

Prioritizing Low Propensity Sample Members in a Survey: Implications for Nonresponse Bias

Jeffrey A Rosen^{*}, Joe Murphy[†], Andy Peytchev[‡], Tommy Holder^{**}, Jill Dever^{††}, Debbie Herget^{‡‡}, Daniel Pratt^{***}

Tags: low propensity respondents, nonresponse bias, nonresponse

Survey Practice

Vol. 7, Issue 1, 2014

Many survey methodologists now agree that simply striving to increase the response rate is not an optimal approach for reducing nonresponse bias in the final survey estimates. Targeting sample cases that are underrepresented can help reduce nonresponse bias. However, the challenge lies in which cases to prioritize when resources are finite, and reducing the risk of nonresponse bias is the goal. We present an approach which identifies, prioritizes, and intervenes on low-propensity-to-respond cases during nonresponse follow-up. Targeted cases were assigned to in-person interviewing. Our results suggest that in-person interviewing can be an effective approach for gaining participation from low-propensity cases. We also find that targeting low-propensity cases could improve representation and therefore should be considered by survey practitioners as a tool for nonresponse bias reduction.

BACKGROUND AND SIGNIFICANCE

To decrease nonresponse bias during data collection, survey data need to be collected from those sample members that may contribute the most to nonresponse bias. This implies two steps (Peytchev et al. 2009): (1) identify the most important cases to focus effort and (2) devise an intervention to gain participation from the targeted sample members. We describe this process, how we operationalized each step, and present results.

TARGETING CASES WITH LOWEST RESPONSE PROPENSITIES

There continues to be interest and debate among researchers regarding strategies for prioritizing survey sample members during nonresponse follow-up (see, Peytchev et al. 2010; Rosen et al. 2010; Wagner 2012). Many agree that data collection effort should not be driven by a desire to simply increase response rates since higher response rates will not necessarily reduce nonresponse bias. Nonresponse follow-up interventions are successful in reducing nonresponse bias to the extent that they secure participation from

* **Institution:** RTI International

† **Institution:** RTI International

‡ **Institution:** RTI International

** **Institution:** RTI International

†† **Institution:** RTI International

‡‡ **Institution:** RTI International

*** **Institution:** RTI International

(underrepresented) nonrespondents who are unlike cases already interviewed (Schouten et al. 2009). If a nonresponse follow-up only increases a response rate by interviewing those who are likely to participate (relative to those who have already participated), bias may not be reduced since newly interviewed cases may be similar to earlier respondents. Prioritizing low response propensity cases may be logical given their possible relationship to nonresponse bias reduction, even if the response rate is not increased as much as when “easier” cases are targeted.

There may be risks involved with prioritizing low-propensity sample members. A long-standing concern has been that reluctant respondents would provide inferior data when interviewed (e.g., Cannell and Fowler 1963). Measurement error is also a concern. Peytchev et al. (2010) found a relationship between response propensity and measurement error for a particular estimate of a highly sensitive question. Olson (2006) demonstrated that measurement error can be statistic-specific, even though for total bias, low-propensity cases seem to be important. It seems that whether the participation of unlikely respondents provides greater measurement error is both statistic- and survey-specific. Our concern in this paper is not about measurement error in reports by reluctant respondents; we alert the reader to this potential undesirable outcome.

IN-PERSON FOLLOW-UP FOR LOWEST PROPENSITY CASES

In prioritizing low-propensity cases, a consideration is the method employed to target those cases. In this study, we utilized in-person computer-assisted personal interviewing (CAPI) for low-propensity targeted cases as an alternative to continuing the main study methods of self-administered web and computer-assisted telephone interviewing (CATI).

RESEARCH QUESTIONS

This study examines two important issues with low propensity to respond cases. We first sought to determine if the participation of low likelihood respondents can be an effective method to reduce nonresponse bias. Second, we sought to determine if CAPI could be an effective approach to encourage participation for low likelihood respondents.

