyla K. Mohadjer, Trena M. Ezzati-Rice, and Joseph Waksberg. 1994. An evaluation of nonresponse bias in NHANES III (1988-91). *Proceedings of the section on Survey Research Methods, American Statistical Association*.

Kreuter, Frauke, Michael Lemay, and Carolina Casas-Cordero. 2007. Using proxy measures of survey outcomes in post-survey adjustments: Examples from the European Social Survey (ESS). *Proceedings of the Joint Statistical Meetings, Survey Research Methods Section, ASA, at Salt Lake City, UT*.

Kreuter, Frauke, Kristen Olson, James Wagner, Ting Yan, Trena M. Ezzati-Rice, Carolina Casas-Cordero, Michael Lemay, Andy Peytchev, Robert M. Groves, and Trivellore E. Raghunathan. 2010. Using proxy measures and other correlates of survey outcomes to adjust for nonresponse: Examples from multiple surveys. *Journal of the Royal Statistical Society Series A: Statistics in Society* 173 (2).

Lepkowski, James M., and Mick P. Couper. 2002. Nonresponse in the second wave of longitudinal household surveys. In *Survey nonresponse*, eds. R. Groves, D. Dillman, J. Eltinge, and R. J. A. Little. New York, NY: Wiley.

Lin, I-Fen, and Nora Cate Schaeffer. 1995. Using survey participants to estimate the impact of nonparticipation. *Public Opinion Quarterly* 59:236–58.

Link, Michael W., and Ali H. Mokdad. 2006. Can web and mail survey modes improve participation in an RDD-based national health surveillance? *Journal of Official Statistics* 22 (2): 293–312.

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

Lorant, Vincent, Stefaan Demarest, Pieter-Jan Miermans, and Herman Van Oyen. 2007. Survey error in measuring socio-economic risk factors of health status: A comparison of a survey and a census. *International Journal of Epidemiology* 36 (6): 1292–99.

Lynn, Peter, and Paul Clarke. 2002. Separating refusal bias and non-contact bias: evidence from UK national surveys. *Journal of the Royal Statistical Society Series D: The Statistician* 51 (3): 319–33.

Martikainen, Pekka, Mikko Laaksonen, Kustaa Piha, and Tea Lallukka. 2007. Does survey non-response bias the association between occupational social class and health? *Scandinavian Journal of Public Health* 35 (2): 212–15.

Merkle, Daniel M., and Murray Edelman. 2009. An experiment on improving response rates and its unintended impact on survey error. *Survey Practice*, March.

Office of Management and Budget. 2006. *Standards and guidelines for statistical surveys 2006*. Available from http://www.whitehouse.gov/omb/inforeg/statpolicy/standards_stat_surveys.pdf (accessed 4 December 2006).

Olson, Kristen. 2006. Survey participation, nonresponse bias, measurement error bias, and total bias. *Public Opinion Quarterly* 70 (5): 737–58.

Peytchev, Andy, Rodney K. Baxter, and Lisa R. Carley-Baxter. 2009. Not all survey effort is equal: Reduction of nonresponse bias and nonresponse error. *Public Opinion Quarterly* 73 (4): 785–806.

Peytchev, Andy, Lisa R. Carley-Baxter, and Michele C. Black. 2011. Multiple sources of nonobservation error in telephone surveys: Coverage and nonresponse. *Sociological Methods & Research* 40 (1): 138–68.

Peytchev, Andy, and Kristen Olson. 2007. Using interviewer observations to improve nonresponse adjustments: NES 2004. *Proceedings of the Joint Statistical Meetings, Survey Research Methods Section, ASA, at Salt Lake City, UT*.

Peytchev, Andy, Emilia Peytcheva, and Robert M. Groves. 2010. Measurement error, unit nonresponse, and self-reports of abortion experiences. *Public Opinion Quarterly* 74 (2): 319–27.

Peytchev, Andy, Jamie Ridenhour, and Karol Krotki. 2010. Differences between RDD telephone and ABS mail survey design: Coverage, unit nonresponse, and measurement error. *Journal of Health Communication: International Perspectives* 15 (1, Suppl. 3): 117–34.

Peytcheva, Emilia, and Robert M. Groves. 2009. Using variation in response rates of demographic subgroups as evidence of nonresponse bias in survey estimates. *Journal of Official Statistics* 25 (2): 193–201.

Schouten, Barry, and Fannie Cobben. 2007. R-indexes for the comparison of different fieldwork strategies and data collection modes. Statistics Netherlands Discussion Paper 07002, Voorburg/Heerlen.

Shin, Hee-Choon, Noelle-Angelique Molinari, and Kirk Wolter. 2008. A dual-frame design for the national immunization survey. *Proceedings of Survey Research Methods Section of the American Statistical Association, at AAPOR Annual Conference*, New Orleans, LA.

Stussman, Barbara, James Dahlhamer, and Catherine Simile. 2005. *The effect of interviewer strategies on contact and cooperation rates in the National Health Interview Survey*. Washington, DC: Federal Committee on Statistical Methodology.

Tourangeau, Roger, and Tom W. Smith. 1996. Asking sensitive questions—The impact of data collection mode, question format, and question context. *Public Opinion Quarterly* 60 (2): 275–304.

Vestbo, J, and F. V. Rasmussen. 1992. Baseline characteristics are not sufficient indicators of non-response bias follow up studies. *Journal of Epidemiology and Community Health* 46 (6): 617–19.

Wolter, Kirk. 2007. Introduction to variance estimation. In *Statistics for social and behavioral sciences*, 2nd edition, eds. S. E. Feinberg and W. J. v. d. Linden. New York, NY: Springer.

Yan, Ting, and Trivellore Raghunathan. 2007. Using proxy measures of the survey variables in post-survey adjustments in a transportation survey. *Proceedings of the Joint Statistical Meetings, Survey Research Methods Section, ASA, at Salt Lake City, UT*.