The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis

Article *in* Public Opinion Quarterly · June 2008

DOI: 10.1093/poq/nfn011

2 authors:

Robert Groves
Georgetown University

Emilia Peytcheva
Research Triangle Institute

The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis[1]

Robert M. Groves[2] and Emilia Peytcheva[3]

Abstract. Fifty-nine methodological studies were designed to estimate the magnitude of nonresponse bias in statistics of interest. These studies use a variety of designs: sampling frames with rich variables, data from administrative records matched to sample case, use of screening interview data to describe nonrespondents to main interviews, followup of nonrespondents to initial phases of field effort, and measures of behavior intentions to respond to a survey. This permits exploration of what circumstances produce a relationship between nonresponse rates and nonresponse bias and what, do not. The predictors are design features of the surveys, characteristics of the sample, and attributes of the survey statistics computed in the surveys.

Introduction

Much survey research follows the inferential paradigm that assumes 100% response rates on a probability sample of a designated frame. That is, the unbiasedness of estimates and of their measured standard errors permit probability statements about population characteristics when all sample elements are measured. When only a subset is measured, none of the properties of probability sampling inference pertain, unless some model of the impact of nonresponse is posited that permits them.

The survey profession is undergoing challenges to this basic paradigm of inference because of the falling response rates in sample surveys throughout the richer countries of the world (deLeeuw and de Heer, 2002). The challenges are exacerbated by the fact that survey designs seeking high response rates are experiencing increasing costs, generated by repeated efforts to obtain access to sample units and to address any concerns of the sample persons.

Groves (forthcoming) examined a set of 30 studies estimating nonresponse bias of descriptive statistics. He finds that the nonresponse

[1] Appreciation for research support from the US National Science Foundation grant SES-0207435.
[2] University of Michigan and Joint Program in Survey Methodology, BGROVES@ISR.UMICH.EDU
[3] University of Michigan

rate, by itself, is a poor predictor of bias magnitudes on the 319 different estimates that can be computed from the studies. Nonresponse rates "explain" only about 11% of the variation in different estimates of nonresponse bias. He notes that a meta-analytic study of a larger number of such studies might be able to examine characteristics of estimates that are related to bias. This paper presents such a meta-analysis.

Theories Linking Nonresponse Rates and Nonresponse Bias

Survey researchers have lamented the lack of theory involving nonresponse bias for some time (Goyder, 1987; Brehm, 1993). Part of this lack of theory may be due to an over-concern among social scientists with nonresponse rates versus nonresponse bias in their theorizing. Indeed, most of the professional concern with nonresponse is with response rates (Bradburn, 1992; Martin, 2004) and the cost implications of attempting to achieve high response rates.

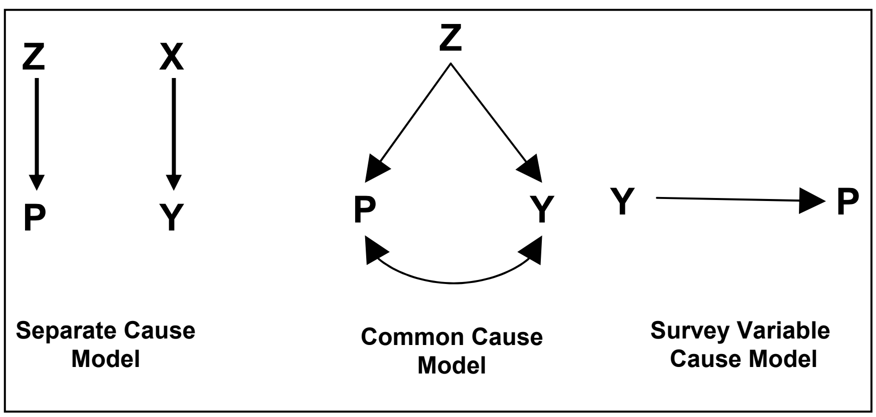

Figure 1. Three Relevant Causal Models Linking Response Propensity with Nonresponse Bias

If one turns from response propensities to nonresponse bias, then a set of causal models becomes of paramount importance (Groves, 2006). As graphically shown in Figure 1, the "separate cause" model asserts that the vector of causes of the *Y* variable is independent of the causes of response propensity, *P*. In this case, expected values of *Y* among respondents would be unbiased estimates of those among all sample persons and corresponds to the "missing completely at random" case

(Rubin, 1987). The "common cause" model asserts that there are shared causes (*Z*) of response propensity and the *Y* variable; this model corresponds to the "missing at random" case. The "survey variable cause" model asserts that *Y* itself is a cause of response propensity; this is the "nonignorable" condition of nonresponse.

All three of these concern possible causal structures underlying nonresponse bias, which, for a simple respondent mean, can be portrayed as $\boldsymbol{\sigma}_{yp} / \bar{\boldsymbol{p}}$, where σ_{yp} is the covariance between a given survey variable, *y*, and the response propensity, *p*; and $\bar{\boldsymbol{p}}$ is the expected propensity over the sample members to be measured (Bethlehem, 2002). The separate cause model would produce a zero covariance; the common cause model would produce a nonzero covariance (but a zero covariance, controlling for *Z*); and the survey variable cause model would have a nonzero covariance.

The expression above reminds us that nonresponse bias varies over different estimates within a survey, as a function of whether the likelihood of survey participation is related to the variable underlying the estimate. In the same survey, some estimates can be subject to large nonresponse biases; others, to negligible biases. The scientific question (the "why" question) associated with this expression is "what causes a correlation between *y* and *p*" or "what causes a survey variable to be correlated to the likelihood to respond?"

Leverage-salience theory can be used to motivate hypotheses about when variation in nonresponse propensity tends to induce nonresponse bias (Groves, Singer, Corning, 2000). That theory asserts that the causes of the survey participation decision vary over persons and over the presentational content of the survey request. Some persons are stimulated to respond because of one feature of a survey request (e.g., the stated purpose of the survey); others, because of some other feature (e.g., the fact the survey is quite short). Some persons are positively disposed to a survey feature (e.g., who sponsors the survey); others, are negatively disposed to the same feature. The impact of each of these features is determined by how salient the given feature is made in the introduction to the survey. For example, if the interviewer emphasizes the confidentiality provisions of the survey data, those concerned about such issues may be positively disposed; if not, they may not. The influence of individual features of the survey request is dependent on how much predispositions of the sample person positively or negatively weight the features. These predispositions are largely unknown prior to the survey request being made, thus, the response rate is a function of the salience given to individual features in an individual encounter with a sample person and the distribution of leverages for the features within the sample population.

When a factor that has great leverage on the survey participation decision for many sample persons is also an item of survey measurement, survey statistics based on it are likely to have large nonresponse bias. Under leverage-salience theory, factors relevant to the self-image of the sample person operate as influences on the participatory decision when they are made salient in the description of the survey request. Thus, both leverage and salience are required.

It is noteworthy (especially for this meta-analysis), that leverage-salience theory suggests few main effects of single influences on nonresponse bias; the theory is inherently one of interaction effects. It notes that different leverages for a given aspect of the survey task exist for different people, and that they should affect the respondent distributions of only the subset of survey variables related to those aspects. If diverse factors influence participation, then bias depends on how the factors link to the survey variables. For example, Lahaut, Jansen, Van de Mheen, and Garretson (2002) show that both teetotalers and heavy alcohol users tend to be reluctant respondents to a survey on alcohol use. This empirical result is compatible with a lack of interest in alcohol among teetotalers influencing their nonresponse and a fear of embarrassment among the heavy consumers. In short, useful theories about nonresponse *bias* versus nonresponse *propensity* require conceptual linkage between individual survey measures and participatory influences. Meta-analyses are strong tools to find main effects of single influences on some phenomenon; they are weaker tools when the phenomenon has complicated, multivariate influences.

