

Addressing Nonresponse and Nonresponse Bias Issues in Surveys

Note: *This guidance is designed to ensure that GAO policies on evidence and generally accepted government auditing standards are met. The guidance conforms to the generally accepted principles and practices of the appropriate disciplines. Statements that particular actions “should” be taken are practices that are expected to be followed, unless there are good reasons for not doing so. Before deviating from a practice expressed as a “should” statement, staff members must consult with an appropriate staff member in Applied Research and Methods (ARM) or a team specialist and must document the consultation.*

Abstract

You should take steps to increase response commensurate with the level of expected nonresponse. The higher the level of nonresponse anticipated, the more resources you should devote to a nonresponse bias analysis, and the more you should consider using multiple and stronger measures to evaluate any possible bias. This paper gives guidance and options on when and what kind of nonresponse bias analysis should be done, how to adjust for the effects of nonresponse in estimation, and how nonresponse affects reporting of results and methodology.

Principles

Nonresponse is a serious and growing concern for survey designers. In addition to reducing the number of observations behind survey findings and the statistical power of estimates, the possibility of material bias always exists – that those who do not answer a question would have answered differently from those that do answer, leading to estimates that differ from true population values. However, research suggests that bias and response rates are not necessarily highly correlated – serious bias may exist at relatively high rates of response, while little bias may exist in some surveys with low rates of response. Also, the presence or absence of bias is sometimes question-specific, not questionnaire-wide. When reporting on nonresponse, the threat of question-specific bias should be emphasized rather than an “acceptable” unit response rate.

Because some methods of detecting bias appear to be better than others, and there are a number of ways to increase response rates in survey design, and adjust for nonresponse in survey analysis, this guidance lays out the choices available to the analyst and makes some suggestions on how to proceed. The literature on this subject is still emerging, so the individual survey designer must make choices based on his or her perceptions of risks and benefits.

As one model to follow, [OMB issued guidance in 2006](#) for Executive Branch agencies that they should: a) take actions to increase response, b) measure, adjust for, report, and analyze the effects of unit and item nonresponse on data quality, c) plan a bias analysis if response rates below 80% are anticipated, and if item response is expected to be below 70%, do this at the item level, d) report weighted response rates in some surveys, e) adjust estimates for nonresponse, f) consider anticipated response rate in the decision to do a survey, although OMB states that there are no strict response rate requirements for this decision.

However, GAO surveys often differ from those of federal statistical agencies, so the OMB guidelines, while good survey practice, may not always be appropriate for GAO work, and we are not obligated to follow them here.

Dealing with nonresponse should occur at all stages of the survey: in planning and survey design, during fieldwork, in a nonresponse analysis phase, and in reporting estimates and methodology.

Planning

1. You should consider the likely level of nonresponse. (See Appendix 1.) If you anticipate a response rate considerably below 60%, you should reconsider the value of doing a survey, or get ARM management input first. For example, RDD telephone surveys or mail surveys with poor contact information may be at particular risk.
2. The lower the anticipated response rate, the more comprehensive the nonresponse reduction activities and bias assessment should be. Plan for nonresponse reduction activities and bias analysis if the anticipated response rate is below 80%. Be mindful of time and resource constraints when planning.
3. Nonresponse is often question-specific and you should also consider the likely level of nonresponse for specific survey items. Depending on the importance of individual estimates and the likelihood of low item response, you could consider: how the items could be evaluated for bias, possible adjustment strategies, or dropping items from the survey.
4. Some nonresponse bias analysis techniques might be more effective if a decision is made to use them before the survey begins. For example, if intensive followup to convert a sample of nonrespondents is considered, procuring the resources and planning the time for this up front could increase the chances of its success. You may also consider searching for reliable auxiliary data for your population that would allow for comparisons of sample respondents to population values.
5. ARM staff should discuss nonresponse issues with team management before finalizing survey design, and decisions should be documented. Team members should be made aware of the risks and implications of low response, how it will affect reporting, and the costs and benefits of proposed methods for reducing or adjusting for it.
6. Consider using the “[Survey Nonresponse Planning and Bias Assessment Document](#)” for creating and documenting a structured approach to nonresponse issues.

