Statistics Netherlands
Division of Methodology and Quality
Department of Methodology The Hague

IMPROVING RESPONSE IN ADAPTIVE SURVEY DESIGNS GIVEN CONSTRAINTS ON RISK OF UNDESIRABLE MEASUREMENT PROFILES

Melania Calinescu, Barry Schouten and Sandjai Bhulai

Abstract: Recent survey literature shows an increasing focus on survey designs that balance survey quality and survey costs. Adaptive and responsive survey designs study the impact of various design features on response propensities and costs, given sample unit characteristics available in registry data or paradata. The mix of design features that yields the highest response quality, without overrunning the budget, represents the optimal design.

An important element in the optimization mechanism behind adaptive survey designs is the set of quality indicators that is used in the trade off with costs. The response rate is considered to be insufficient as a survey quality measure. The alternative is to aim at high and representative response, i.e., high response rates that show little variation over relevant population subgroups. However, in surveys that employ multiple survey modes, a focus on nonresponse is too simple. Mode-specific measurement error must be quantified and constrained.

In this paper we investigate how adaptive survey designs can help reduce nonresponse bias by implementing optimal designs. We present results from adaptive survey designs that address nonresponse error and measurement error in the Dutch Labor Force Survey.

Keywords: Nonresponse bias; Response bias; Adaptive Survey Design.

1 Introduction

Recent survey literature has put forward responsive and adaptive survey designs (ASDs) as means to make efficient tradeoffs between survey quality and survey costs. ASDs have however additional appeal. They take into consideration the fact that design features have a different impact on different persons and households (see [3] and [6]). The survey sample is divided into subpopulations based on available characteristics (e.g., age, gender, ethnicity) and paradata. The impact of survey design features on the propensity to respond is estimated from historical data given the subpopulations. Moreover, ASDs provide a mathematical framework to address the cost-effectiveness problem that traditional designs often encounter. The mix of design features that yield most quality (over all subpopulations) and it is not more expensive than the available budget gives the ASD optimal strategy. ASDs that employ paradata observations for the subpopulation definition are called dynamic and can be used in ongoing survey data collection. Such designs resemble the responsive designs (see [2]).

Currently, the quality indicators that ASDs use are the response rate and representativity of the response, thus a complete focus on nonresponse error. However, moving towards multiple survey mode designs, other errors become influential, such as measurement error and coverage error. The challenge ASDs face now is quantifying the impact and interaction of multiple survey errors. In the current paper, we investigate techniques to include measurement errors (MEs) in the ASD framework. To this end we extend the approach in [1] and [5] and use the Dutch Labor Force Survey (LFS) data from 2008 as a case study.

The remainder of the paper is structured as follows. Section 2 gives details on how MEs occur in the LFS and discusses the implications of accommodating MEs within the ASD framework. Section 3 provides numerical examples of the extended ASD model and Section 4 sketches directions for future research.

2 Adaptive survey designs for nonresponse and measurement

In [1] the authors develop the ASD mathematical framework and consider the survey mode and the contact protocol (timing and number of contact attempts) as the design features of interest. Other survey features such as different types of reporting, different advance letters, different interview skills can be treated in the same manner. The optimal strategy is found as a solution to the optimization problem where the expected response rate is maximized subject to a budgetary constraint and a constraint on the total number of contact attempts. The decision variables are the allocation probability of a given subpopulation to a survey mode per time slot. For computational reasons, the decision variables are binary.

In the current paper, we extend this setting by adding two constraints. First, a variability constraint that supplements the response rate as a quality indicator. This constraint, added via the R-indicator (representativity indicator, see for details [4]), limits the variance of subpopulation response rate, creating a more representative respondent sample, with respect to the specified characteristics. The second constraint deals with the measurement error. In constructing this constraint, we followed the approach in [5] that investigates the relationship between nonresponse error and measurement error as a function of several survey design features, in particular the reporting type. The measurement error is decomposed into a measurement profile risk and response errors. Response errors represent the actual differences between the observed and true values, while the measurement profile risk is the probability that a respondent will show a certain response style that may lead to response errors. For the sake of simplicity, we consider only the measurement profile risk and we assume that there is only one measurement profile. The model can be extended to include several measurement profiles. The error constraint limits the probability for a measurement profile (i.e., the proportion of respondents showing the undesirable measurement profiles) to a given threshold. We additionally investigate the possibility of modifying the objective function to maximize the expected rate of error-free response (i.e., if the measurement profile is observed, the response thus obtained is treated as nonresponse).

