WORKING PAPER SERIES

WORKING PAPER NO 9, 2005

ESI

The variance of some common estimators and its components under nonresponse

By

Sara Tångdahl

http://www.oru.se/esi/wps

SE-701 82 Örebro
SWEDEN

ISSN 1403-0586

The variance of some common estimators and its components under nonresponse

Sara Tångdahl

Abstract

In most surveys, the risk of nonresponse is a factor taken into account at the planning stage. Commonly, resources are set aside for a follow-up procedure which aims at reducing the nonresponse rate. However, we should pay attention to the effect of nonresponse, rather than the nonresponse rate itself. When considering nonresponse error, i.e. bias and variance, it is not obvious that the resources spent on nonresponse rate reduction efforts are time and money well spent. In this paper we address this issue, continuing the work begun in Tångdahl (2004), now focusing on the effect of follow-ups on estimator variance. The components of the variance for some common estimators are derived under a setup that allows us to take into account the data collection process, and follow-up efforts in particular.

Contents

1 Introduction

1.1 The problem

In survey planning, decisions must be made on how to allocate limited resources between various survey operations, all with the purpose of minimizing the total survey error. Examples of such survey operations are frame construction, questionnaire testing and editing. In particular, if we have reason to suspect that nonresponse will occur in the survey, efforts to deal with this must be decided upon. A common strategy for dealing with nonresponse is to use follow-ups or callbacks with the purpose of reducing the nonresponse rate, implicitly assuming that a lower nonresponse rate results in smaller nonresponse error. This leads survey administrators to consider reduction of the nonresponse rate as a major goal during data collection, despite evidence from empirical studies that indicate that there is not necessarily an immediate relation between nonresponse rate reduction and decrease in nonresponse error. Examples of such studies are Curtin, Presser, and Singer (2000) and Keeter, Miller, Kohut, Groves, and Presser (2000). Hence, before trying to minimize the nonresponse rate within the available budget, we should focus on the effect of nonresponse. The main concern in the presence of nonresponse is the obvious risk of bias, but the variance of estimators will also be inflated. In addition, not only point estimators, but also variance estimators may be biased.

Strategies to deal with nonresponse are generally combinations of preventive, reductive or adjustive efforts. The meaning of these terms should be obvious; preventive actions are taken prior to data collection and aim at preventing nonresponse from ever occurring, reductive actions take place during, or rather, as part of the data collection with the purpose of reducing the nonresponse that ultimately occurs. Adjustment include any attempts to compensate for the nonresponse at the estimation stage.

In the allocation of resources between different efforts, we need to consider whether the effect of the efforts really outweigh the costs. To be able to balance reduction efforts, adjustment efforts and costs under budget constraints we will need to study the effect of reductive measures on estimator properties. Some of these issues are discussed in Tångdahl (2004), where the purpose is to study and evaluate the effect on estimator bias of nonresponse rate reduction efforts used as part of the data collection process. A framework based on a fairly general response distribution that incorporates the effect of reduction efforts is introduced. Approximate expressions for the change in nonresponse bias for six generally formulated estimators under certain assumptions about the true response distribution are derived, and it is shown that the bias does not necessarily decrease as the nonresponse rate decreases.

The study of the change in bias alone, however, does not provide a complete evaluation of the data collection procedure. In considering excluding some of the nonresponse rate reduction efforts, the effect on the variance of a given point estimator and the change in the variance must also be studied. In this paper, the estimator variances are evaluated under the assumed true, fairly general, sequence of response distributions, *RDs*, introduced in Tångdahl (2004). A complete discussion of the effect of truncation of field efforts must also be based on costs. We will return to this and related issues in subsequent papers. The reasonable view taken here is that the individual response probabilities are subject to influence by the survey operations, and in this case in particular by the nonresponse reduction efforts. Thus, in the model we adopt, the response probabilities are allowed to change during the data collection process. We define a sequence of response distributions, all influenced by both general survey conditions and the specific survey settings, but also by the nonresponse reductive measures taken by the survey administrator. This means that for every possible choice of time to terminate data collection, we define a corresponding response distribution, *RD*. Formal details are given in section 1.2. In section 2, the estimators under study are presented briefly. Section 3 contains a summary of the approach and results in Tångdahl (2004). A general discussion of the estimator variance and its components is given in section 4, while explicit variances of the studied estimators are found in section 5. Section 6 is a discussion of alternative error measures, and section 7 contains a summary and some conclusions.