During the Survey

1. You should take action to increase response rate to the extent practicable. Consider a well-developed, multi-stage (and multi-mode) contact strategy, adding techniques for higher levels of nonresponse. For example, you could begin with mailed reminders and replacement questionnaires, then progress to telephone reminder and problem-solving followup calls using contractors, and add personal followup calls or visits by professional GAO staff to encourage response if response rates are still lagging.
2. Consider using other strategies for increasing response rate for a survey achieving less than an 80% response rate:
 - a. Research the specific causes of nonresponse in the survey, and take action to address them. For example, determine the noncontact rate due to poor address or email information, and correct contact information to the extent practicable.
 - b. Commission neutral but encouraging endorsement messages from third parties that can establish legitimacy or value of the GAO survey.
 - c. Conduct abbreviated questionnaire interviews to capture key variables if the burden of the full survey is a factor in nonresponse.
 - d. Offer monetary or nonmonetary incentives for response (monetary incentives may be feasible in only a subset of contractor-conducted surveys).
 - e. Attempt data collection with proxies or replacements for initially identified respondents within households or organizations. For example, redirect survey request to organization officials at a higher level.
 - f. Lengthen the field period if the burden of the survey tasks and seasonal availability of the respondent population is a factor in nonresponse.
 - g. Establish or strengthen confidentiality pledges (Note: requires pre-survey planning and approval) if sensitivity is a factor in nonresponse.
3. To aid later nonresponse bias analysis or adjustment, expand the capture of auxiliary information for as much of the sample as possible during nonresponse followup contacts. For example: characteristics observable during in-person interview attempts, or discoverable from an initial reminder call or attempt to complete the survey, even if unsuccessful.

Assessing Nonresponse Bias

1. You should conduct some kind of nonresponse bias analysis when response rates are below 80%. Consider examining nonresponse bias for response rates above 80% if the levels of importance and risk warrants – for example, if it is likely that differences between respondents and nonrespondents are large, or when making an estimate of a rare characteristic that could be influenced by even a small number of nonresponding cases.

2. The purpose of non-response analysis is to identify the potential for bias in survey estimates, and if possible, describe its direction and magnitude.
3. Multiple methods are preferable.
4. Some methods are more informative than others. In particular, time-of-return analysis and comparison of respondents to nonrespondents on variables (such as generic demographics) that are only weakly related to key questionnaire items may or may not be informative and should not be relied upon to rule out nonresponse bias. Respondents and nonrespondents can have similar demographic distributions and significant bias may still exist. On the other hand, differences between respondents and nonrespondents on a few demographic characteristics, may suggest that nonrespondents are not “missing at random,” and there may be similar differences on key survey variables. But there may be little or no bias after all if the key variables aren’t related to those demographics.
5. In the same vein, a common component of many nonresponse bias analysis techniques (see below) is comparing respondents to some other group – nonrespondents or the population as a whole – on some auxiliary variable that is thought to be related to some or all of the key survey items. For example, employee year of birth might be highly related to likelihood to retire in the next year. How informative that analysis can be is determined by the availability and quality of those related variables.
6. Consider the following analysis techniques. The decision of which technique(s) to use should be driven by the levels of importance and risk, the amount of available data, time and other resources available.
 - a. Compare respondents to the entire population on the values of known auxiliary variables which are related to the key survey items. Since the values of the auxiliary variables are known with certainty, or near-certainty, for both groups, differences between respondents and the rest of the population may be an indicator of the presence or absence of response bias, and may allow calculation of the size and direction of the bias.
 - b. Compare respondents to only the sample (or directly to the nonrespondents in the sample) on the values of known auxiliary variables. This is the same approach as in a), but it is not as strong because it is limited to the sample. This approach might be used if it is too costly to find the value of the auxiliary variables for every member of the entire population.
 - c. Compare respondents to only a subset of the sample, which in turn is comparatively less powerful than approaches a) and b).
 - d. Benchmarking - Compare survey estimates to similar ones from other sources, such as earlier studies or administrative data.

