Stockholms universitet

Research Report

Department of Statistics

No. 2018:2

Design-based "Optimal" Calibration Weights Under Unit Nonresponse in Survey Sampling

Per Gösta Andersson

Department of Statistics, Stockholm University, SE-106 91 Stockholm, Sweden

Design-based "Optimal" Calibration Weights Under Unit Nonresponse in Survey Sampling

Per Gösta Andersson, Department of Statistics, Stockholm University

Abstract

High nonresponse is a very common problem in sample surveys today. In statistical terms we are worried about increased bias and variance of estimators for population quantities such as totals or means. Different methods have been suggested in order to compensate for this phenomenon. We can roughly divide them into imputation and calibration and it is the latter approach we will focus on here. A wide spectrum of possibilities is included in the class of calibration estimators. We explore linear calibration, where we suggest using a nonresponse version of the design-based optimal regression estimator. Comparisons are made between this estimator and a GREG type estimator. Distance measures play a very important part in the construction of calibration estimators. We show that an estimator of the average response propensity (probability) can be included in the "optimal" distance measure under nonresponse, which will help reducing the bias of the resulting estimator. To illustrate empirically the theoretically derived results for the suggested estimators, a simulation study has been carried out. The population is called KYBOK and consists of clerical municipalities in Sweden, where the variables include financial as well as size measurements. The results are encouraging for the "optimal" estimator in combination with the estimated average response propensity, where the bias was highly reduced for the Poisson sampling cases in the study.

KEY WORDS: Distance measure; Nonresponse bias; Response propensity.

1 Introduction

In a survey the response (nonresponse) mechanism for units is in reality unknown. To avoid defining a proper probability measure which might not be meaningful or realistic, one usually discusses the nonresponse situation in terms of a propensity for a unit to participate. To be able to take into account the possible nonresponse effect on estimators, it is however practice to treat the propensities as probabilities to be estimated (e.g. propensity scores). This can be done for individual units, for groups of units or as an "average" over the whole response set.
In e.g. Haziza and Lesage (2016) two main approaches are discussed: calibration weighting with and without foregoing propensity score weighting. The authors warn us about potential negative effects on the bias and variance for the resulting estimators when not taking into account the propensities. (These two options of weighting are referred to by the authors as two-step and one-step procedures, respectively not to be mistaken for the two- and single-step calibrations as defined by Särndal and Lundström (2005).) However, in the simulation study by Haziza and Lesage (2016) the sampling design plays no role, since there $n = N$ and the focus is solely on how the auxiliary information relates to the study variable and the nonresponse mechanism.
In this paper we propose to use a nonresponse version of what in the full response case is called the (design-based) optimal regression estimator. The underlying distance measure is a quadratic form with a more complex structure than the one leading to the GREG estimator (see Deville and Särndal (1992)). As it turns out there is also room for refinement in terms of the average response propensity (probability) when constructing the distance measure under nonresponse, which leadis to a modified "optimal" estimator.

1.1 Outline of the paper

Section 2 starts with an introduction to the calibration idea under full response before dealing with the nonresponse situation. In total there are three estimators of a population total which are considered: the GREG related estimator and two versions of the "optimal" estimator. Some theoretical results for the resulting bias follows. Section 3 contains a simulation study where simple random sampling and Poisson sampling are used for illustration. The Poisson design enables us to construct and investigate a situation where the auxiliary information is involved in the design as well as in the nonresponse mechansim. We end with concluding remarks in Section 4.

1.2 Notation and setup

We will start with a population U of size N from which we take a probability sample s of size n_s with inclusion probabilities $\pi_1, \dots, \pi_N$. Nonresponse means that we only observe the response set r of size n_r. Our aim is to estimate the study variable total $t_y = \sum_U y_k$. We assume access to an auxiliary variable vector $\boldsymbol{x}$ of dimension J, where either $\boldsymbol{x} = \boldsymbol{x}^*$ and $(\boldsymbol{x}^*_k)_{k \in U}$ are known (the population level) or $\boldsymbol{x} = \boldsymbol{x}^o$ and $(\boldsymbol{x}^o_k)_{k \in s}$ are known (the sample level) or possibly a mixture of these cases: $\boldsymbol{x} = (\boldsymbol{x}^{*'}, \boldsymbol{x}^{o'})'$.