Much of the literature on nonresponse focuses on attributes of the survey design (e.g., mode, the use of incentives). We note that in the meta-analysis such attributes can offer insight only into the *among-survey* component of variation in nonresponse bias. That is, since all estimates *within* a survey have the same value of a given design attribute, it cannot "explain" variation in bias among those estimates within the same survey. In this meta-analysis only a small part of the variation in bias components lies between surveys versus within surveys (between estimates). Thus, it is unlikely that such design features will greatly illuminate differences across survey estimates in nonresponse bias.

This paper reports on a meta-analysis of correlates of nonresponse bias based on 59 studies designed to produce such estimates. We address three questions: a) are there characteristics of survey design that are systematically related to nonresponse bias?, b) are the properties of target populations related to nonresponse bias?, and c) are there characteristics of survey estimates that are systematically related to nonresponse bias?

Research Design

Meta-analyses involve the study of a set of research projects from a target population of research studies, where the inferences about the common or typical findings are desired. It is important to understand the characteristics of the studies included before examining their results.

Target Population of Research Studies and Selection Mechanisms

The meta-analysis defined the target population of research as studies conducted since 1978. This date was chosen rather arbitrarily to include enough years to have hopes of collecting a sufficiently large number of studies and to include "modern" survey methods that apply to current practice.

The articles result from a search of a wide variety of electronic databases for literature on survey nonresponse, including The Scholarly Journal Archive (JSTOR), Gale/Info Trac Expanded Academic ASAP, ABI/INFORM Global, LexisNexis, Proquest Research Library, SilverPlatter databases, OCLC Social Science Abstracts, ECO and ArticleFirst databases, SocioFile, ISI Web of Knowledge, Web of Science Social Sciences Citation Index and ISI Proceedings, and ScienceDirect. Searches of journals with a specific focus on survey methodology, such as *Public Opinion Quarterly* and *Journal of Official Statistics*, and searches of survey methodology reference books, such as *Nonresponse in Household Interview Surveys*, were also performed. We reviewed proceedings of the American Statistical Association Survey Research Methods Section and papers presented at the 1999 International Conference on Survey Nonresponse. In addition, we conducted general Google Internet searches for survey nonresponse literature and specific searches for nonresponse studies from the Survey of Consumer Finances and National Center for Education Statistics surveys. Then, references to other work cited in the found articles were pursued. Much of the literature exists in journals in the biomedical field, possibly because of the existence of record bases, which are used as gold standards in the studies. The effort resulted in 47 articles that fit the criterion; in total, 59 separate studies were reported in the articles (see the Appendix for a complete list of references).

To be eligible, the research needed to have produced estimates of nonresponse bias for a set of population means or percentages. Acceptable techniques for producing these were:

1. sample frame data (i.e., where records were available both on respondents and nonrespondents), and means on the record variables were estimated;
2. supplemental data, for both respondents and nonrespondents, linked to the sample person's data;
3. screener interview data, used to compare respondents and nonrespondents to a later larger interview;
4. followup studies of sample persons who were nonrespondents to a survey, comparing the earlier respondent group to those former nonrespondents measured in the followup; and
5. reports of intentions to respond to a later survey, comparing those who report agreeing to respond with those who decline to respond.

The studies examined a wide variety of target populations, including US national populations, communities, health service members, physicians, employees of an organization, company customers, low income women, visitors to a recreational lake, disabled people, university students and alumni, special interest groups, voters, new parents, and others. The most prevalent topic is health (59%), followed by employment (11%). Most estimates arose from self-administered surveys (56%); 27% from face-to-face surveys; 17% from telephone surveys. The vast majority of studies are documented in peer-reviewed journals (81%).

In addition to recording the nonresponse bias estimates, we attempted to record the following characteristics of the surveys: year of publication, survey length, survey topic, topic saliency, survey sponsor, evidence for respondents' involvement with the survey sponsor, prenotification about the survey request, incentives, mode of data collection, mode of the nonresponse followup. We coded population type, sample characteristics such as mean age, gender and majority/minority distributions, and urbanicity of the sample. We coded each reported estimate by type of statistic (percentage, mean, median), relevance of the statistic to the survey topic, and type of measure (attitudinal, behavioral).

IMPUTATION FOR ITEM MISSING DATA IN PREDICTOR VARIABLES

We sought to obtain auxiliary characteristics above as predictors of nonresponse bias. Meta-analyses are often plagued by item missing data on predictor variables because the analyses are dependent on what attributes are documented in the scientific paper. The pressures on editors to reduce the number of pages of journal articles and the lack of uniform standards for documentation produce holes in metadata.

Despite repeated efforts to contact those responsible for the research studies to learn about undocumented attributes of the studies or their estimates, we failed in many cases. Faced with item missing data, we preferred to avoid the complete-case analysis option. Instead, we chose to build imputation models for the item missing data on the predictors. The use of sequential regression techniques in imputation permits the construction of a complete-case data set with all of the covariance properties of the original data set (Raghunathan *et al.*, 2001). By using this technique in the context of a multiple imputation design, we can estimate the impact of the imputation variance on estimates computed on the imputed data.

Of the predictor variables discussed in this paper, the variables indicating survey sponsorship and whether the sample had prior involvement with the sponsor of the survey had values for 2 of 59 studies imputed. Values for urbanicity of the sample were imputed for 15 of the 59 studies, while values for subcultural mix of the sample were imputed for 41 of the 59 studies.

CHARACTERISTICS OF THE META-ANALYSIS OBSERVATIONS

Prior to presenting the results of a meta-analysis, the assembly of nonresponse bias estimates needs to be scrutinized. We note that across all the estimates the nonresponse rates of the studies range from 14% to 72%, with a mean nonresponse rate of 36%. Most of the estimates come from studies using non-survey records (24% from the sampling frame, 32% from a supplementary data set); 28% come from studies using followup of nonrespondents with some extraordinary effort. The remainder is mostly studies using screener interview data (14%). A very small percentage (2%) use reports of intentions to be a respondent or nonrespondent.

A serious weakness of the past literature we examine in the meta-analysis is that nonresponse due to failure to access the sample case (i.e., noncontact, delivery failure) is not separated from nonresponse due to refusal or other reasons. There is now a well-established empirical literature to show that the various types of nonresponse are differentially productive of bias in statistics of different types (e.g., Groves and Couper, 1998; Campanelli, Sturgis and Purdon, 1997). Unfortunately, the articles used in the meta-analysis combine all the types of nonresponse into one category and present differences between respondents and the total set of nonrespondents.

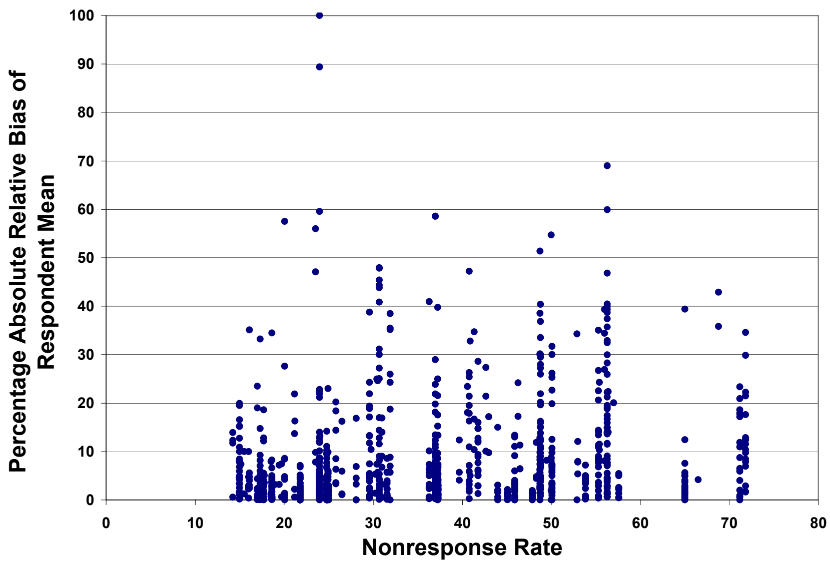

Figure 2. Percentage Absolute Relative Nonresponse Bias of 959 Respondent Means by Nonresponse Rate of the 59 Surveys in Which They Were Estimated

Figure 2 presents a scatterplot of one function of nonresponse bias:

$$\left| \frac{100 * (\bar{y}_r - \bar{y}_n)}{\bar{y}_n} \right|$$

where the numerator contains the difference between respondent and full sample means, and the denominator is the full sample mean. This is the percentage absolute relative nonresponse bias, portraying the bias as a portion of the full sample estimate.