1.2 Notation and definitions

Let $U = \{1, \ldots, k, \ldots, N\}$ be the finite population of interest and let N denote the number of elements in U. The study variable is denoted y, with the value y_k for element k. The parameter of interest is the population total $t_y = \sum_U y_k$ where $\sum_U$ is shorthand for $\sum_{k \in U}$. A random sample s of size n is selected from U according to the design $p(\cdot)$ with positive first and second order inclusion probabilities π_k and π_{kl}. We will assume that direct element sampling is used and allow the special case $\pi_k = 1$ for all $k \in U$.

For notational convenience, division by π_k will be denoted by $\check{}$, e.g. $\check{y}_k = y_k/\pi_k$.

Let $a = 1, \ldots, A$ denote an arbitrary point of time during the data collection period, and let A denote the final data collection time point in the current survey setup. Ideally, all elements in the sample respond to the initial survey request, but commonly in practice some nonresponse still remains at time A, and consequently at time a. Since we limit the discussion to only one study variable, we will not need to make a distinction between item and unit nonresponse. As the data collection progresses, successive response sets $r^{(1)} \subset \cdots \subset r^{(a)} \subset \ldots r^{(A)} \subset s$ are generated, where $r^{(A)}$ is the final response set in the current survey setup. The response set $r^{(a)}$ is assumed to have been generated by response distribution $RD^{(a)}$, with individual response probabilities $\theta^{(a)}_{k|s} = Pr(k \in r^{(a)}|s)$. We will assume that elements respond independently of each other and that $\theta^{(a)}_{k|s} \geq \theta^{(a-1)}_{k|s}$ for all k. The assumption of independence implies that $Pr(k\&l \in r^{(a)}|s) = \theta^{(a)}_{k|s}\theta^{(a)}_{l|s}$ for every pair $k \neq l \in s$.

Throughout, quantities based on response set $r^{(a)}$, the response set at time a, will be denoted with a superscript (a).

Most of the estimators studied in the following utilize auxiliary information. In a situation with nonresponse, as in a two-phase sampling situation, the auxiliary information can be available on two levels as follows:

1. Let $\mathbf{x}_{1k}$ be a J_1-vector with values known for all $k \in s$ and with the population total $\sum_U \mathbf{x}_{1k}$ known.

2. Let $\mathbf{x}_{2k}$ be a J_2-vector with values known for all $k \in s$. The complete information at this level is combined in the vector $\mathbf{x}_k = (\mathbf{x}'_{1k}, \mathbf{x}'_{2k})'$ of length J.

Note that the value of the study variable y_k is known only for $k \in r^{(a)}$. In the most general case, both $\mathbf{x}_1$ and $\mathbf{x}$ are used as an attempt to improve estimation. There are several ways to combine the available information. Two important special cases are the case where only $\mathbf{x}_1$ is used, which will be referred to as special case 1, and the case where no auxiliary information is known at the population level, referred to as special case 2.

2 The estimators under study

There are many different methods available for handling nonresponse at the estimation stage, one being reweighting. It is common to refer to this as adjusting for nonresponse, although the adjustment is not necessarily successful. In Tångdahl (2004), estimators based on two reweighting methods, weighting by response homogeneity groups (RHGs) and calibration for nonresponse were studied. These techniques were chosen because they represent two large and widely used classes of estimators. With appropriate auxiliary information, they are both powerful methods. The approaches differ in how the nonresponse is handled. In the RHG approach, the response distribution is explicitly modeled, allowing the use of two-phase sampling methods. In the calibration for nonresponse approach, no modeling of the response distribution is done. Since this paper is a continuation of the work in Tångdahl (2004), the same estimators will be studied here.