2 Calibration estimation

2.1 Calibration estimators under full response

Starting with the full response situation ($r = s$) and following the procedure as established by Deville and Särndal (1992), the calibration estimator is defined as

$$\hat{t}_{y\,cal} = \sum_s w_{ks} y_k,$$

where the sample dependent weights w_{ks} are chosen so that

$$\sum_s w_{ks} \boldsymbol{x}_k = \boldsymbol{t}_{\boldsymbol{x}}, \text{ (the calibration equation)} \tag{1}$$

while also minimizing the quadratic distance measure

$$(\boldsymbol{w}_s - \boldsymbol{w}'_{0s}) \boldsymbol{R} (\boldsymbol{w}_s - \boldsymbol{w}_{0s}),$$

where $\boldsymbol{w}_s = (w_{ks})_{k \in s}$, $\boldsymbol{w}_{0s} = (1/\pi_k)_{k \in s} = (d_k)_{k \in s}$ and $\boldsymbol{R}$ is diagonal. (Alternative distance measures are considered in both Deville and Särndal (1992) and Haziza and Lesage (2016).)

In other words, given the constraint (1) the w_{ks} should be "as close as possible" to the design weights d_k, which is desirable since $\sum_s d_k y_k$ is an unbiased estimator of t_y.

The resulting weights are

$$\boldsymbol{w}_s = \boldsymbol{w}_{0s} + \boldsymbol{R}^{-1} \boldsymbol{x}' (\boldsymbol{X} \boldsymbol{R}^{-1} \boldsymbol{X}')^{-1} (\boldsymbol{t}_{\boldsymbol{x}} - \hat{\boldsymbol{t}}_{\boldsymbol{x}})$$

It turns out that the model assisted homoskedastic GREG estimator $\hat{t}_{y\,r}$ (Särndal, Swensson and Wretman (1992)) is a calibration estimator for which

$$\boldsymbol{R} = (\boldsymbol{w}_{0s}\boldsymbol{I}_{n_s})^{-1},$$

where I_{n_s} is the unit diagonal matrix of size n_s.
Another calibration estimator is the optimal regression estimator $\hat{t}_{y\,opt}$ (see e.g. Rao (1994) and Montanari (1998)), for which

$$\boldsymbol{R} = \left(\frac{\pi_{kl} - \pi_k\pi_l}{\pi_{kl}\pi_k\pi_l}\right)^{-1}_{k,l\in s},$$

as shown by Andersson and Thorburn (2005).
Asymptotically, this estimator has (in a design-based sense) minimum variance among linear regression type estimators.

2.2 Calibration estimators under nonresponse

In the nonresponse case, a possible calibration estimator is

$$\sum_r w_{kr} y_k,$$

where it should hold that

$$\sum_r w_{kr}\boldsymbol{x}_k = \boldsymbol{X}, \tag{2}$$

where $\boldsymbol{X} = \sum_U \boldsymbol{x}^*_k$, if the auxiliary information is known up to the population level. Otherwise, $\boldsymbol{X} = \sum_s d_k\boldsymbol{x}^o_k$, the unbiased estimator of $\boldsymbol{t_x}$. (We can also combine the two types of information in the constraint $\boldsymbol{X}$.)
For a variety of cases weights fulfilling the requirement (2) are presented by e.g. Särndal and Lundström (2005). Using the direct approach, where all information is used in one single calibration, we get

$$w_{kr} = d_k(1 + \boldsymbol{x}'_k(\sum_r d_k\boldsymbol{x}_k\boldsymbol{x}'_k)^{-1}(\boldsymbol{X} - \sum_r d_k\boldsymbol{x}_k)) \tag{3}$$