This figure contains a point for each of the means reported in the 59 studies, with complementary percentages for binary variables.[4] The plot is a series of vertical sequences of points, representing different estimates computed from the same survey. The figure clearly shows: a) large relative nonresponse biases exist in the studies, b) most of the variation in nonresponse lies across estimates within the same survey; and, as implied by that observation, c) the nonresponse rate of a survey, by itself, is a poor predictor of the absolute relative nonresponse bias.[5] In short, insight into the linkage between nonresponse rates and nonresponse bias needs more information about the circumstances of the survey measurement.

Analytic Plan

We have an opportunity with the meta-analysis observations to estimate a variety of functions of nonresponse bias. However, there are problems facing some options because of lack of documentation. Each article provides estimates of the respondent mean, $\bar{y}_r$; the nonrespondent mean, $\bar{y}_m$; and the full sample mean, $\bar{y}_n$. The number of cases for each mean is usually cited or could be computed from other reports in the article. Sample designs are reported. However, the element variances of the y variables are not generally reported, nor are standard errors of the estimated means.

We will view the data set as observations from 59 clusters (studies) of observed nonresponse biases. The observations are heteroskedastic because the measurements are subject to different element variances in the studied population and they are subject to different sampling variances because of sample design and size. Further, imputation variance must be reflected for some variables.

---

[4] For each binary variable, two percentages can be computed. The smaller of the two tends to generate higher relative nonresponse bias. Hence, the figure presents the nonresponse bias of both complementary percentages.

[5] If a naïve OLS regression line were fit to the scatterplot, the R^2 would be .04.

We note that the respondent mean, $\bar{y}_r$, and the nonrespondent mean, $\bar{y}_m$, can be viewed as independent observations. When y is a binary variable, its element variance is a function of its mean value, and we can properly estimate the standard error of each of the observations. For count and continuous variables, we cannot do that. Hence, this paper limits the statistical analysis to 566 estimates of transformed *y* variables, using standardized variables that have equal element variances. These 566 estimates come from 44 of the 59 studies. We present estimated absolute values of differences between respondent and nonrespondent standardized means, $\left|\bar{y}_r - \bar{y}_m\right|$, weighting each observation by its sample size, to reflect unequal sampling variances. This statistic is a direct measure of how the attributes of respondents and nonrespondent differ. The differences can be interpreted in units of standard deviations of the standardized variables.

We used IVEWARE (Raghunathan, 2002) within a SAS format, with 20 replicate imputed data sets, with studies as a clustering factor, to estimate the standard errors of the estimated coefficients.

The reader should note that prior conceptual work in nonresponse suggests that the mechanisms productive of nonresponse bias are inherently multivariate (e.g., Groves, Presser, and Dipko, 2004). For example, incentives are seen to reduce nonresponse bias when the survey topic is made highly salient in the survey request (Groves, Couper, Singer, Tourangeau, Piani Acosta, Nelson, 2006). We have fit multivariate models to the meta-analytic observations, but have found many of the coefficients are rather unstable, given the sparseness of data set for various contrasts. In this paper, we present bivariate relationships with nonresponse differences and comment on properties of the data set that evoke some cautions to the conclusions. We hope that the addition of other studies, with characteristics now underrepresented in the literature, will permit future multivariate modeling of the data.

The questions we posed to this meta-analytic data set are related to three possible linkages between nonresponse rates and nonresponse bias: a) attributes of the survey design, b) attributes of the sample population, and c) characteristics of the individual statistic estimated in the survey.

Attributes of the Survey Design

NONRESPONSE DIFFERENCES AS FUNCTIONS OF NONRESPONSE RATES

Some survey researchers have speculated that the differences between respondents and nonrespondents are themselves functions of the response rate (Keeter *et al.*, 2000). This is implied by the hypothesis of the "continuum of resistance," a notion largely unsupported in empirical studies (Lin and Schaeffer, 1995). Sometimes the argument is made that in surveys with higher nonresponse rates, there is a more heterogeneous mix of nonrespondents. With very low nonresponse rates, the argument goes, nonrespondents are quite different from the bulk of the sample. Figure 3 groups the 44 studies into thirds by their nonresponse rates and presents the weighted mean differences between the respondent and nonrespondent estimates for 566 standardized estimates. (For economy of language, we will use the more compact phrase, "nonresponse differences" instead of "differences between respondent and nonrespondent means.") Given the standard errors of the three means, it is clear that nonresponse differences are largely similar across the range of nonresponse rates found. We note that a separate analysis (not presented) shows that the nonresponse *bias* estimates themselves do not reliably vary across these three groups (again, merely reflecting large variation across estimates in surveys with similar nonresponse rates).

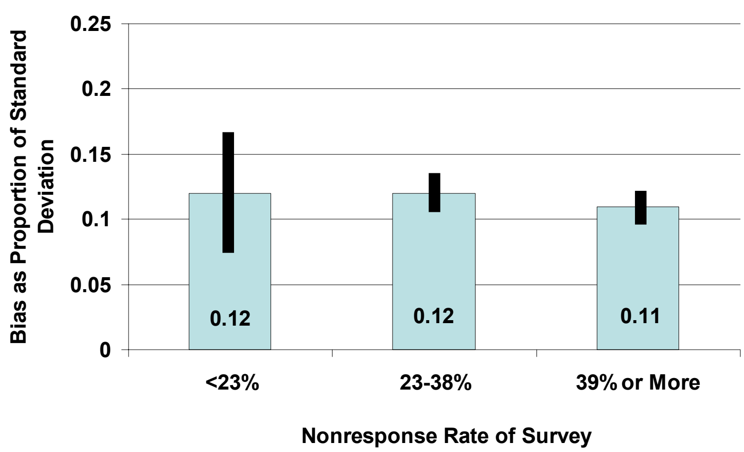

Figure 3. Average Nonresponse Differences $|(\bar{y}_r - \bar{y}_m)|$ for 566 Standardized Estimates from 44 Studies Grouped into 3 Nonresponse Rate Groups (black lines reflect one standard error above and below the group mean, with standard errors reflecting the clustering of observations into studies)

VARIATION IN NONRESPONSE DIFFERENCES FROM THE METHOD USED IN THE NONRESPONSE STUDY

Upon first examination of the patterns of nonresponse differences we noticed an unexpected pattern related to the method used in the nonresponse study (see Figure 4). Two techniques used variables that were not themselves part of the survey measurement – one set used sampling frame variables; another, some supplementary data set matched to the sample. These two methods of studying nonresponse bias tend to report estimates of lower bias (nonresponse differences of .08 and .10, respectively). In contrast, when the nonresponse bias study compared respondents and nonrespondents to a main interview based on variables measured on both groups in a screening interview, the average nonresponse difference is .19. Similarly, if the nonresponse bias study compared early to later respondents during the course of followup, the average nonresponse difference is .14. In short, the "screener" and "followup" methods tend to be distinctive from the other methods.

