ARTICLES

Issues Facing the Field: Alternative Practical Measures of Representativeness of Survey Respondent Pools

Robert M Groves^{*}, J Michael Brick[†], Mick Couper[‡], William Kalsbeek^{**}, Brian Harris-Kojetin^{††}, Frauke Kreuter^{‡‡}, Beth-Elle Pennell^{***}, Trivellore Raghunathan^{†††}, Barry Schouten^{‡‡‡}, Tom Smith^{****}, Roger Tourangeau^{††††}, Ashley Bowers^{‡‡‡‡}, Matthew Jans^{*****}, Courtney Kennedy^{†††††}, Rachel Levenstein^{‡‡‡‡‡}, Kristen Olson^{*****}, Emelia Peytcheva^{††††††}, Sonja Ziniel^{‡‡‡‡‡‡}, James Wagner^{*****}

Keywords: survey practice

<https://doi.org/10.29115/SP-2008-0013>

Survey Practice

Vol. 1, Issue 3, 2008

Issues Facing the Field: Alternative Practical Measures of Representativeness of Survey Respondent Pools

It is increasingly clear (e.g., Curtin et al. 2000; Groves 2006; Keeter et al. 2000) that nonresponse rates are poor indicators of nonresponse error of survey estimates. However, the field (and the public at large) has been taught to use response rates as quality indicators. Indeed, with the exception of the sample size, the response rate is probably commonly held to be the most important criterion of survey quality. As of yet, however, there is no alternative indicator of relevance to nonresponse error that has been proposed.

We thus seem to be in a moment of uncertainty – where old procedures have been questioned, but no new options have been forwarded. The value of the response rate historically was

1. its simplicity (a single number to characterize quality), and

* University of Michigan and Joint Program in Survey Methodology

† Westat

‡ University of Michigan and Joint Program in Survey Methodology

** University of North Carolina

†† US Office of Management and Budget

‡‡ Joint Program in Survey Methodology

*** University of Michigan

††† University of Michigan and Joint Program in Survey Methodology

‡‡‡ Statistics Netherlands

**** National Opinion Research Center

†††† University of Michigan and Joint Program in Survey Methodology

‡‡‡‡ University of Michigan

**** University of Michigan

††††† University of Michigan

‡‡‡‡‡ University of Michigan

***** University of Michigan

†††††† University of Michigan

‡‡‡‡‡‡ University of Michigan

***** University of Michigan

2. its transparency (it was not a complicated statistical function of sample observations).

The common perspective that higher response rates produce lower nonresponse error led to the popular use of response rates.

The plague of nonresponse research is the fact that very little information is typically available on nonrespondents. Based on recent meta-analyses it's clear that there is commonly more variation in nonresponse error across estimates of the same survey than across surveys with different response rates. For example, since

1. the key to nonresponse error is the correlation of the survey variables with response propensities,
2. these correlations vary as a function of covariances with the causes of propensities and the variables, therefore,

no single measure of nonresponse (at the survey level) can inform the researcher to the diverse levels of nonresponse error that various estimates might hold. This is not dissimilar to the problem the field faces with characterizing the effect of clustering on standard errors. Design effects vary across different estimates within a single survey. While no one design effect is appropriate for all estimates, sometimes average design effects are useful.

Alternative Approaches. To get the conversation moving, it seems that there are two “ideal-type” approaches of alternative indicators:

1. a single indicator at the survey level
2. individual indicators at the estimate level

Indicators at the survey level differ from the simple response rate computations by incorporating into the computation some key auxiliary variables that are correlates of the survey variables (or at least desirable measures on which the respondent pool should be balanced). Indicators at the estimate level are specific to a single estimate, providing some information based on variation in success of measuring groups that vary on the expected value of the estimate.

Regardless of its merits, the notion of having a separate indicator of nonresponse error for each estimate may be too large a leap in complexity for current practitioners. It seems useful, therefore, to evaluate compromises between a single indicator and estimate-specific indicators. These would be indicators of nonresponse errors for sets of estimates (ones that might have similar correlations with different auxiliary variables).

A SINGLE INDICATOR AT THE SURVEY LEVEL

A single indicator to characterize nonresponse error must perforce assert that a set of auxiliary variables measured on nonrespondents and respondents usefully capture the major correlates of *all* the survey estimates. We know this to be strictly false, but it is the smallest useful step away from response rates.

There are several that could be forwarded:

- a) Variance functions of nonresponse weights (e.g., coefficients of variation of nonresponse weights)

The value of this would be to indicate how variable the level of nonresponse is across sample cases. It could incorporate observations on respondents and nonrespondents as part of the data collection effort. Comment: these measure variation in response propensities without linkage to the survey variables and are dependent on how rich the poststratification weights are.

- b) Variance functions of post-stratification weights (e.g., coefficients of variation of poststratification weights)

This would probably be most appealing when there were no other nonresponse weights based on sample quantities and population totals served as the basis of both nonresponse and coverage adjustments. These would assert that the weights were correlated with the variables of interest and that the measures of these variables in the survey were equivalent to the measures used from the population data (e.g., that education measured in a survey was equivalent to the measure of education in the decennial census). Comment: these measure variation in response propensities without linkage to the survey variables and are dependent on how rich the poststratification weights are.