DATA AND METHODS

We developed and implemented a response propensity-based responsive design in the High School Longitudinal Study of 2009 (HSLs:09), sponsored by the National Center For Education Statistics, U.S. Department of Education. HSLs:09 follows a cohort of ninth-grade students through their high school career, postsecondary education, military career, and/or the workforce. Data collection occurred in the fall of 2009 (base year) and the spring of 2012 (first follow-up). The HSLs:09 first follow-up included a parent survey.¹

Response propensities for the parent cases, the focus of this research, were

calculated 6 weeks into data collection – immediately following a 3-week early web response period and a subsequent 3-week CATI period in order to allow time for all cases to be worked in the less costly early phases of the data collection period. The dependent variable on which response propensities were based was a parent’s response outcome during these first 6 weeks of data collection. At the conclusion of the CATI period, the propensity model was fit, and all nonresponding cases in the lowest quartile of propensity scores were identified for treatment.

MODEL SPECIFICATION

We developed a response propensity model that incorporated paradata² and sampling frame variables to estimate a parent’s likelihood of response. As predictors, a range of paradata, student, parent, and school characteristics and panel maintenance results were considered. We investigated the use of survey (y) variables; however, none had all known values for both respondents and nonrespondents. Imputation for nonrespondents was considered, but ultimately rejected since the level of missing values was determined to be too high.

At the time of model implementation, 3,385 parents of the student sample members had responded, leaving 8,065 pending cases that were available for consideration for the CAPI intervention. All pending cases were assigned a propensity score. To assign propensity scores, a logistic regression model³ was fit to the data and predicted probabilities were used to score pending sample cases. Variables retained in the propensity model are listed in Table 1. Table 1 also lists the odds ratios and associated confidence intervals for all retained variables used in the propensity score calculation.

1 For more information on HSLS:09 and the study methodology, see: <http://nces.ed.gov/surveys/hsls09/>.

2 Paradata, in this context, refers to information related to locating, contacting and interviewing sample members (e.g., results of call attempts, status of mailings, and whether refused initially).

3 The model fit the data well; adjusted r squared for the final iteration was 0.47. The model was run unweighted and did not account for the complex survey design, although it included the stratification variables as predictors. We considered the following variables for inclusion in the propensity model: base year response status for the parent and student; outcomes (response status and mode) for the first follow-up panel maintenance activity; school enrollment status (dropout, transfer, homeschooled, in same school in base year and first follow-up); early graduate status; whether the student was a refusal or absent during the base year; call counts in the base year and first follow-up; contact attempts in base year and first follow-up; whether the case made a hard or soft appointment; whether or not the case logged into but did not complete the Web interview; gender; race; school type; metro type, and region.

Table 1 Modeling the likelihood of being a respondent – odds ratios and confidence intervals for retained variables.

Retained variable	Odds ratio	95% Confidence interval
Student was a respondent in the base year	1.81	1.53, 2.14
Parent was a respondent in the base year	2.14	1.93, 2.38
Responded to the first follow-up panel maintenance activity	2.93	2.52, 3.41
Was a web respondent to the first follow-up panel maintenance activity	1.73	1.41, 2.13
First follow-up dropout	1.49	1.10, 2.00
In the same school in first follow-up and base year	1.78	1.59, 1.99
Parent refused during first follow-up	0.13	0.08, 0.19
Call count during first follow-up	0.96	0.96, 0.96
Lives in a rural area	1.27	1.14, 1.41
Lives in the Midwest	1.54	1.37, 1.73
Lives in the South	1.33	1.20, 1.47
Logged into the web interview (first follow-up) but did not complete	0.19	0.11, 0.35
Number of first follow-up contact attempts	1.37	1.34, 1.41
Made a hard appointment	1.86	1.68, 2.07
Made a soft appointment	1.71	1.54, 1.89

IMPLEMENTATION OF IN-PERSON INTERVIEWING

The lowest propensity cases were sent to the field for a CAPI attempt. These selected cases were assigned to designated field interviewers who were responsible for working the case until disposition. A total of 2,051 (determined by project resources) out of the 8,065 pending cases were targeted which corresponded roughly to the lowest quartile of response propensities. The remaining 6,014 nonresponding cases continued in the CATI and Web modes. A small number of cases selected for CAPI were not able to have in-person visits due to excessive costs associated with those cases. However, in these few instances, the field interviewers made contact attempts by telephone.