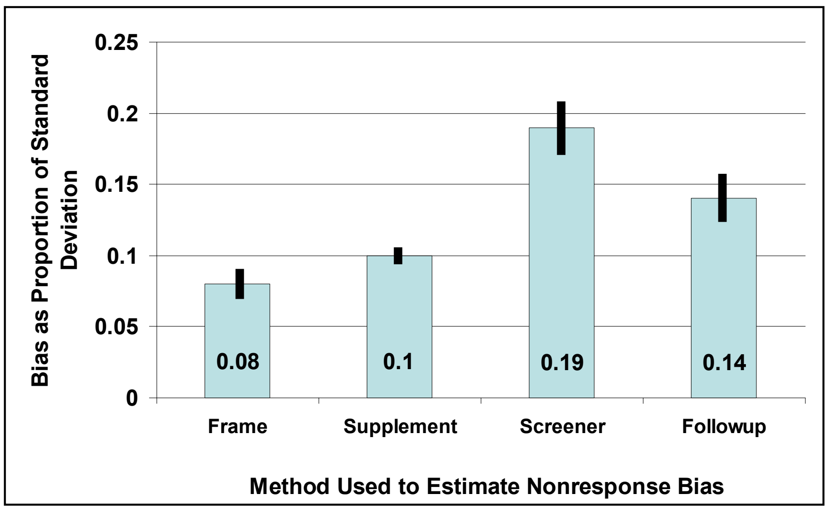

Figure 4. Average Nonresponse Differences ,$|(\bar{y}_r - \bar{y}_m)|$, for 566 Standardized Estimates from 44 Studies Grouped by Methods Used to Estimate Nonresponse Bias (black lines reflect one standard error above and below the group mean, with standard errors reflecting the clustering of observations into studies)

There are quite likely diverse causes of these differences. Some may reflect different kinds of variables measured in the studies (e.g., items relevant to the topic of the survey versus others). Some may reflect different modes of data collection in the secondary phase of data collection for the "screener" and "followup" techniques. Different modes of data collection may bring with them different measurement error patterns across modes. If this were true, then the "nonresponse bias" estimates using the screener and followup methods may actually be the net magnitude of nonresponse bias and mode bias differences.

From a different perspective, the higher nonresponse differences for screener and followup studies may be real. *Survey* variables (versus frame or supplemental data variables) sometimes act as influencers for the participatory decision. If the knowledge of what is to be measured influences cooperation, then nonresponse biases could result. Persons with different values on the variables could vary in their response propensities, producing biased respondent estimates.

The value of these two methods of estimating nonresponse bias (screener and followup) is that they permit bias estimates on the items in the survey questionnaire themselves. The weakness of the frame and supplementary data techniques is that the measures on which nonresponse bias is estimated are, in general, not the focus of the survey (i.e., if the purpose of the survey were to gather data solely on those variables, the records would have been used directly!).

Indeed, the apparent sensitivity of bias estimates to the method used to estimate them displayed in Figure 4 may be an important finding of the meta-analysis, and one that deserves further exploration. For later analyses of the data set presented below, we have decided to pool across the four methods. For each, however, we have replicated the analysis on the combined "frame" and "supplemental data" studies only. When discrepancies arose between the full data set and those two methods, we note them below.

INFLUENCES ON RESPONSE RATES

The survey methodological literature is replete with techniques to increase response rates; for example, prenotification (de Leeuw, Hox, Korendijk, Lensvell-Mulders, Callegaro, 2006) and incentives (Singer, 2002). From a conceptual point of view, few if any of these techniques should have general value of reducing nonresponse differences on all types of survey estimates. However, there is some indication of tendencies for poorer persons to be more sensitive to incentive effects (and hence, measures of socioeconomic status might have larger nonresponse bias with vs. without incentives).

Table 1 shows the average difference between respondents and nonrespondents, using the 566 standardized percentage estimates in the 44 studies. The estimates from studies using prenotification have a mean difference of .11 standard deviations; those not using prenotification, .13. The standard error of the difference of .019 is .028. Although we know that prenotification acts to increase response rates, it appears not to be associated, in general, with the magnitude of differences between respondents and nonrespondents on the survey variables. Another way of stating the result is that attributes measured in the survey are largely unrelated to prenotification effects.

We urge readers not to overinterpret this failure to reject the null hypothesis. It would be erroneous to infer that prenotification *never* changes the nonresponse bias of any variables; the variation over different estimates within the meta-analysis is very large and contributes to the large standard errors above. For example, we can imagine prenotification emphasizing specific purposes of the survey leading to more biases on items most closely-related to those purposes.

Table 1. Weighted Absolute Nonresponse Differences, $|\bar{y}_r - \bar{y}_m|$, by Nonresponse Rate of Study by Subgroup for Standardized Percentage Estimates

	Average Absolute Differences between Standardized Respondent and Nonrespondent Mean		
	$\lvert\bar{y}_r - \bar{y}_m\rvert$	(Std. Error)	n
Prenotification			
Yes	0.11	(0.010)	*229*
No	0.13	(0.026)	*337*
Difference	-0.019	(0.028)	*566*
Incentives			
Yes	0.16	(0.069)	*48*
No	0.11	(0.011)	*518*

Difference	0.050	(0.070)	566
Respondent's Involvement with Sponsor			
Yes	0.093	(0.0051)	*200*
No	0.14	(0.021)	366
Difference	-0.052**	(0.022)	566
Sponsorship			
Government	0.15	(0.026)	*203*
Other	0.10	(0.0010)	*363*
Difference	0.048*	(0.028)	*566*
Mode			
self-administered	0.10	(0.010)	*275*
interviewer-administered	0.14	(0.021)	*291*
Difference	-0.040*	(0.024)	*566*
Topic			
Health	0.11	(0.017)	*358*
Other	0.13	(0.013)	*208*
Difference	-0.018	(0.023)	*566*
Population Type			
General	0.17	(0.032)	*159*
Specific	0.10	(0.0059)	*407*
Difference	0.075**	(0.033)	*566*
Urbanicity			
Urban	0.11	(0.014)	*224*
Mixed	0.12	(0.021)	*342*
Difference	-0.013	(0.027)	*566*
Subculture			
Majority	0.11	(0.018)	*236*
Other	0.14	(0.052)	*330*
Difference	-0.030	(0.061)	*566*
Question Type			
1.behavioral	0.11	(0.019)	*304*
2.attitudinal	0.24	(0.012)	*25*
3.demographic	0.11	(0.010)	*237*

1-2 difference	-0.14***	(0.022)	*329*
1-3 difference	-0.0081	(0.019)	*541*
2-3 difference	0.13***	(0.014)	*262*
Topic Saliency			
1.yes	0.10	(0.017)	*270*
2. no	0.17	(0.061)	*53*
3. undetermined	0.11	(0.013)	*243*
1-2 difference	-0.075	(0.064)	*323*
1-3 difference	-0.012	(0.021)	*513*
2-3 difference	0.063	(0.063)	*296*
Statistic's Relevance to Topic			
Yes	0.12	(0.023)	*315*
No	0.11	(0.0089)	*251*
Difference	0.06	(0.023)	*566*

***p<.10, ** p<.05, ***p<.01**

Similar results apply to incentives -- use of the incentive is not reliably related to the magnitude of nonresponse differences. (We note that few studies offered incentives and hence, the standard errors of the contrast are large.)

SPONSORSHIP OF THE SURVEY

Sponsors of the survey are often policy-makers or advocates for the topics of the surveys they sponsor (e.g., central governments measure employment *and* execute policies on economic welfare; companies conduct customer satisfaction surveys *and* manage the service contacts with customers). When the sample person judges that the sponsor has an identifiable "point of view" on the survey topic, reaction to that point of view by the sample persons can prompt them to respond or refuse. Sample persons who have prior connections with the sponsor are most likely to experience these influences. For survey variables that are related to that point of view, nonresponse bias can result. Measures unrelated to engagement in the organization can be relatively immune to nonresponse bias.

We rated the surveys relative to whether the entire sample had had prior involvement with the sponsor. Examples in the data set of surveys with prior involvement include a satisfaction and general health survey by the Hospital of the University of Pennsylvania mailed to patients who had undergone total knee arthroplasty (Kim, Lonner, Nelson and Lotke, 2004) and a survey sent via e-mail to subscribers to a computer network managed by the survey sponsor (Walsh, Kiesler, Sproull and Hesse, 1992). In the Table 1 analysis, when the full sample had prior involvement with the survey sponsor, on average there were lower nonresponse biases (difference of .05 of a standard deviation, p=.05).