The resulting estimator will henceforth be denoted $\hat{t}_{y\,cal}$. (Other approaches, including two-step procedures, are presented and investigated by e.g. Andersson and Särndal (2016).)
An evident question to ask is: What is the underlying distance measure generating these weights? Särndal and Lundström (2005) do not comment on this particular issue, but according to Lundström and Särndal (1999),

we should choose "w_k 'as close as possible' to the d_k", which does not seem quite adequate under nonresponse. Going back to Lundström (1997) we will find that the corresponding distance measure *is* actually

$$(\boldsymbol{w}_r - \boldsymbol{w}_{0r})'(\boldsymbol{w}_{0r}\boldsymbol{I}_{n_r})^{-1}(\boldsymbol{w}_r - \boldsymbol{w}_{0r}),$$

where $\boldsymbol{w}_r = (w_{kr})_{k\in r}$ and $\boldsymbol{w}_{0r} = (d_k)_{k\in r}$.
If we have a random mechanism generating the response set r from the sample s with probabilities θ_k of inclusion, we can view the nonresponse situation as a two-phase design and this is the assumption we will make in the following. Then we should minimize the distance between w_{kr} and $d_k \cdot (1/\theta_k)$. Using some modelling θ_k can be estimated by $\hat{\theta}_k$, to be put to use for the distance minimization. But in this paper we will not go in the direction of model-based inference. In order to reduce the bias effect under nonresponse one could instead in the distance measure think of comparing w_{kr} not with d_k, but with $d_{k,alt} = d_k \cdot c$, where c is a constant larger than 1, aiming to compensate for the "average" nonresponse effect.
However, Lundström (1997) shows that in many important cases, namely when one can find a vector $\boldsymbol{\mu}$ for which $\boldsymbol{\mu'}\boldsymbol{x}_k = 1$, for all k, the multiplicative increase in $d_{k,alt}$ implies the same resulting calibration weights w_{kr}. This follows from the result that if $\boldsymbol{\mu'}\boldsymbol{x}_k = 1$, for all $k \in U$, we can simplify the expression of w_{kr} as

$$w_{kr} = d_k\boldsymbol{x}_k'(\sum_r d_k\boldsymbol{x}_k\boldsymbol{x}_k')^{-1}\boldsymbol{X}$$

Thus, we have an invariance property for the weights. The result holds also when the population is partitioned into groups and the initial weights are inflated with a constant within each group. Note that if we include a constant, e.g. "1", as a first component of the auxiliary vector $\boldsymbol{x}_k$, we can simply let $\boldsymbol{\mu'} = (1, 0, \ldots, 0)$ to achieve $\boldsymbol{\mu'}\boldsymbol{x}_k = 1$.
With this as a background we propose to use alternative "optimal" weights resulting from the distance measure

$$(\boldsymbol{w}_r - \boldsymbol{w}_{0r})'\left(\frac{\pi_{kl} - \pi_k\pi_l}{\pi_{kl}\pi_k\pi_l}\right)^{-1}_{k,l\in r}(\boldsymbol{w}_r - \boldsymbol{w}_{0r}),$$

leading to $\hat{t}_{y\,opt}$. (π_{kl} denotes the inclusion probability for the pair (k, l)). It is to be observed that as for the full response situation, there are cases for which the "optimal" weights are identical to (3), as e.g. under simple random sampling.

Using quotation marks around *optimal* is deliberate, but under full response *optimal* has a very clear meaning. As mentioned earlier, the optimal regression estimator has asymptotically minimum variance among linear regression estimators. Adding nonresponse where the nonresponse mechanism is at least partially unknown, makes it difficult to define optimality criteria in a proper way.
For this "optimal" measure it might be fruitful to replace d_k with $d_{k,alt}$, where we include in $d_{k,alt}$ the reciprocal of an estimate of the average response probability $\bar{\theta}_U = \sum_U \theta_k / N$. One simple candidate is

$$\hat{\bar{\theta}}_U = n_r / n_s,$$

thus yielding $d_{k,alt} = d_k \cdot (n_s/n_r)$. Another natural choice is

$$\hat{\bar{\theta}}_U = \sum_r d_k / \sum_s d_k, \tag{4}$$

since $E(\sum_s d_k) = N$ and $E(\sum_r d_k) = \sum_U \theta_k = N\bar{\theta}$, which lead to $E(\sum_r d_k / \sum_s d_k) \approx \bar{\theta}_U$. The resulting modified estimator is denoted by $\hat{t}_{y\,optm}$. (Also observe that $E(n_r/n_s) \approx \sum_U (\theta_k/d_k) / \sum_U 1/d_k$)
In the following simulation study we will focus on a sampling design where generally $\hat{t}_{y\,cal} \neq \hat{t}_{y\,opt}$, namely Poisson sampling. The independence of drawings simplifies the "optimal" distance measure:

