Relationship of Response Level, Survey Costs, and Bias: Survey of Science and Engineering Research Facilities and Survey of R&D Expenditures at Colleges and Universities

J. Michael Brick, Leslie Christovich, Kenneth Herrell, Benmei Liu, Timothy Smith
Westat, Inc.
National Science Foundation

Abstract
This paper focuses on the relationships between response rates and survey costs and response rates and nonresponse bias. The analysis used data from the NSF Survey of R&D Expenditures at Colleges and Universities and the NSF Survey of Science and Engineering Research Facilities.

Keywords: Response Rates, Survey Costs, Nonresponse Bias

Introduction
This research focuses on the relationships between response rates and survey costs and response rates and nonresponse bias. Nonresponse bias is a function of two components, the nonresponse rate and the difference between nonrespondents and respondents to the survey. Estimates from a survey with high nonresponse rates can still produce estimates with low nonresponse bias if the difference between respondents and nonrespondents is small. Conversely, low nonresponse rates in a survey can give estimates with high nonresponse biases if respondents and nonrespondents differ dramatically.

In household or human population surveys there has been considerable research into the effects of curtailing nonresponse follow-up efforts. The findings from household surveys (e.g., Curtin, Presser, and Singer 2000) have raised questions about long-held beliefs about the relationship between response rates and nonresponse bias. Traditionally, response rates have been used as indicators of survey quality, based on the assumption that decreases in response rates lead to increases in nonresponse bias. However, numerous population studies found little relation between variation in response rates and changes in nonresponse error.

In establishment surveys, research into the effects of curtailing nonresponse follow-up efforts on costs and data quality has been limited. There is some evidence that differences between respondents and nonrespondents at different stages of data collection may vary and induce nonresponse bias. However, the bias is specific to each variable and might be negligible depending on the relationship between the statistic of interest and the response pattern (e.g., Wright et al. 2001).

Background
The Science Resources Statistics Division (SRS) of the U.S. National Science Foundation has a legislative mandate to “provide a central clearinghouse for the collection, interpretation, and analysis of data on scientific and engineering resources, and to provide a source of information for policy formulation by other agencies of the Federal Government.” This research uses data from two of the periodic, establishment censuses conducted by SRS: the fiscal year (FY) 2003 and FY 2004 cycles of the Survey of Research and Development Expenditures at Colleges and Universities (Expenditures survey) and the FY 2003 and FY 2005 cycles of the Survey of Science and Engineering Research Facilities (Facilities survey).

The Expenditures survey is a census of U.S. colleges and universities granting a bachelor’s degree or higher in science and engineering (S&E) fields and with at least $150,000 in research and development (R&D) expenditures.1 The Facilities survey uses the Expenditures survey as a frame but limits the population to a census of all colleges and universities with at least $1 million in R&D expenditures.2

For the four survey cycles used in this analysis, data are used on the time of submission of the completed survey to examine what might have happened to the estimates of key statistics if the data collection had been terminated before the actual end of the data collection cycle.

---

1 The Expenditures survey also collects data from 37 Federally Funded R&D Centers; these are excluded from this paper’s analyses.
2 The Facilities survey also collects data from a census of biomedical research institutions receiving research funds from the U.S. National Institutes of Health; these are excluded from this paper’s analyses.
1. Response Rates

Since the goal of the research was to examine the relationships between response rates and costs and response rates and nonresponse bias, we created scenarios with different response rates: 75, 88, and 90 percent. We artificially created lower response rates than those actually reached by assuming that institutions responding after the date when a specific response rate (e.g., 88%) was achieved were nonrespondents. Respondent institutions were classified into three response groups based on the date of their submission.

For both surveys, we applied weighting procedures to the data sets that would have resulted if data collection had ceased when response rates reached the specified response level. Also, the imputed data for item nonresponse were treated as if they were actual responses.

2. Survey Estimates

We investigated potential nonresponse bias using the estimated absolute relative differences of key survey estimates. Three basic measures were computed for each estimate. The first was an estimated total, \( \hat{Y} \), the key statistic of interest in the survey. The estimated totals were computed for the full sample (\( \hat{Y}_{\text{full}} \)) and for each response level (\( \hat{Y}_l \) for \( l = 75, 88, \) and 90%).

The second measure was the estimated difference between the total for response level \( l \) and the full response level, written as
\[
\hat{D}_l = \hat{Y}_l - \hat{Y}_{\text{full}}
\]
If the full response level had included all the institutions in the frame, then this would actually be an estimate of the nonresponse bias for the estimated total using the lower level of response \( l \). Since that is not the case, it is the best estimate of bias under the circumstances.

The third measure is the estimated absolute relative difference of the estimate. It is the absolute value of the difference divided by the estimate from the full response level, written as
\[
\overline{ARD}_l = \left| \frac{\hat{D}_l}{\hat{Y}_{\text{full}}} \right| = \left| \frac{\hat{Y}_l - \hat{Y}_{\text{full}}}{\hat{Y}_{\text{full}}} \right|
\]

This measure is essentially a standardized difference and was used to compare totals across variables that had different measurement scales.

For the Facilities surveys, we examined three survey statistics: net assignable square feet of research space (NASF) by S&E field; completion costs of repair/renovation projects started in FY 2002 or FY 2003 by S&E field; and cost and NASF resulting from new construction projects started in FY 2002 or FY 2003 by S&E field. For the Expenditures surveys, we examined three statistics: S&E R&D expenditures by source of funds; S&E R&D expenditures by S&E field; and S&E R&D expenditures for purchase of equipment by S&E field.