3 A case study: the Labor Force Survey

The Dutch LFS targets people with age between 15 and 64 years. Proxy reporting is allowed by members of the same household in order to increase the response rate and decrease traveling costs. Households that refuse participation are not re-approached for refusal conversion. Traditionally, the survey mode in LFS is CAPI. In the current paper we focus on differences that appear given the reporting type (self-report vs. proxy-report). We also explore the impact of the number of visits. For the sake of simplicity, we make a number of assumptions. First, the choice of design features does not impact the traveling distances and times. Second, the interviewer workloads are not influenced by the choice of designs features, given that the interviewers handle multiple surveys at the same time.

We consider a sample of size $N = 10000$ and given the age of the sample unit we differentiate three subpopulations, namely $G = \{15-25, 26-55, 56-65\}$. We assume the data collection period is divided in $T = 10$ time slots, which allows us to analyze the impact of the contact protocol up to ten visits. The weights of each subpopulation in the total population are $W = \{0.196, 0.624, 0.18\}$. Let $M = \{0, 1, 2\}$ be the reporting type, where 0 denotes no visit, 1 self-report and 2 proxy-report. Successful contact is defined as contact with the sampled person (in case of self-report) and contact with the household the sampled person belongs to (in case of proxy-report). Table 1 gives the estimates for contact, cooperation and measurement profile probabilities from LFS 2008. For illustration purposes, these probabilities are averaged over all the time slots.

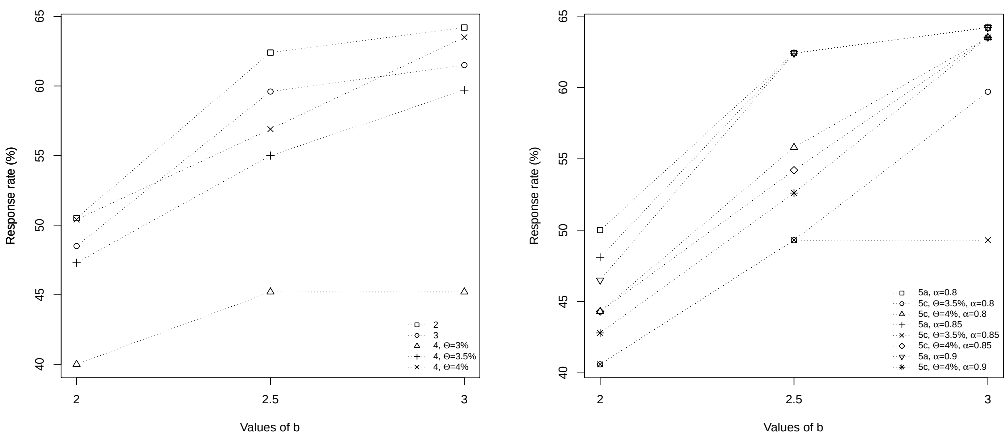
Probabilities	Reporting type	Group		
		$15-25$	$26-55$	$56-65$
Contact	$m = 1$	0.261	0.303	0.434
	$m = 2$	0.392	0.367	0.461
Cooperation		0.594	0.651	0.690
Measurement profile	$m = 1$	0.060	0.028	0.041
	$m = 2$	0.078	0.035	0.047

Table 1: Estimated input parameters

We investigate several settings, starting from a baseline setting (i.e., the reporting type is fixed as well as the average number of visits per sample unit) and adding one by one the aforementioned constraints. This gives rise to the following six settings:

- Setting 1 (baseline) - The average number of visits is fixed to b for each sample unit and only one reporting type is allowed (setting 1a - self-report, setting 1b - proxy-report)
- Setting 2 - Optimization: maximize expected response rate subject to a maximum number of visits $B = bN$. The reporting type is no longer fixed.
- Setting 3 - Optimization: maximize the expected error-free response rate subject to the same constraints as in Setting 2.
- Setting 4 - Add the error constraint to Setting 2.
- Setting 5 - Add the variability constraint to Settings 2-4.
- Setting 6 - Allow at most one change in the reporting type. Add this constraint to Settings 2-5.