- c) Variance functions of response rates on subgroups defined for all sample cases (both respondents and nonrespondents)

These could be coefficients of variation of response rates across key subgroups in the sample. These might result from efforts of the survey to induce interviewer observations or other design features to capture such variables, ideally correlated with the key survey variables. Comment: These indicators are generally sensitive to the size of the samples in the subgroups and the level of response rates.

- d) Goodness of fit statistics on propensity models

It is more and more common to build response propensity models (often logistic regression models predicting the likelihood of a complete). When these models fit the data very well, then de facto predictor variables have identified groups with very different response rates. This indicator would assert that if such a case existed, there is evidence of nonresponse bias. Comment: the value of this is probably limited to estimates using survey variables highly correlated with the predictor variables of the response propensity model.

e) R-indexes, which are model-based equivalents of the above

Shouten defines a whole set of indexes, which he calls R-indexes. Among these is a marginal R-index using a standardized regression coefficient in a logistic propensity model, to indicate the degree of imbalance among respondents and nonrespondents on an auxiliary variable. There is a large class of such indexes, some overlapping with the above. The distinctive subset would posit multivariate propensity models with auxiliary variable predictors. Comment: a current challenge in these indexes is their tendency to be affected by the level of response rates as well as the variation in rates. (<http://www.cbs.nl/en-GB/menu/methoden/research/discussionpapers/archief/2007>).

INDICATORS AT THE LEVEL OF INDIVIDUAL ESTIMATES

Clearly the field would be on the soundest theoretical footing if each survey estimate produced had its own nonresponse error indicator. However, such a stand would be a giant leap in complexity for users of survey estimates. Moving to this level permits us, however, to examine various functions of the relationship between response propensities and individual survey variables. We use “ y ” as the designator of an individual survey item. We use the term, “auxiliary variables,” to describe any attributes that are known for respondents and nonrespondents

a) Comparisons of respondents and nonrespondents on auxiliary variables

Any variables available on the sampling frame or used for postsurvey adjustment are measured on both respondents and nonrespondents. These means on these variables for respondents can be compared to those of nonrespondents (as well as the total sample). In essence, these comparisons permit nonresponse bias estimates on such variables.

Comment: to the extent that these variables are correlated to survey variables these indicators may be predictive of nonresponse error on the survey variables. A useful exercise might be to group survey variables by the level of correlation with the auxiliary variables, as measured among respondents, to comment on which auxiliary variables are most like various survey variables.

b) Correlation between post-survey nonresponse adjustment weights and y , measured on the respondent cases

All surveys that use some adjustment procedure in hopes of reducing nonresponse error of estimates possess a variable on every respondent that acts to increase the influence of underrepresented cases and decrease the influence of overrepresented cases within the respondent pool. Correlations between the resulting “postsurvey” adjustment weights and the survey variables (estimated on the respondent cases) may be informative about the relationship between response propensity and those survey variables. Comment: the correlations

may not be the same within the nonrespondent pool and thus may mislead the researcher about the extent of nonresponse bias in the adjusted and unadjusted estimates.

c) Examine the means of survey variable y within deciles of the survey weights

With a graphical display of mean across deciles of nonresponse weights, this is a visual equivalent of the correlation in b) above. Comment: the variation in means across propensity groups may not be the same within the nonrespondent pool as within the respondent pool and thus may mislead the researcher about the extent of nonresponse bias in the adjusted and unadjusted estimates.

d) Fraction of missing information on y

This is based on the ratio of between-imputation variance of an estimate and the total variance of an estimate, based on imputing values for all the nonrespondent cases in a sample (Little and Rubin 2002). The imputation models can be diverse, but one of some appeal is that of a sequential regression imputation, utilizing all of the variables in the data set. The percentage of missing information would be high when the percentage of variation in the total estimate due to imputation was large. It would be small when the percentage of variation in the total estimate due to imputation was small. Comment: The value of this indicator is a function of the quality of imputation model, both its variance properties and its bias properties.

INDICATORS AT THE LEVEL OF SETS OF VARIABLES/ESTIMATES

Instead of each estimate in a survey having its own indicator of nonresponse error, for convenience, sets of similar estimates could be defined, such that the set would share a value of a nonresponse indicator. These would be based on a prior analysis, most likely among the respondent cases only, by which the researcher would establish the magnitude of relationships (covariances or other measure of the relationship) between the auxiliary variables and the likelihood of participation. The researcher would identify sets of survey variables or estimates that share high correlations with different auxiliary variables *on the respondent cases*. Based on that indirect information, separate variance functions above would be presented for each class of estimates. In some sense, this would resemble the technique of identifying sets of estimates subject to similar design effects (e.g., see the CPS Technical Paper 63RV, <http://www.census.gov/prod/2002pubs/tp63rv.pdf>).

SUMMARY

The above is a listing of alternative indicators that might be useful to explore across several ongoing surveys simultaneously, in order to aid judgments about whether different indicators do indeed supply more information about possible nonresponse errors in survey estimates.

REFERENCES

- Curtin, R., S. Presser, and E. Singer. 2000. "The Effects of Response Rate Changes on the Index of Consumer Sentiment." *Public Opinion Quarterly* 64 (4): 413–28.
- Groves, R.M. 2006. "Nonresponse Rates and Nonresponse Bias in Household Surveys." *Public Opinion Quarterly* 70 (5): 646–75.
- Keeter, S., C. Miller, A. Kohut, R.M. Groves, and S. Presser. 2000. "Consequences of Reducing Nonresponse in a National Telephone Survey." *Public Opinion Quarterly* 64: 125–48.