Another sponsorship attribute is what sector of the society the sponsor represents. In many countries of the world surveys sponsored by governments tend to achieve higher response rates than those of other sponsors (deLeeuw and de Heer, 2002). In the data, some examples of government surveys include the Fundus Photography component of the US National Health and Nutrition Examination Survey (Khare, Mohadjer, Ezzati-Rice and Waksberg, 1994) and a study on nutrition and health conducted in Roskilde, Denmark (Osler and Schroll, 1992). The meta-analytic finding is that government-sponsored surveys tend to generate larger nonresponse differences (mean .15) than do other sponsors (mean .10). The difference is beyond traditional ($p < .05$) levels of statistical significance; however, when only the studies using frame or supplemental data are examined, the difference disappears.

MODE OF DATA COLLECTION

Some self-administered modes (e.g., a mailed paper questionnaire survey) permit the sample person to examine the questions prior to making the participatory decision. The inspection of questions prior to the decision to participate permits survey content to influence response propensities. Mechanisms facilitating that influence may include negative emotions connected to the topic (e.g., fear of revelation of socially undesirable traits) or assessment of high burden of the questions (e.g., construction of complicated reports of past behaviors, lookup of household records). If persons with such reactions tend to have different distributions on the y variables, then nonresponse bias should be induced by the decision-making circumstances.

Conversely, interviewers are commonly trained to emphasize the purpose of the survey and, under informed consent procedures common to many surveys, to describe the general content of the questions. Further, when interviewers tailor their remarks to the concerns of the sample person, they often try to relate the topic of the survey to the concerns of the respondent. In short, there are reasons to expect differences in either direction.

Table 1 shows that interviewer administered surveys generate on average a .14 standard deviation difference between respondents and nonrespondents and self-administered, a .10 difference (the .04 difference between the two estimates has a standard error of .024, $p<.10$).

TOPIC OF THE SURVEY

Health surveys often attain higher response rates than surveys on other topics; electoral behavior surveys commonly have lower response rates (e.g., Voogt and Van Kempen, 2002). We coded the surveys as either health-related topics or something else. In Table 1, there was no reliable difference between the two types of surveys in the nonresponse differences their variables displayed. (When the frame and supplemental data methods are examined, health surveys display lower average nonresponse differences than surveys on other topics.)

Attributes of the Sample Population

TYPE OF SAMPLE POPULATION

Some of the surveys use general population sampling frames; others are specific to members of an organization, students of a school, patients of a hospital, etc. It is very common for surveys of such specific populations to generate higher response rates, *ceteris paribus*. Further, it is common to note that membership survey respondents tend to be more attached to the organization than the nonrespondents (Rogelberg, Luong, Sederburg, and Cristol, 2000). Because general population surveys usually do not have rich frames to study such nonresponse tendencies, the literature tends not to have such findings documented. Hence, the impression that membership surveys tend to suffer from unusually large nonresponse biases may be fallacious.

Table 1 shows that surveys of the general population tend to generate larger average nonresponse differences (mean = 0.17) than surveys of specific populations (mean = .10). The difference of .075 has a standard error of .033 ($p<.05$). Estimates from general population surveys have higher nonresponse differences than those from specific population using frame or supplemental data also, but the contrast is not as sharp.

URBANICITY OF TARGET POPULATION

One of the most common correlates of response rates in household surveys is the urbanicity of the population sampled (e.g., House and Wolf, 1979; Steeh, 1981). The mechanisms that produce this correlation are not well-understood. First, many of the other attributes related to low response rates tend to cluster in cities (e.g., single person households, households without children). Second, social psychologists have observed that the pace of urban life, filled with fleeting, superficial interactions with strangers, sharply contrasts with the deeper, multidimensional relationships with others that exist in non-urban settings. Little of the past research identifies mechanisms that might link these concepts to what types of survey variables might be subject to nonresponse bias induced from urban-rural contrasts.

We coded the studies by whether they sampled only urban populations or mixed populations (we had no examples of purely rural samples). Table 1 shows no differences between the two groups in the nonresponse differences on survey variables. We remind the reader, however, that this finding is subject to rather high imputation variance of the urbanicity variable.

SUBCULTURES REPRESENTED IN THE SURVEY

A common speculation in surveys is that racial and ethnic minorities tend to have lower response rates than majority groups (Brehm, 1993). There is some evidence that this results from unusually high noncontact rates and specific types of survey content (Groves and Couper, 1998). If survey variables tend to be correlated with minority status, then we might hypothesize that surveys studying majority populations would have higher nonresponse differences than those focusing on minority subcultures.

Table 1 shows that the two types of surveys yield similar average nonresponse differences on survey variables. We note, however, that imputation variance on this predictor variable is quite high.

Attributes of the Survey Estimates

Figure 2 shows that most of the variation in nonresponse bias of survey estimates lies within surveys, across estimates. Given the dependence of bias on the relationship between response propensities and the survey variable, it seems clear that attributes of individual estimates must play a part in the explanation of nonresponse bias. There are several lines of argument present in the past literature.

MEASUREMENT OF SUBJECTIVE VERSUS OBJECTIVE PHENOMENA

There are speculations that statistics based on attitudinal measures might be more subject to nonresponse bias than those based on objective phenomena (Stinchcombe, Jones, and Sheatsley, 1981). This might flow from a perspective that a set of attitudinal states influences the survey participation decision and when attitudes are measured in the survey, they tend to be correlated with those attitudes. If behaviors are measured, the reasoning continues, they would be less correlated with the attitudes influencing participation.

We find that behavioral measures have lower average biases than measures of nonobservable attributes (by .12 standard deviation, p=.10). The attitudinal variables tend to come from studies using screener variables or followup studies comparing early and late responders. There are only 25 estimated attitudinal measures, and despite the standard errors indicating reliable differences, we are concerned that the results disproportionately come from one type of design, which itself tends to have higher mean nonresponse biases.

TOPIC INTEREST OR SELF-INTEREST RELATED TO PARTICIPATION

Some persons, when the topic of the survey is made salient in the request for participation, become positively disposed because the topic itself is of interest to them (Groves, Presser, Dipko, 2004). They have learned through experience that cognitive engagement in the topic brings them some satisfaction. This could be because it stimulates a set of memories that are pleasurable – achievements in social, educational, or occupational terms. It could be because the topic permits them to engage in a display of knowledge about a field – political affairs, the state of their community or nation. It could be because the topic concerns an activity that the persons willingly engage in repeatedly in their lives and they have a large set of well-rehearsed memories that are stimulated by the mention of the topic. Because this reaction of self-interest in the topic is influential for these persons' decision to participate, the respondent pool tends to be disproportionately "interested" persons.

When interest is correlated with different distributions on the variable, nonresponse bias can be induced. For the variables concerning the stated topic in the questionnaire, the respondent distribution is likely to be different from the nonrespondent distribution. (Conversely, for variables thematically unrelated to the stated topic, no such nonresponse bias influences are expected.)

We created two ratings relevant to these hypotheses: a) is the topic of general interest to the population studied (e.g., health symptoms among a sample of patients as in Marcera, Jackson, Davis, Kronenfeld and Blair, 1990), and b) is the estimate measured on a variable that is key to the topic of the survey (e.g., percent drinking less than once a week in a survey on alcohol use as in Wild, Cunningham and Adlaf, 2001). Neither on the study-level test of this nor the estimate-level test supported this hypothesis. (Indeed, when just the frame and supplemental data studies are included, the estimate-level analysis shows *lower* average nonresponse differences for items judged relevant to the topic!)

Statistical Attributes of Estimates

It is common to hope that biases "cancel out" when measures of relationships or differences over time in a statistic are examined (Cochran, 1977, pp. 379-380; Martin, 2004; Goudy, 1976). For example, when the mean of subclass 1, $\bar{y}_{r1}$, is biased and the mean of subclass 2, $\bar{y}_{r2}$, is biased, it is hoped that the difference of the two subclass means, $\bar{y}_{r1} - \bar{y}_{r2}$, enjoys some canceling of biases. For example, consider the estimated difference between the percentage of females reporting excellent health and percentage of males reporting excellent health. If one asserts that each subclass mean is subject to the same nonresponse bias, then the difference of means is free of nonresponse bias. This hope taken to the extreme leads some practitioners to believe that if their survey statistics are comparisons of subgroups (or measures of relationships between variables), their analysis is immune to nonresponse bias.