3. Results

We examined nonresponse differences in key estimates using two approaches. First, for all of the many measures, we examined the \( p \) values for the t-tests of whether the difference between the estimated total for a response level and the full response level were statistically significant. Even though there were a huge number of estimates, the number of statistically significant differences was very small. In addition, because of the positive correlation between the estimates for the specific response level and the full response, the standard errors of the differences were substantially smaller than they would have been if the two estimates were independent. Thus, even relatively small differences had small standard errors and were more likely to be statistically significant. Despite the small standard errors, very few differences were statistically significantly different from zero.

We also graphed all of the estimated absolute relative differences computed for NASF, costs for repair/renovation, costs for new construction projects, and NASF for new construction projects for the institutions in the FY 2003 and FY 2005 Facilities survey. Each panel in Figure 1 corresponds to one of these FY 2003 key statistics, and the distributions are plotted for the 75, 88, and 90 percent response levels. Even though most differences are not statistically significant, the pattern suggests that the deviation from zero of absolute differences is greater for the 75 percent level than for the 88 percent or 90 percent level.

---

3The final academic response rates for the Facilities surveys were 92 and 94 percent; the final rate for the Expenditures surveys was 94 percent for both cycles.

4Although this report is limited to findings on these six estimates, numerous other estimates were also included in this analysis. Information on these other estimates is available from the authors.
Figure 1. Relative Absolute Differences by Response Levels for Estimates of Academic Institutions From the FY 2003 Facilities Survey

Figure 2 covers the FY 2005 Facilities survey. The distribution of absolute relative differences for these estimates does not show as much of the same obvious increase in the spread for the 75 percent response level as in Figure 1.

Figure 2. Relative Absolute Differences by Response Levels for Estimates of Academic Institutions From the FY 2005 Facilities Survey
Figure 3 graphs all the estimated absolute relative differences computed for R&D expenditures by source, R&D expenditures by field, and R&D expenditures for research equipment by field for the FY 2004 Expenditures survey. Across the three panels, there is again some indication of more spread of the absolute relative differences at the 75 percent response level, but no large and persistent pattern. The results were similar for the FY 2003 statistics.

**Figure 3. Relative Absolute Differences by Response Levels for Estimates From the FY 2004 Expenditures Survey**

Across all figures, the 75 percent response level generally exhibited larger absolute relative differences than the other response levels, although the patterns were not consistent. For some of the graphs of statistics we examined, the differences for the 75 percent response level were less than the other levels. No significant bias is exhibited at the 88 percent level or higher. One caveat is that with graphical analysis it is very hard to visually account for the correlation between the response-level estimates. As the response level approaches full, the differences must decrease since they are defined to be zero at the full response level. Consequently, we should expect response levels that differ (e.g., the 75% and full levels) to have a greater dispersion than those that are close together (e.g., the 90% and full levels). The statistical tests discussed earlier dealt with these correlations in ways that are difficult to do when viewing the estimates.

In summary, the analyses show that for the selected survey items, no significant relative bias actually would be caused by discontinuing the data collection once reaching an 88 percent response rate. Even stopping when 75 percent is reached, significant biases were rarely detected, although the graphs suggest the relative biases might be higher.

**4. Survey Costs**

The cost analysis was limited to the FY 2003 and FY 2005 cycles of the Facilities survey. Survey costs were defined as the labor costs associated with data collection and with data retrieval. Since a large number of institutions used the Web to respond to the survey, most of these costs are from staff trying to contact the institutions by telephone to encourage participation or to resolve inconsistent or missing data.

To calculate costs for each institution, we first determined the number of days during each month when data collection staff were in contact with the institution in an effort to obtain survey submission. This determination was based on the number of days in the month during which the institution was classified.
as a nonrespondent. Next, we calculated the total data collection costs for each month and the cost per active day in each month. The total data collection costs were a function of the labor hours billed in each labor category, the hourly rate for each labor category, and the number of days in the month. Total data collection costs for a month reflected the sum of all staff costs incurred during the month.

Data retrieval costs varied by institution depending upon the amount of inconsistent and missing data contained in their survey submission. Cost per retrieval item was calculated and applied to each institution to reflect the overall effort to resolve the case. Total cost per data retrieval element was a function of the labor hours billed in each labor category, the hourly rate for each labor category, and the number of data retrieval elements across all institutions.

Costs for each response level included both data collection and data retrieval costs and included costs for both respondent and nonrespondent institutions. The increases for both the total costs and the component costs by response level were found to be gradual and were rather smooth functions of the response rate. Neither data collection nor data retrieval costs dramatically increased in attempting to achieve higher response levels. No point was found at which either of the costs increased rapidly.

4.1 Survey Costs and Nonresponse Bias

Figure 4 plots the total costs and mean relative absolute differences of NASF, one of key estimates from the facilities surveys, for the three response levels. The mean relative absolute difference is the mean of all the relative absolute differences computed for the NASF total across all fields of study. Figure 4 shows that for both years, the slope for the cost curve is relatively steep while the nonresponse bias (as well as bias captured in the mean relative absolute difference) is nearly flat across the response levels. Similar results were found for the estimates of total costs for repair/renovation for the two survey cycles.

Figure 4. Ratio of Cost to Full Effect Cost and Mean Relative Absolute Differences of NASF by Response Levels for Estimates From Academic Institutions From the FY 2003 and FY 2005 Facilities Surveys

The survey cost analysis found that the costs of data collection and data retrieval grew smoothly from the lowest response level to the full response level for the Facilities survey. The nonresponse bias analysis found no strong evidence of nonresponse bias of the estimates related to the response level. The analyses generally supported the conclusion that the relative reduction in costs is greater than the increase in relative bias of the estimates once response rates of at least 75 percent have been obtained.

References