RESULTS

EFFECTS OF CAPI ON LOW-PROPENSITY CASES

To assess the effectiveness of the approach on participation rates, we measured the relationship between response propensity and actual response to the survey for cases above and below a response propensity cutoff determined by available resources. Those under the cutoff received CAPI follow-up. Figure 1 shows the achieved completion rates for cases by response propensity percentile. The lines represent linear regression results using a regression discontinuity model (Bloom 2009) to show effects separately by propensity class. The boost in response among the low-propensity group shows that the CAPI treatment was likely effective in increasing the rate of completion among the lowest propensity cases to at least the levels among the mid-range propensity cases.

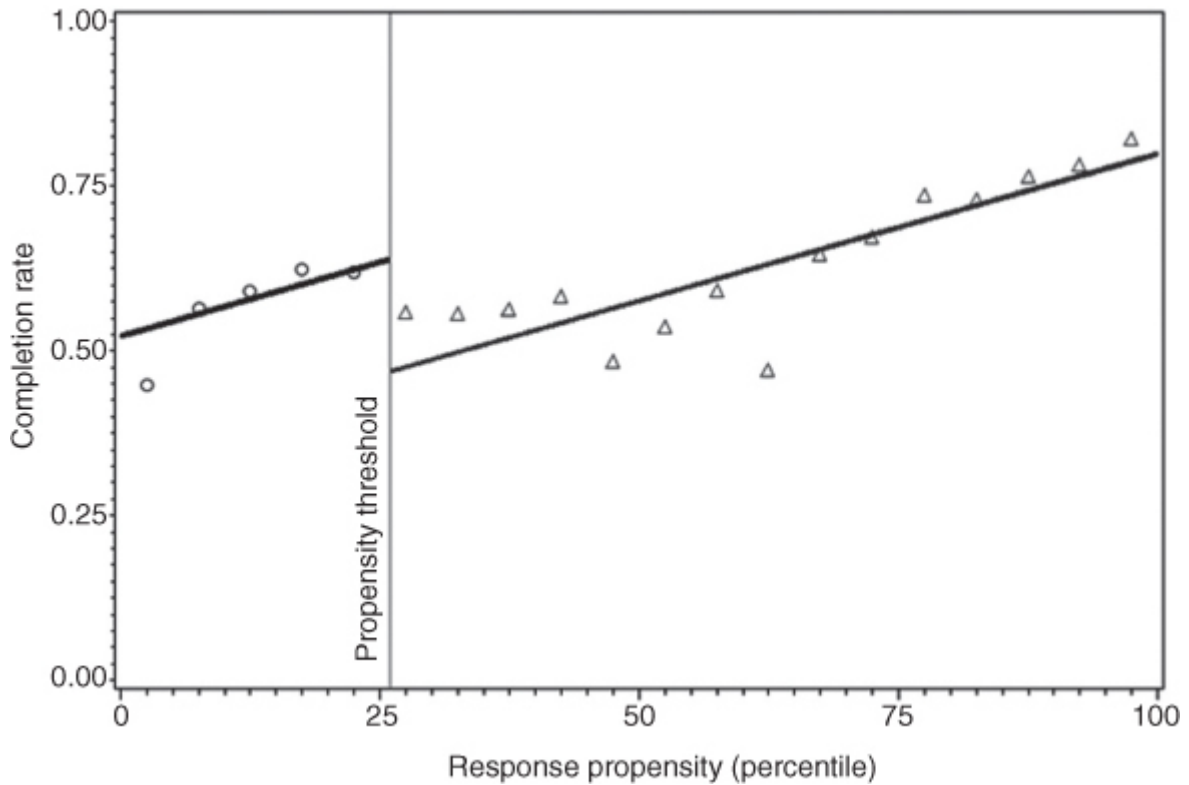


Figure 1 The effects of CAPI treatment (low – propensity percentiles) on completion rates.