$$\sum_r \frac{\pi_k^2}{1-\pi_k}(w_{kr} - d_k)^2 = \sum_r \frac{(w_{kr} - d_k)^2}{d_k(d_k - 1)}$$

and minimization yields

$$w_{kr} = d_k(1 + (d_k - 1)\boldsymbol{x}_k'(\sum_r d_k(1 - d_k)\boldsymbol{x}_k\boldsymbol{x}_k')^{-1}(\boldsymbol{X} - \sum_r d_k \boldsymbol{x}_k))$$

For the modified "optimal" estimator d_k is replaced by $d_{k\,alt} = d_k \cdot (1/\hat{\bar{\theta}})$, with $\hat{\bar{\theta}}_U$ as in (4).

2.2.1 Bias for calibration estimators under nonresponse

We can write $\hat{t}_{y\,cal}$ as

$$\hat{t}_{y\,cal} = \sum_r d_k y_k + \hat{\boldsymbol{B}}_{U;\theta}(\boldsymbol{X} - \sum_r d_k \boldsymbol{x}_k), \tag{5}$$

where $\hat{\boldsymbol{B}}_{U;\theta} = (\sum_r d_k \boldsymbol{x}_k' y_k)(\sum_r d_k \boldsymbol{x}_k \boldsymbol{x}_k')^{-1}$. In order to arrive at an approximate expression for the bias of $\hat{t}_{y\,cal}$ and subsequently $\hat{t}_{y\,opt}$ and

$\hat{t}_{y\,optm}$, we follow the derivation in Särndal and Lundström (2005) and first note that $\hat{t}_{y\,cal}$ can be rewritten as

$$\hat{t}_{y\,cal} = \sum_r d_k y_k + \boldsymbol{B}_{U;\theta}(\boldsymbol{X} - \sum_r d_k \boldsymbol{x}_k) + (\hat{\boldsymbol{B}}_{U;\theta} - \boldsymbol{B}_{U;\theta})(\boldsymbol{X} - \sum_r d_k \boldsymbol{x}_k),$$

where $\boldsymbol{B}_{U;\theta} = (\sum_U \theta_k \boldsymbol{x}'_k y_k)(\sum_U \theta_k \boldsymbol{x}_k \boldsymbol{x}'_k)^{-1}$.
If we let $\hat{t}_{y\,cal} - t_y = A_1 + A_2$, where
$A_1 = \sum_r d_k y_k - t_y + \boldsymbol{B}_{U;\theta}(\boldsymbol{X} - \sum_r d_k \boldsymbol{x}_k)$ and
$A_2 = (\hat{\boldsymbol{B}}_{U;\theta} - \boldsymbol{B}_{U;\theta})(\boldsymbol{X} - \sum_r d_k \boldsymbol{x}_k)$, it can further be shown that

$$A_1 = \sum_r d_k e_{\theta k} - \sum_U e_{\theta k} + \boldsymbol{B}^o_{U;\theta}(\sum_s d_k \boldsymbol{x}^o_k - \sum_U \boldsymbol{x}^o_k),$$

where $e_{\theta k} = y_k - \boldsymbol{B}_{U;\theta}\boldsymbol{x}_k$ and $\boldsymbol{B}^o_{U;\theta} = (\sum_U \theta_k \boldsymbol{x}^o_k \boldsymbol{x}^{o'}_k)^{-1} \sum_U \theta_k \boldsymbol{x}^o_k y_k$.
Then

$$E(\hat{t}_{y\,cal}) - t_y \approx E(A_1) = \sum_U \theta_k e_{\theta k} - \sum_U e_{\theta k} = -\sum_U (1 - \theta_k) e_{\theta k},$$