- e. Compare response rates across subgroups in the sample. This approach does not require finding auxiliary variable information for the sample; it relies only on subgroups defined by characteristics already in the sample. Since nonresponse bias results only when subgroups with different characteristics have different response rates, differences in response rates may indicate bias.
- f. Intensive followup to convert all of a subset of nonrespondents, to create a comparison group that is representative of other nonrespondents. This will only work if all or almost all nonrespondents in the pool can be converted. Differences between original respondents and converted nonrespondents on key variables are direct measures of bias. This method elevates the risk of upsetting frequently-contacted nonrespondents.
- g. Alternative weighting – compare estimates using the original base weights to the same estimates using nonresponse-adjusted weights. Nonresponse-adjusted weights are created by identifying those characteristics most related to nonresponse and applying them to subgroups. The assumption is that the respondents and nonrespondents have similar responses within these subgroups and large differences between them. Within subgroups, an adjustment factor is applied to the weights for the respondents to compensate for the nonrespondents

If there are large differences between original and adjusted estimates, it is possible that using the adjustment would reduce the bias in the estimates. If there are no differences, then it is possible that the respondent sample was not very different from the nonrespondents, and the concern over nonresponse bias is lessened.

- h. “What if” simulation creates upper and lower bounds on the potential bias by assuming that all the non-respondents differ from the respondents in a manner that would introduce the maximum possible bias into the estimates. This technique is most useful when responses to survey variables of interest are bounded. For example, when dealing with an attribute estimate (such as “47% of all firms are in Category A”) it is possible to assume that all of the non-respondents either possess or do not possess the attribute. By assuming they possess the attribute, an upper bound on the potential bias is created, and by assuming they all do not possess the attribute, a lower bound is created. If the differences between the estimate and the bounds are large then it is possible that the non-respondents could introduce significant bias. However, if the differences are small then the concern over non-response bias is lessened.
- i. “Time of return” or “Level of effort” analysis consists of comparing characteristics of early respondents to those responding very late or who respond only after intense follow up. The assumption in this type of analysis is that the late respondents are similar to the nonrespondents.

7. If response rates fall considerably below 60%, you should consider reporting the survey as based on a nonprobability sample, or not using it.
8. If some key questions have low item response rates (for example, under 70%), you should perform nonresponse bias analyses for those questions, or if item response rates fall considerably below 60%, you could consider not using the results from such questions.

Adjusting for Nonresponse

Consult an ARM/CDMA statistician to explore post-survey adjustment methods.

1. Weighting, an implicit form of imputation where it is assumed that all responses within a weighting class are missing at random, is one approach to adjustment.
2. Imputation, or replacing missing answers with values derived in a number of ways, could be considered to reduce bias or to improve accuracy and consistency.

Reporting on Nonresponse and Nonresponse Bias

1. The report should include the following response rate information (See guidance paper [“Calculating and Reporting Survey Response Rates”](#) for details.)
 - a. Weighted unit response rates for each stratum, and weighted overall response rate, if a stratified sample design was used (or whenever elements in a sample are selected with different probabilities) should be considered. In general, weighted response rates should be reported when applicable. Weighted response rates will differ from unweighted response rates when subpopulations with different weights have different response rates.
 - b. Unweighted unit response rates can also be reported in stratified samples, and should be reported for all other samples, following the definitions for [AAPOR RR1 or RR3](#).
 - c. A blanket statement of the range of item response rates for items included in the report.
 - d. Individual item response rates for items included in the report for items with item response rates below 70 percent.(The combined unit and item response rate immediately tells the reader what proportion of the sample answered any particular question.)

2. Nonresponse caveats should be made in the report, depending on the nature and extent of nonresponse, and the outcome of any nonresponse bias analyses. Some of these examples are very comprehensive and can be modified:

“In our analysis, we did detect a small amount of nonresponse bias among small public companies traded over the counter. We analyzed the result of this nonresponse on selected estimates. We concluded that the nonresponse did not affect our findings or conclusions.”