Some of the studies assembled have nonresponse bias estimates for subclasses, and the appropriate documentation for computation of the bias of the difference of subclass means. Not all of the articles collected present subclass estimates, but there are 234 subclass means that are reported that also have estimated nonresponse biases. One statistical version of the speculation is that the difference of the subclass means would have an absolute value of nonresponse bias that would be larger than the *average* absolute value of nonresponse bias of the two constituent subclass means.

Figure 5 is a scatterplot of 117 estimates -- each comparison of two subclass means (of, say, $\bar{y}_{r1} - \bar{y}_{r2}$) contributes one point to the scatterplot. The *x-axis* is the mean of the absolute value of the biases of the two subclass means (the average bias of the subclass means). The *y-axis* value is the nonresponse bias of the difference of the two subclass means. To ease the reading of the plot, the line representing $y = x$ is placed into the plot, corresponding to the case when the bias of the contrast between the two subclass means equals the average bias of the two subclass means themselves. Points below the diagonal line are those where the differences of subclass means have lower nonresponse bias than the subclass means, on average. This is desired state for all researchers doing analytic comparisons.

Of the 117 differences of subclass means in Figure 5, only 41 have lower nonresponse biases than those of the contrasted subclass means. (The ridge of values on the *y* dimensions arises when the subclass means have opposite signs.) In short, there seems little evidence biases tend to cancel when comparing two subclasses.

Why doesn't the rule of thumb always work? There are reasons for the biases of the two subclass means to vary. Returning to the expression, $\sigma_{yp} / \bar{p}$, it is possible that both the response rates and the covariances are different for the two subclasses. For example, many of the comparisons published in the articles are contrasts of male and female means on some health variable. The bias of the male mean tends to be a different sign or magnitude than that of the female mean, perhaps reflecting different response rates and/or different causes of nonresponse.

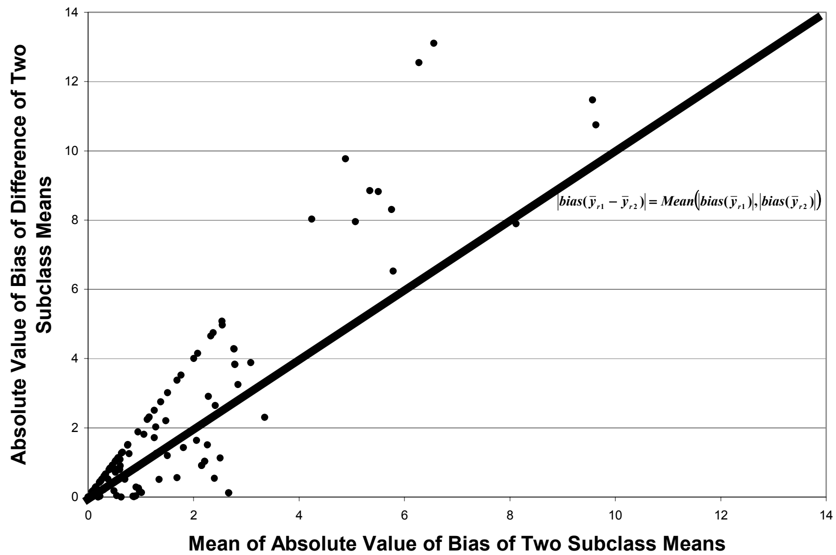

Figure 5. Absolute Value of Bias of 117 Differences of Two Subclass Means , $\left|Bias(\bar{y}_{r1} - \bar{y}_{r2})\right|$, by Mean of Absolute Value of the Biases of the Two Contrasted Subclass Means, $Mean\left\{\left|Bias(\bar{y}_{r1})\right|, \left|Bias(\bar{y}_{r2})\right|\right\}$

Conclusions and Discussion

As with all meta-analyses, conclusions must be made with considerable caution. Meta-analyses may be good tools when single variables influence a phenomenon, but rarely do they have complete representation of possible patterns of predictor variables to support complex multivariate analysis. We are very concerned about undetected confounds in the set of published articles we were able to assemble. Thus, we order our conclusions by the strength of our belief that they will withstand replication attempts.

1. Large nonresponse biases can happen in surveys.

   High response rates can reduce the risks of bias. They do this less when the causes of participation are highly correlated with the survey variables. Indeed, in the studies we assembled, some surveys with low nonresponse rates have estimates with high relative nonresponse bias.

2. The search for mechanisms that link nonresponse rates and nonresponse bias should focus on the level of individual measures not on the level of the survey.

   The meta-analysis shows much variability in nonresponse differences within surveys, across estimates. We know from statistical expressions that influences on survey participation that are themselves measured in the survey will show the largest nonresponse bias. To predict what survey estimates are most susceptible to nonresponse bias, we need to understand how each survey variable relates to causes of survey participation.

3. Differences of subclass means do not, in general, enjoy lower nonresponse biases than their constituent subclass means.

   We cannot rely on full or partial canceling of nonresponse biases when we subtract one subclass mean from another. The bias of the difference is a function of differences of response rates and covariances between response propensities of the subgroups and the survey variable. Too many components have to align themselves in a beneficial way for the desired outcome to occur.

4. How we estimate nonresponse bias may make a difference.

   We found that nonresponse differences in the literature tend to be higher when screener data and data from followup efforts are used (relative to using frame or supplemental data sources). These techniques try to estimate bias on the survey variables as actually measured in the data collection. Thus, they are informative about bias in the key survey estimates themselves. Further, when knowledge of the survey items influences the decision to participate, our theory predicts larger nonresponse biases on those items. On the other hand, the screener and followup methods often employ different modes of data collection or other changes in measurement conditions. Given the documentation of studies in the literature, we can’t easily separate the measurement errors from the nonresponse errors.

Given the uniqueness of this meta-analytic data set, we offer the reader some further cautions and suggestions. First, we observed high correlations among a set of attributes of the assembled studies a) using screener or followup technique to estimate bias, b) studying general populations, c) having government sponsorship, and d) not having prior involvement of the target population with the sponsor. All of these attributes relate to higher nonresponse differences and therefore nonresponse biases. Unfortunately, the meta-analytic data set has too few cases with variation on those four attributes. We need more studies with different mixes of these attributes to have more confidence in the statistical findings of Table 1.

Second, our theoretical framework suggests one linkage between nonresponse rates and bias that should lie at the estimate-level. When the estimate is based on survey items highly relevant to the topic of the survey as presented at the time of recruitment *and* the topic was a very salient attribute of the survey request, then the conditions for nonresponse bias should exist. Unfortunately, the documentation about recruitment protocols is almost nonexistent in the printed literature we found. Hence, we remain cautious about the findings involving our coding of estimates on relevance to the topic. Gathering more documentation on the nature of the survey introduction to the sample could be of benefit.

Third, even though the nonresponse differences between interviewer-administered surveys versus self-administered surveys were only marginally significant, we find this result very intriguing. We believe this finding may itself be a function of the nature of how interviewers recruited the respondents and what information guides compliance with self-administered survey requests. That is, we suspect different mechanisms may produce the covariance between response propensities and survey variables in the two modes. This is a rich area for study, given our field's movement to mixed mode surveys.

References

Bradburn, Norman (1992) "A Response to the Nonresponse Problem," *Public Opinion Quarterly*, Vol. 56, pp. 388-418.

Brehm, John (1993) *The Phantom Respondents: Opinion Surveys and Political Representation*, Ann Arbor, University of Michigan Press.

Campanelli, P., Sturgis, P. and Purdon, S. (1997) *Can You Hear Me Knocking: An Investigation into the Impact of Interviewers on Survey Response Rates*. London: S.C.P.R.