since it can be argued that $\hat{\boldsymbol{B}}_{U;\theta}$ is a consistent estimator of $\boldsymbol{B}_{U;\theta}$ and therefore $E(A_2) \approx 0$.
The approximation for the bias of $\hat{t}_{y\,cal}$ is called the nearbias:

$$\text{nearbias}(\hat{t}_{y\,cal}) = -\sum_U (1 - \theta_k) e_{\theta k}$$

The nearbias of $\hat{t}_{y\,cal}$ is zero if $\theta_k = 1$, for all $k \in U$ and/or $y_k = \boldsymbol{B}_{U;\theta}\boldsymbol{x}_k$, for all $k \in U$.
Then, if we consider $\hat{t}_{y\,opt}$, we have that

$$\hat{t}_{y\,opt} = \sum_r d_k y_k + (\boldsymbol{X} - \sum_r d_k \boldsymbol{x}_k)\hat{\boldsymbol{C}}_{U;\theta}, \tag{6}$$

where

$$\hat{\boldsymbol{C}}_{U;\theta} = \Big(\sum_{k\in r}\sum_{l\in r} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_{kl}} \frac{\boldsymbol{x}'_k}{\pi_k} \frac{y_l}{\pi_l}\Big)\Big(\sum_{k\in r}\sum_{l\in r} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_{kl}} \frac{\boldsymbol{x}_k}{\pi_k} \frac{\boldsymbol{x}'_l}{\pi_l}\Big)^{-1}$$

Since $\hat{t}_{y\,opt}$ can be written as (6), which is of the same form as for $\hat{t}_{y\,cal}$ in (5), we will again arrive at the nearbias expression

$$\text{nearbias}(\hat{t}_{y\,opt}) = -\sum_U (1 - \theta_k) e_{\theta k},$$

where $e_{\theta k} = y_k - \boldsymbol{C}_{U;\theta}\boldsymbol{x}_k$ and with θ_{kl} denoting the response probability for the pair (k, l):

$$\boldsymbol{C}_{U;\theta} = \Big(\sum_{k\in U}\sum_{l\in U}\theta_{kl}(\pi_{kl} - \pi_k\pi_l)\frac{\boldsymbol{x}'_k}{\pi_k}\frac{y_l}{\pi_l}\Big)\Big(\sum_{k\in U}\sum_{l\in U}\theta_{kl}(\pi_{kl} - \pi_k\pi_l)\frac{\boldsymbol{x}_k}{\pi_k}\frac{\boldsymbol{x}'_l}{\pi_l}\Big)^{-1}$$

If we use the alternative weighting $d_{k,alt} = d_k \cdot (1/\hat{\bar{\theta}}) = d_k \cdot (\sum_s d_k / \sum_r d_k)$, we get that

$$\text{nearbias}(\hat{t}_{y\,optm}) = E(\sum_r d_{k,alt}e_{\theta k} - \sum_U e_{\theta k}) \approx \sum_U \frac{\theta_k}{\bar{\theta}_U}e_{\theta k} - \sum_U e_{\theta k} = -\sum_U (1 - \frac{\theta_k}{\bar{\theta}_U})e_{\theta k},$$

where $\sum_U (1 - (\theta_k/\bar{\theta}_U)) = 0$.
Unless $\boldsymbol{\mu}'\boldsymbol{x}_k = 1$, for all $k \in U$, an equivalent expression can be obtained for $\hat{t}_{y\,cal}$. On the other hand, if the restriction $\boldsymbol{\mu}'\boldsymbol{x}_k = 1$, for all $k \in U$ *does* hold, it can be shown (Särndal and Lundström (2005)) that

$$\text{nearbias}(\hat{t}_{y\,cal}) = -\sum_U e_{\theta k},$$

which holds independently of the sampling design and which is a result completely in line with the aforementioned invariance property of the calibration weights.