“Our analysis suggested that those small firms responding were different from those that did not, in terms of geography and number of clients. We were concerned that some small firms did not respond because the prospect of more auditing work for publicly held clients did not appeal to them and, thus, they found the survey request irrelevant. The small firms that responded could have answered the survey questions differently than the nonresponding small firms would have. As a result, reporting percentages based on responding small firms with one to four clients could introduce bias into results if those results were generalized to all accounting firms that audited at least one publicly traded company. Therefore our results are not generalizable to the entire population of small firms.”

And:

“After the survey was closed, we analyzed the survey respondents to determine if there were any differences between the responding districts, the nonresponding districts, and the population. We performed this analysis for three characteristics—total number of students enrolled, total number of special education students, and total number of English language learner students. We determined whether sample-based estimates of these characteristics compared favorably with the known population values. The population value for all of the characteristics we examined fell within the 95 percent confidence intervals for the estimates from the survey respondents. On the basis of the 77 percent response rate and this analysis, we chose to include the survey results in our report and produce sample-based estimates to the population of districts required to provide SES in the 2004-2005 school year.”

And:

“Because only 76 percent of the population provided usable responses, bias from nonresponse may result. If the responses of those who did not respond would have differed from the responses of those who did on some survey questions, the estimates made solely from those who did respond would be biased from excluding parts of the population with different characteristics or views. To limit this kind of error, we made multiple attempts to gain the participation of as many banks as possible. To assess the likelihood of significant bias, we compared characteristics such as asset size, regulator, and minority type—which may be related to the substance of answers to our survey questions—of nonrespondents to respondents. We did not detect a significant difference between those who chose to respond and those who did not based on these characteristics. To further assess the potential extent of nonresponse bias, we compared the response rates of the subgroups of those characteristics in our population, and determined that response rate did not differ markedly between categories of these

subgroups, suggesting that banks of certain types were not materially more likely to participate or not participate than others. Finally, we analyzed the patterns in response between those who answered in the earlier part of the fieldwork period and those who responded only after repeated follow-up attempts. It is possible that the latter group resembles nonrespondents. No significant difference in the answers between the groups was detected, which may suggest that actual nonrespondents would not have answered in a substantially different way from those who did. While the possibility exists that the true results for the entire population might be different from those we estimated in our report, we feel that on the basis of our analysis, nonresponse bias is unlikely.”

3. The limitations of what we can conclude from nonresponse bias analysis should be described. Some methods are more informative than others; be careful not to overextend the analyses. Also, a nonresponse bias analysis result of “no bias” should not be a statement that there definitely is no bias, but instead indicates that we did not detect any using these methods.

Appendix I: Anticipating Response Rates

Although response rates generally cannot be predicted, either general experience or early survey-specific information may provide useful estimates for planning purposes. Such information can be gained from pretesting and considering population characteristics. GAO’s or others’ experience with surveys of similar populations can be used for guidance and insight.

To identify characteristics that tend to lower the response rates, consider the items in table 1. This provides the team with a useful framework. It is not possible to weight the characteristics listed in the table, since their importance will vary across surveys. In addition, while the table may be a useful start for anticipating response rate problems, other factors may be equally or more important in particular surveys.

Table 1: Framework for Anticipating Low Response Rates: Characteristics that tend to lower the response rate

Responding individual or organization	
1	Not interested in questionnaire topics
2	Opposed to answering questionnaires generally
3	Fears contact with GAO or groups involved in the survey
4	Fears that answers might be made public
5	Fears that answers could lead to an unfavorable report conclusion for the respondent
6	Unlikely to be available during data collection
7	Is frequently surveyed or regularly reports government data
8	Was recently surveyed by GAO
9	Has especially low or especially high social status
Survey Design	
10	Survey imposes high respondent burden (e.g., it is time-consuming, complex, or contains many open-ended questions)
Survey Administration	
11	Survey request is addressed to a location or position, not a person
12	Request goes through an intermediary, not directly to the respondent
13	Contact information in sample frame is low in quality
14	Field period is short or at a time of year when respondents are more likely to be unavailable.