Cochran, William G. (1977) *Sampling Techniques*, 3rd Edition, New York: Wiley.

de Leeuw, Edith and De Heer, Wim. 2002. Trends in Household Survey Nonresponse: A Longitudinal and International Comparison. In *Survey Nonresponse*, edited by R. Groves, D. Dillman, J. Eltinge and R. J. A. Little. New York: Wiley.

de Leeuw, Edith, Hox, Joop, Korendijk, Elli, Lensvell-Mulders, Gerty, Callegaro, Mario (2006) "The Influence of Advance Letters on Response in Telephone Surveys: a Meta-Analysis," paper presented at the International Conference on Telephone Methodology, January.

Goudy, Willis (1976) "Nonresponse Effects on Relationships Between Variables," *The Public Opinion Quarterly*, Vol. 40, pp. 360-369.

Goyder, John C. (1987), *The Silent Minority: Nonrespondents on Sample Surveys,* Boulder: Westview Press.

Groves, Robert M. (2005) "Research Synthesis: Nonresponse Rates and Nonresponse Bias in Household Surveys," ms.

Groves, Robert M., and Couper, Mick P. 1998. *Nonresponse in Household Inteview Surveys.* New York: John Wiley and Sons.

Groves, Robert M., Couper, Mick P., Presser, Stanley, Singer, Eleanor, Tourangeau, Roger, Piani Acosta, Giorgina, Nelson, Lindsey (2006) "Experiments in Producing Nonresponse Error," unpublished ms.

Groves, Robert M., Presser, Stanley and Dipko, Sarah. 2004. The Role of Topic Interest in Survey Participation Decisions. *Public Opinion Quarterly* 68 (1):2-31.

Groves, Robert M., Singer, Eleanor and Corning, Amy. (2000) "Leverage-Saliency Theory of Survey Participation - Description and an Illustration," *Public Opinion Quarterly* 64 (3):299-308.

House, James, and Wolf, Sharon (1979) "Effects of Urban Residence on Interpersonal Trust and Helping Behavior," *Journal of Personality and Social Psychology*, 36, pp. 1029-1043.

Keeter, Scott, Miller, Carolyn, Kohut, Andrew, Groves, Robert M., and Presser, Stanley. 2000. "Consequences of Reducing Nonresponse in a National Telephone Survey," *Public Opinion Quarterly* 64, pp.125-148.

Khare, Meena, Mohadjer, Leyla K., Ezzati-Rice, Trena, and Waksberg, Joseph (1994) "An Evaluation of Nonresponse Bias in NHANES III (1988-91)," *Proceedings of the section on Survey Research Methods, American Statistical Association,* pp. 949-954.

Kim, Jane, Lonner, Jess H., Nelson, Charles L., and Lotke, Paul A. (2004) "Response Bias: Effect on Outcomes Evaluation by Mail Survey After Total Knee Arthroplasty," *Journal of Bone and Joint Surgery*, 86A:1, pp. 15-21.

Lahaut, Vivienne M.H.C.J., Jansen, Harrie A.M., Van de Mheen, Dike and Garretson, Henk F.L. (2002) "Non-response Bias in a Sample Survey on Alcohol Consumption," *Alcohol and Alcoholism*, 3, pp. 256-260.

Lin, I-Fen; Schaeffer, Nora Cate (1995) "Using Survey Participants to Estimate the Impact of Nonparticipation," *The Public Opinion Quarterly*, Vol. 59, pp. 236-258.

Macera Caroline A, Jackson, Kirby L, Davis Dorothy R, Kronenfeld, Jennie J. and Blair, Steven N. (1990) "Patterns of non-response to a mail survey". *J Clin Epidemiol*.;43:1427–1430.

Martin, Elizabeth (2004) "Unfinished Business," *Public Opinion Quarterly*, 68, pp. 439-450

Osler, Merete, Schroll, M. (1992) "Differences between participants and nonparticipants in a population study on nutrition and health of the elderly," European Journal of Clinical Nutrition. 46, pp.189-95.

Raghunathan, Trivellore, Lepkowski, James, VanHoewyk, John, and Solenberger, Peter (2001) "A multivariate technique for multiply imputing missing values using a sequence of regression models" *Survey Methodology*, Vol. 27, pp. 85-95.

Raghunathan, Trivellore, VanHoewyk, John, and Solenberger, Peter (2002) *IVEware: Imputation and Variance Estimation Software User Guide,* Ann Arbor: Survey Research Center.

Rogelberg, Steven, Luong, Alexandra, Sederburg, Matthew, and Cristol, Dean (2000) "Employee Attitude Surveys: Examining the Attitudes of Noncompliant Employees," *Journal of Applied Psychology*, 85, pp. 284-293.

Rubin, Donald B. (1987) *Multiple Imputation for Nonresponse in Surveys,* New York: John Wiley and Sons.
Singer, Eleanor (2002) "The Use of Incentives to Reduce Nonresponse in Household Surveys," in *Survey Nonresponse*, edited by R. M. Groves, D. A. Dillman, J. L. Eltinge and R. J. A. Little. New York: Wiley.
Steeh, Charlotte (1981) "Trends in Nonresponse Rates, 1952-1979" *The Public Opinion Quarterly*, 45, pp. 40-57.
Stinchcombe, Arthur, Jones, Calvin, and Sheatsley, Paul (1981) "Nonresponse Bias for Attitude Questions," *The Public Opinion Quarterly*, Vol. 45, pp. 359-375
Walsh, John P., Kiesler, Sara, Sproull, Lee S., and Hesse, Bradford W. (1992) "Self-selected and Randomly Selected Respondents in a Computer Network Survey," *Public Opinion Quarterly*, 56, pp. 241-244.
Wild, Cameron T., Cunningham, John and Adlaf, Edward (2001) "Nonresponse in a follow-up survey to a representative telephone survey of adult drinkers," *Journal of Studies on Alcohol*, 62, pp.257-261.

APPENDIX A

References from Meta-Analysis of Nonresponse Bias Studies

Assael, H., and Keon, J. (1982) "Nonsampling vs. Sampling Errors in Survey Research," *Journal of Marketing,* 46, pp. 114-123
Axelsson G, Rylander R. (1982) "Exposure to anaesthetic gases and spontaneous abortion: response bias in a postal questionnaire study," Int J Epidemiol ;11:250–6.
Barchielli, A., and Balzi, D. (2002) "Nine-year Follow-up of a Survey on Smoking Habits in Florence (Italy): Higher Mortality Among Non-Responders," *International Journal of Epidemiology*, 31, pp. 1038-1042
Benfante, R., D. Reed, C. MacLean and Kagan, A. (1989) "Response bias in the Honolulu Heart Program. Am J of Epidemiology," 130: 1088-1100
Bergstrand R, Vedin A, Wilhelmsson C, et al. (1983) "Bias due to nonparticipation and heterogeneous sub-groups in population surveys," J Chronic Dis;36:725–8.

Bolstein, Richard (1991) “Comparison of the Likelihood to Vote Among Preelection Poll Respondents and Nonrespondents,” *Public Opinion Quarterly*, 55, pp. 648-650

Cohen, G. and Duffy, J. (2002) “Are Nonrespondents to Health Surveys Less Healthy Than Respondents?” *Journal of Official Statistics*, 18, pp. 13-23

Criqui, M., Barrett-Connor, E. and Austin, M. (1978) “Differences between respondents and non-respondents in a population-based cardiovascular disease study,” *American Journal of Epidemiology,* 108, pp. 367-372

Drew, James, and Groves, Robert, (1989) “Adjusting For Nonresponse In A Telephone Subscriber Survey,” *Proceedings of the American Statistical Association, Survey Research Methods Section,” pp. 452-456*

Etter, J-F., and Perneger, T. (1997) “Analysis of Nonresponse Bias in a Mailed Health Survey,” Journal of Clinical Epidemiology, 50:10, pp. 1123-1128

Gershen, J.A. and McCreary, C.P. (1983) “Personality comparisons of responders and nonresponders to a mailed personality inventory,” Psychology Reports, 52, pp. 555-562.