POTENTIAL BIAS IN LOW-PROPENSITY CASES

To assess the effectiveness of the approach on unit level bias, a nonresponse bias analysis was conducted using frame variables both before and after the inclusion of the low-propensity cases to determine if the inclusion of low-propensity cases reduced nonresponse bias. A successful reduction in bias would be identified if frame variables showed a statistically significant bias before, but not after, the inclusion of the low-propensity cases. As shown in Table 2, this pattern occurred for several frame variables. For example, consider the race variable. Prior to the inclusion of low-propensity cases, the Black (estimated bias=-1.88) and Other (estimated bias=2.73) race categories showed significant bias. After the inclusion of low-propensity cases, bias levels for both Black (estimated bias=-0.51) and Other (estimated bias=1.70) was reduced and no longer significant. The mean relative bias for the variables given in Table 2 showed a reduction from 7.91 (prior to the inclusion of low-propensity cases) to 3.98 (after the inclusion of low-propensity cases). The mean relative bias across all the variables for which bias analyses was completed (variables in Table 2 as well as 9th grade enrollment, census regions, grade range of school, urbanicity, and gender) also showed a small reduction from 7.64 to 7.15. This overall pattern suggests that the inclusion of low-propensity cases for these variables reduced bias. It should also be noted that none of the variables listed above were included as a predictor in the propensity model. Notably, no variables showed a pattern of significant bias occurring after the inclusion of low-propensity cases, although this may have been the result if the

CAPI-assigned cases were continued under the Web and CATI protocol.

Table 2 Weighted bias analysis before and after the inclusion of low-propensity Cases.^{1,2}

Domain category	Bias analysis before identifying and targeting low propensity cases (n=11,450)			Bias analysis including low propensity cases (n=8,310)		
	Nonresponse rate	Overall weighted mean	Bias ^a	Nonresponse rate (final)	Overall weighted mean	Bias ^a
School type						
Public	71.2	92.9	-1.15*	27.9	92.5	-0.37
Private	56.6	7.1	1.15*	25.6	7.5	0.37
Charter school						
Yes	83.2	1.4	-0.01	32.7	1.6	0.03
No	71.0	91.2	-1.14*	27.8	90.9	-0.40
Private	56.6	7.3	1.15*	25.6	7.5	0.37
Religious affiliation						
Yes	56.6	6.8	1.14*	25.1	7.2	0.39
No	56.7	0.3	0.00	36.2	0.3	-0.02
Public	71.2	92.9	-1.15*	27.9	92.5	-0.37
School is regular secondary						
Yes	56.4	6.5	1.01*	25.5	6.8	0.31
No	59.2	0.7	0.14	26.7	0.7	0.07
Public	71.2	92.9	-1.15*	27.9	92.5	-0.37
Race						
Hispanic	69.8	4.0	-0.25	31.2	3.7	-0.29
Asian	86.0	16.7	-0.60	28.8	15.8	-0.90
Black	80.8	14.6	-1.88*	31.6	14.1	-0.51
Other	63.8	64.7	2.73*	26.4	66.4	1.70

*Result was significant at $\alpha=0.05$.

¹Table includes only variables which showed estimated bias becoming insignificant after the inclusion of low-propensity cases. Some frame variables did not show this pattern. They include, 9th grade enrollment, census regions, grade range of school, urbanicity, and gender.

²We assumed that these frame variables were related to some key survey variables of interest. Lowering bias in the frame variables, we think, will lower bias in the survey variables we cannot include in bias analyses due to missing information for nonrespondents. Table includes only variables that showed estimated bias becoming insignificant after the inclusion of low-propensity cases. Some frame variables did not show this pattern. They include, 9th grade enrollment, census regions,

grade range of school, urbanicity, and gender. The nonresponse bias analysis used study weights and accounted for the complex survey design.