3 A simulation study

Properties of the estimators were studied by means of a Monte Carlo simulation. We used an authentic population called KYBOK, which consists of $N = 832$ clerical municipalities in Sweden in 1992. (This population was also used for simulation purposes in Särndal and Lundström (2005) and Andersson and Särndal (2016).)
The study variable y_k is "Expenditure on administration and maintenance" ($t_y = 1\,023\,983$). The population is divided into four groups with respect to size, from the smallest to the largest. The group sizes are $N_1 = 218$, $N_2 = 272$, $N_3 = 290$ and $N_4 = 52$. The moon vector is $\boldsymbol{x}^o_k = (x^o_{1k}, \ldots, x^o_{4k})'$, where $x^o_{ik} = 1$ if the unit k belongs to population group i and otherwise 0, $i = 1, \ldots, 4$. The quantitative star variable x^*_k is the square root of "Revenue advances", which is highly positively correlated with y_k.
The sample size/expected sample size was 300 and we used the exponential response probability $\theta_k = 1 - \exp(-c \cdot x^*_k)$, $k \in U$, where c is chosen according to the desired average response probability; in this study varying

between 0.55 and 0.86. Two sampling designs have been considered separately: simple random sampling and Poisson sampling. In the latter case $\pi_k \propto x_k^*$. For each combination of design, sample size/expected sample size and average response probability, 10 000 samples were generated. For each such sample s, a response set r was created by performing independent Bernoulli trials with probability θ_k of success, $k \in s$.
An arbitrary estimator $\hat{t}_y$ is assessed by the empirical (simulation estimated) bias ($\widehat{B}$), variance ($\widehat{V}$) and mean squared error ($\widehat{MSE}$):

$$\widehat{B} = \widehat{E}(\hat{t}_y) - t_y = \frac{1}{K}\sum_{j=1}^{K} \hat{t}_{yj} - t_y$$

$$\widehat{V} = \frac{1}{K}\sum_{j=1}^{K}(\hat{t}_{yj} - \widehat{E}(\hat{t}_y))^2$$

$$\widehat{MSE} = \widehat{B}^2 + \widehat{V},$$

where $K = 10\,000$.
Observe that expressions as "the bias has increased" should be interpreted in the following as an increase of the bias in absolute value.

3.1 Results

As a benchmark for the study where auxiliary information is not used at the design stage, let us first consider the results for simple random sampling in Table 1. This is a case where $\hat{t}_{y\,cal} = \hat{t}_{y\,opt}$. (Actually, to get equality the "star" information is $\boldsymbol{x}_k^* = (1, x_k^*)'$ for $\hat{t}_{y\,cal}$.) As expected the results for the bias and variance are substantially worse for $\bar{\theta}_U = 0.70$ than for $\bar{\theta}_U = 0.86$.
Looking instead at the results in Table 1 for Poisson sampling, we can first observe that for $\hat{t}_{y\,cal}$, with the exception of the bias when $\bar{\theta}_U = 0.70$, both the bias and variance are higher than under simple random sampling. Thus, although the Poisson sampling design incorporates the highly explanatory auxiliary variable x_k^*, $\hat{t}_{y\,cal}$ does not make sufficient use of it. $\hat{t}_{y\,opt}$ on the other hand benefits from this design and shows improvement in the bias and variance, resulting in a clear reduction of the mean squared error.
Finally in Table 1 we observe a great improvement using $\hat{t}_{y\,optm}$ instead of $\hat{t}_{y\,opt}$ regarding the bias. However, compared with $\hat{t}_{y\,opt}$ the variance of $\hat{t}_{y\,optm}$ increases, leading to a trade-off effect, which implies only a modest

reduction of the mean squared error. The most interesting effect of $\hat{t}_{y\,optm}$ is the change of sign for the bias: negative when $\bar{\theta}_U = 0.86$ and positive (as opposed to the other estimators) when $\bar{\theta}_U = 0.70$. So, what happens in between? Is there a monotonic behaviour of the bias for changing average response probabilities? The answer to the last question is yes and Table 2 shows additional results for the bias of $\hat{t}_{y\,optm}$ for other values of the average response probabilities. The bias is actually zero for this estimator when $\bar{\theta}_U \simeq 0.79$.