Goldberg, M., Chastang, J., Leclerc, A., Xins, M., Bonenfant, S., Bugel, I., Kaniewski, N., Schmaus, A., Niedhamer, I, Piciotti, M., Chevalier, A., Godard, C., and Imbernon, E. (2001) “Socioeconomic, Demographic, Occupational, and Health Factors Association with Participation in a Long-term Epidemiologic Survey: A Prospective Study of the French GAZEL Cohort and Its Target Population,” *American Journal of Epidemiology*, 154, pp. 373-384

Grosset, J (1994), "The Biasing Effects of Nonresponses on Information Gathered by Mail Surveys," Institutional Report 78, Philadelphia:Community College of Philadelphia

Grotzinger, K. M, B.C. Stuart, and F. Ahern. (1994) “Assessment and Control of Nonresponse Bias in a Survey of Medicine Use by the Elderly,” Medical Care, 32, 989-1003.

Hudson, D., Seah, L-H., Hite, D., and Haab, T. (2004) “Telephone Presurveys, Self-Selection, and Non-response Bias to Mail and Internet Surveys in Economic Research,” *Applied Economics Letters*, 11, pp. 237-240

Hutchinson, J., N. Tollefson, and H. Wigington. (1987) “Response bias in college freshmen’s responses to mail surveys,” Research in Higher Education, 26(1), pp. 99-106.

Kendrick, D., Hapgood, R., and Marsh, P. (2001) “Do Safety Practices Differ Between Responders and Non-Responders to a Safety Questionnaire,” *Injury Prevention*, 7, pp. 100-103.

Kennickell, Arthur, and McManus, D. (1993) “Sampling for Household Financial Characteristics Using Frame Information on Past Income,” *Proceedings of Survey Research Methods Section of the ASA, pp 88-97*

Khare, Meena, Mohadjer, L., Ezzati-Rice, Trina, and Waksberg, Joseph (1994) "An Evaluation of Nonresponse Bias in NHANES III (1988-91)," *Proceedings of the section on Survey Research Methods, American Statistical Association,* pp. 949-954

Kim, J., Lonner, J., Nelson, C., and Lotke, P. (2004) "Response Bias: Effect on Outcomes Evaluation by Mail Survey After Total Knee Arthroplasty," *Journal of Bone and Joint Surgery*, 86A:1, pp. 15-21

Kleven, Oyvin and Tor M. Normann (2005) “Living with Bias in Political Surveys: Findings from the Norwegian Election Surveys 1997 and 2001”, Paper presented at the Nonresponse Workshop in Tallberg, Sweden, 2005.

Lahaut, V.M.H.C.J., Jansen, H.A.M., Van de Mheen, D and Garretsen, H.F.L. (2002) “Non-response bias in a sample survey on alcohol consumption,” *Alcohol & Alcoholism,* 3, pp. 256-260

Lasek, Rebecca, W. Barkley, D.L. Harper and G. Rosenthal. 1997. An Evaluation of the Impact of Nonresponse Bias on Patient Satisfaction Surveys. Medical Care, Vol. 35 (6), June, pp.646-652

Lemmens, P. H., S. E. Tan and R.A. Knibbe (1988) “Bias due to NR in a Dutch survey on alcohol consumption,” British J of Addiction 83, 1069-1077.

Lin, I-F. and Schaeffer, Nora (1995) “Using Survey Participants to Estimate the Impact of Nonparticipation,” *Public Opinion Quarterly*, 59, pp. 236-258.

Macera Caroline A, Jackson, Kirby L, Davis Dorothy R, Kronenfeld, Jennie J. and Blair, Steven N. (1990) “Patterns of non-response to a mail survey”. *J Clin Epidemiol.*;43:1427–1430.

McNutt, L-A., and Lee, R. (2000) “Intimate Partner Violence Prevalence Estimation using Telephone Surveys: Understanding the Effect of

Nonresponse Bias," *American Journal of Epidemiology*, 152, pp. 438-441.

Melton, L., Dyke, D., Karnes, J., and Obrian, P. (1993) "Nonresponse Bias in Studies of Diabetic Complications: The Rochester Diabetic Neuropathy Study," *Journal of Clinical Epidemiology*, 46, pp. 341-348

Messonnier, M., Bergstrom, J., Cornwell, C., Teasley, R., and Cordell, H. (2000) "Survey Response-Related Biases in Contingent Valuation: Concepts, Remedies, and Empirical Application to Valuing Aquatic Plant Management," *American Journal of Agricultural Economics*, 83, pp. 438-450

Norton MC, Breitner JC, Welsch KA, et al. (1994) "Characteristics of nonresponders in a community survey of the elderly," *J Am Geriatr Soc*.;42:1252–1256.

Osler, Merete, Schroll, M. (1992) "Differences between participants and nonparticipants in a population study on nutrition and health of the elderly," Eur Journal of Clinical Nutrition. 46, pp.189-95.

Paganini-Hill, A., Hsu, G., Chao, A. and Ross, R. (1993) "Comparison of Early and Late Respondents to a Postal Health Survey Questionnaire," *Epidemiology*, 4, pp. 375-379

Pedersen, P. (2002) "Non-Response Bias – A Study Using Matched Survey-register Labour Market Data," Working Paper 02-02, Aarhus, Denmark: Centre for Labour Market and Social Research

Perneger, T., Chamot, E., and Bovier, P. (2005) "Nonresponse Bias in a Survey of Patient Perceptions of Hospital Care," *Medical Care*, 43, pp. 374-380

Potter, D. (1989) "Nonresponse in a Survey of Nursing Home Residents," *Proceedings of the American Statistical Association, Survey Research Methods Section,* pp. 440-445

Reigneveld, S., and Stronks, K. (1999) "The Impact of Response Bias on Estimates of Health Care Utilization in a Metropolitan Area: the Use of Administrative Data," *International Journal of Epidemiology*, 28, pp. 1134-1140

Rockwood, K., P. Stolee, D. Robertson, and E.R. Shillington (1989) "Response Bias in a health status survey of elderly people," Age and Ageing. 18, pp.177-82.

Rogelberg, S., Luong, A., Sederburg, M., and Cristol, D. (2000) "Employee Attitude Surveys: Examining the Attitudes of

Noncompliant Employees," *Journal of Applied Psychology*, 85, pp. 284-293

Ronmark E, Lundqvist A, Lundback B, Nystrom L. (1999) "Non-responders to a postal questionnaire on respiratory symptoms and diseases," Eur J Epidemiol; 15:293–99.

Sheikh, K., and Mattingly, S. (1981) "Investigating Non-response Bias in Mail Surveys,"*Journal of Epidemiology and Community Health*, 35, pp. 293-296

Strayer M, Kuthy R, Sutton S. (1993) "Elderly nonrespondents to a mail survey: a telephone follow-up," *Spec Care Dentist.*;13:245-8.

Teitler, J., Reichman, N., and Sprachman, S. (2003) "Costs and Benefits of Improving Response Rates for a Hard-to-Reach Population," *Public Opinion Quarterly*, 67, pp. 126-138.

Van Kenhove, P., Wijnen, K., and De Wulf, K. (2002) "The Influence of Topic Involvement on Mail-Survey Response Behavior," *Psychology and Marketing,* 19, pp. 293-301

Voogt, R., and van Kempen, H. (2002) "Nonresponse Bias and Stimulus Effects in the Dutch National Election Study," *Quality and Quantity,"* 36, pp. 325-345.

Walsh, John P., Kiesler, Sara, Sproull, Lee S., and Hesse, Bradford W. (1992) "Self-selected and Randomly Selected Respondents in a Computer Network Survey," *Public Opinion Quarterly*, 56, pp. 241-244.

Wild, Cameron T., Cunningham, John and Adlaf, Edward (2001) "NR in a follow-up survey to a representative telephone survey of adult drinkers," Journal of Studies on Alcohol, 62, pp.257-261.

Wildner M. (1995) " Lost to follow-up," *J Bone Joint Surg Br.* ;77:657.