^aThis is the estimated bias which is defined as the weighted item nonresponse rate among all eligible sample members multiplied by the difference between the survey variable mean for respondents and nonrespondents.

DISCUSSION

Our results suggest a few important things for survey practice. First, if low-propensity cases can be accurately identified, they may represent a sensible set of cases to target (in terms of potential bias reduction) during nonresponse follow-ups. Low-propensity nonresponding cases, in our study, appeared to be different in terms of representation from those who responded in the earlier phases of data collection. And, the fact that we were successful in securing their participation to a greater degree than we would have been had we not intervened with CAPI (as evident in Figure 1), suggests that CAPI efforts yielded benefits in terms of increased participation from challenging cases. The primary implication for survey data quality is that our intervention seems to have improved survey representation as evidenced by less bias in the frame variables. Second, CAPI could be an effective intervention for difficult cases in studies using other survey modes. In our study, CAPI produced much greater participation among the most challenging sample members. Of course, the cost of CAPI is significant so the challenge for survey teams is to assess whether the cost is acceptable. When CAPI can only be applied to a portion of the sample, it might be worthwhile to focus CAPI efforts on the types of cases likely to introduce bias if not interviewed.

A limitation to note about the evaluation of the reduction of nonresponse bias is the lack of a control condition. Since response rates were apparently increased among the lowest propensity cases (based on the discontinuity analysis) and biases were reduced, we conclude that the intervention was effective; however, the magnitude of the bias without the intervention would likely be smaller if the Web and CATI modes were continued, as they were continued for the rest of the nonrespondents. Future studies might consider including a control condition if feasible. That would better allow the evaluation of the impact on other statistics of interest, such as the variances of weighted estimates. Future studies might also consider identifying a propensity cut point at which personal interviews are more effective than CATI in terms of cost and survey response. Since personal interviews are costly, it may be worthwhile to identify precisely the point at which for example, a low propensity case should be removed from CATI and put into another mode. Unfortunately, our data did not allow us to examine this issue.

Caution is warranted for survey managers since low-propensity cases may introduce other sources of survey errors that need to be evaluated, such as levels of item nonresponse and measurement error. Other selection criteria should also be considered particularly with respect to bias reduction and cost tradeoff, such as avoiding the most extreme propensities, or using the more traditional random selection of nonrespondents in two-phase designs.

Despite these issues, our results suggest that the targeting of low-propensity cases for nonresponse bias reduction should be considered by survey practitioners.

REFERENCES

- Bloom, H. 2009. *Modern regression discontinuity analysis*. MDRC, New York.
- Cannell, C.F. and F.J. Fowler. 1963. Comparison of a self-enumerative procedure and a personal interview: a validity study. *Public Opinion Quarterly* 27(2): 250–264.
- Olson, K. 2006. Survey participation, nonresponse bias, measurement error bias, and total bias. *Public Opinion Quarterly* 70(5): 737–758.
- Peytchev, A., R.K. Baxter and L.R. Carley-Baxter. 2009. Not all survey effort is equal: reduction of nonresponse bias and nonresponse error. *Public Opinion Quarterly* 73(4): 785–806.
- Peytchev, A., E. Peytcheva and R.M. Groves. 2010. Measurement error, unit nonresponse, and self-reports of abortion experiences. *Public Opinion Quarterly* 74(2): 319–327.
- Peytchev, A., S. Riley, J.A. Rosen, J.J. Murphy and M. Lindblad. 2010. Reduction of nonresponse bias in surveys through case prioritization. *Survey Research Methods* 4(1): 21–29.
- Rosen, J.A., J.J. Murphy, A. Peytchev, S. Riley and M. Lindblad. 2010. The effects of differential interviewer incentives on a field data collection effort. *Field Methods* 23(1): 24–36.
- Schouten, B., F. Cobben and J. Bethlehem. 2009. Indicators for the representativeness of survey response. *Survey Methodology* 35(1): 101–113.
- Wagner, J. 2012. A comparison of alternative indicators for the risk of nonresponse bias. *Public Opinion Quarterly* 76(3): 555–575.