For each of these settings, we consider the following parameter values: average number of visits per sample unit $b = \{2, 2.5, 3\}$; maximum proportion of respondents showing the measurement profile $\Theta = \{3\%, 3.5\%, 4\%\}$ and minimum level of representativity $\alpha = \{0.8, 0.85, 0.9\}$. Figure 1 gives a brief overview of the resulting response rate in Settings 2-5.



(a) Impact of error constraint

(b) Impact of variability constraint

Figure 1: Impact of error and variability constraints on the expected response rate.

The maximum expected response rate is 64.2% obtained in Settings 1b and 2 for $b = 3$, with a corresponding R-indicator of 0.939 and proportion of 4.51% of the respondents showing the measurement profile. All age groups receive visits in all the 10 time slots, which leads us to the conclusion that 64.2% is the maximum possible. This level decreases to 45.2% in Setting 4 when $\Theta = 3\%$ and we reach infeasibility when the variability constraint is added. For milder thresholds in the error constraint (e.g., $\Theta = 3.5\%$) the feasible regions still contains points and the optimal solution yields 49.3% response rate for $\alpha = 0.85$. In the case of self-report (Setting 1a), the maximum number of visits for $b = 3$ is not enough to visit all subpopulations in all time slots, yielding 60.9% response rate, with a corresponding R-indicator of 0.851 and proportion of erroneous response of 3.54%. In Setting 3, the level of response increases to 61.5%, with 3.95% proportion of erroneous response and 0.928 level of representativity.

4 Open Issues in Current Research

Survey research has tried to improve designs such that the quality of the estimates is high. To this end, multiple issues have to be taken into account such as presence of nonresponse, measurement errors, lack of representativity in responses and budget overruns. The current paper was motivated by the necessity of addressing measurement errors that arise when moving towards mixed-mode surveys. Given the ASD framework, three major research questions emerge when including analysis of measurement errors:

- How to design an optimization model that accounts for nonresponse as well as measurement errors? How to balance on two dimensions of survey error?
- How to select appropriate measures to investigate the impact of measurement errors?

- Would we allow for low response rates or low representativity if we can obtain low measurement errors?

Section 3 attempted to answer these questions. Further discussion is necessary to better understand the implications of combining measurement and nonresponse errors in adaptive survey designs. Given the results of the case study we have the following questions:

- The accuracy of the input parameters are crucial for a meaningful analysis. How can we obtain robust solutions?
- Is the measurement profile a good method to account for measurement errors and how would it relate to nonresponse?
- What paradata should we collect?

References

[1] M. Calinescu, S. Bhulai, and B. Schouten. Optimal scheduling of contact attempts in mixed-mode surveys. Discussion paper 201131, CBS, The Hague, 2011.

[2] S. G. Heeringa and R. M. Groves. Responsive design for household surveys: tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 169(3), 2006.

[3] B. Schouten, M. Calinescu, and A. Luiten. Optimizing quality of response through adaptive survey designs. Discussion paper 201118, CBS, The Hague, 2011.

[4] B. Schouten, F. Cobben, and J. Bethlehem. Indicators for the Representativeness of Survey Response. *Survey Methodology*, 35(1), 2009.

[5] B. Schouten and K. Leufkens. Response representativity and measurement profile risks in the lfs. Discussion paper 201104, CBS, The Hague, 2011.

[6] J. R. Wagner. *Adaptive Survey Design to Reduce Nonresponse Bias*. PhD thesis, University of Michigan, 2008.