Table 1. Empirical bias ($\widehat{B}$), variance ($\widehat{V}$) and mean squared error ($\widehat{MSE}$) for $\hat{t}_{y\,cal}$ (Cal) , $\hat{t}_{y\,opt}$ (Opt) and $\hat{t}_{y\,optm}$ (Optm) under simple random sampling ($n = 300$) and Poisson sampling ($E(n) = 300$) with average response probabilities 0.86 and 0.70.

Simple random sampling (Cal=Opt)

	$\widehat{B}$	$\widehat{V}$	$\widehat{MSE}$
$\bar{\theta}_U = 0.86$			
Cal	$-2.44 \cdot 10^4$	$8.40 \cdot 10^8$	$1.44 \cdot 10^9$
$\bar{\theta}_U = 0.70$			
Cal	$-4.00 \cdot 10^4$	$9.59 \cdot 10^8$	$2.57 \cdot 10^9$

Poisson sampling

	$\widehat{B}$	$\widehat{V}$	$\widehat{MSE}$
$\bar{\theta}_U = 0.86$			
Cal	$-2.50 \cdot 10^4$	$3.89 \cdot 10^9$	$4.52 \cdot 10^9$
Opt	$-2.41 \cdot 10^4$	$6.53 \cdot 10^8$	$1.23 \cdot 10^9$
Optm	$-6.77 \cdot 10^3$	$1.00 \cdot 10^9$	$1.05 \cdot 10^9$
$\bar{\theta}_U = 0.70$			
Cal	$-3.87 \cdot 10^4$	$4.40 \cdot 10^9$	$5.90 \cdot 10^9$
Opt	$-2.97 \cdot 10^4$	$7.24 \cdot 10^8$	$1.61 \cdot 10^9$
Optm	$3.72 \cdot 10^3$	$1.58 \cdot 10^9$	$1.59 \cdot 10^9$

Table 2. Empirical bias $(\widehat{B}) * 10^{-4}$ for $\hat{t}_{y\,opt}$ (Opt) and $\hat{t}_{y\,optm}$ (Optm) under Poisson sampling ($E(n) = 300$) with average response probabilities from 0.86 to 0.55.

$\bar{\theta}_U$	0.86	0.75	0.65	0.55
Opt	-2.41	–2.67	-3.43	-5.08
Optm	-0.677	0.144	0.535	0.669

4 Concluding remarks

The family of linear calibration techniques in survey sampling contains a variety of alternative weightings under full response, including GREG estimators and the optimal regression estimator. The nonresponse situation offers still more options and challenges and we have studied the "optimal" estimator while also taking into account average response propensities (probabilities). The approach has been design-based since the "optimal" estimator can be motivated by asymptotic argumentation and we have furthermore not used any modelling for the response propensities. The results are encouraging, especially concerning reduction of the bias for the suggested estimator. Further work will include the construction of a variance estimator, which should be valid conditionally on the size of the response set.

References

Andersson, P. G., and Thorburn, D. (2005), "An optimal calibration distance leading to the optimal regression estimator," *Survey Methodology*, 31(1), 95-99.

Andersson, P. G., and Särndal, C. E. (2016), "Calibration for nonresponse treatment: In one or two steps?," *Statistical Journal of the IAOS*, 32, 375-381.

Deville, J. C., and Särndal, C-E. (1992), "Calibration estimators in survey sampling," *Journal of the American Statistical Association*, 87, 376-382.

Haziza, D., and Lesage, É. (2016), "A discussion of weighting procedures for unit nonresponse," *Journal of Official Statistics*, 32(1), 129-145.

Lundström, S. (1997), *Calibration as a standard method for treatment of nonresponse*, Ph.D. thesis, Department of Statistics, Stockholm University.

Lundström, S., and Särndal, C. E. (1999), "Calibration as a standard method for treatment of nonresponse," *Journal of Official Statistics*, 13, 305-327.

Montanari, G. E. (1998), "On regression estimation of finite population means," *Survey Methodology*, 23, 69-77.

Rao, J. N. K. (1994)," Estimating totals and distribution functions using auxiliary information at the estimation stage," *Journal of Official Statistics*, 10, 153-165.

Särndal, C. E., and Lundström, S. (2005), *Estimation in surveys with nonresponse*. Chichester, UK: Wiley.