The RHG model is formulated as follows: assume that a partitioning of the realized sample can be made such that response probabilities are constant within groups s_h, $h = 1, \ldots, H_s$. It is also assumed that the response probabilities are positive for all elements and that elements respond independently. The partitioning need not be the same for different samples, but for a given sample, the grouping is always the same. In the present particular setting, we will assume that the same model applies for any given time a, only the levels change. The model can be stated formally as

$$Pr(k \in r^{(a)}|s) = \theta^{(a)}_{k|s} = \theta^{(a)}_{hs} > 0 \text{ for all } k \in s_h$$
$$Pr(k\&l \in r^{(a)}|s) = \theta^{(a)}_{k|s}\theta^{(a)}_{l|s} \text{ for all } k \neq l \in s$$

for $h = 1, \ldots, H_s$ and for a given time $a = 1, \ldots, A$.

Let n_h be the size of s_h and let $r^{(a)}_h$ of size $m^{(a)}_h$ be the responding subset of s_h at time a. Conditioning on the response count vector $\mathbf{m}^{(a)}$, estimated first and second order response probabilities are

$$\hat{\theta}^{(a)}_{k|s} = \hat{\theta}^{(a)}_{hs} = \frac{m^{(a)}_h}{n_h} \text{ for all } k \in s_h$$

and

$$\hat{\theta}^{(a)}_{kl|s} = \begin{cases} \dfrac{m^{(a)}_h(m^{(a)}_h - 1)}{n_h(n_h - 1)} & \text{for } k \neq l \in s_h \\ \dfrac{m^{(a)}_h}{n_h}\dfrac{m^{(a)}_{h'}}{n_{h'}} & \text{for } k \in s_h, l \in s_{h'}; h \neq h' \end{cases}$$

Naturally, the properties of estimators based on RHGs depend on the formulation of the groups. To eliminate the nonresponse bias, the statistician must be able to identify groups such that the model assumptions hold, or at least nearly so. The choice and use of auxiliary information will also influence the estimator properties since strong auxiliary information can provide a robustness against RHG model specification error as well as a reduction of the estimator variance. That strong auxiliary information can lead to a substantial reduction of the sampling variance is well known; evidence can be found in most standard textbooks in survey sampling as well as in numerous papers.

An estimator based on RHGs but with no use of auxiliary information is the simple RHG estimator

$$\hat{t}^{(a)}_{yc\pi^*} = \textstyle\sum_{r^{(a)}} \dfrac{y_k}{\pi_k \hat{\theta}^{(a)}_{k|s}} = \displaystyle\sum_{h=1}^{H_s} \frac{n_h}{m_h^{(a)}} \textstyle\sum_{r_h^{(a)}} \check{y}_k \tag{1}$$

Using auxiliary information we can form a regression type estimator. In the general case, where both $\mathbf{x}_1$ and $\mathbf{x}$ are available, the regression based RHG estimator is

$$\hat{t}^{(a)}_{ycreg} = \textstyle\sum_U \hat{y}^{(a)}_{1k} + \textstyle\sum_s \dfrac{\hat{y}^{(a)}_k - \hat{y}^{(a)}_{1k}}{\pi_k} + \displaystyle\sum_{h=1}^{H_s} \frac{n_h}{m_h^{(a)}} \textstyle\sum_{r_h^{(a)}} \dfrac{y_k - \hat{y}^{(a)}_k}{\pi_k} \tag{2}$$

with predictions

$$\hat{y}^{(a)}_{1k} = \mathbf{x}'_{1k}\hat{\mathbf{B}}^{(a)}_{1r} = \mathbf{x}'_{1k}\left(\sum_{h=1}^{H_s} \frac{n_h}{m_h^{(a)}} \textstyle\sum_{r_h^{(a)}} \dfrac{\mathbf{x}_{1k}\mathbf{x}'_{1k}}{\sigma^2_{1k}\pi_k}\right)^{-1} \displaystyle\sum_{h=1}^{H_s} \frac{n_h}{m_h^{(a)}} \textstyle\sum_{r_h^{(a)}} \dfrac{\mathbf{x}_{1k} y_k}{\sigma^2_{1k}\pi_k}$$

and

$$\hat{y}^{(a)}_{k} = \mathbf{x}'_{k}\hat{\mathbf{B}}^{(a)}_{r} = \mathbf{x}'_{k}\left(\sum_{h=1}^{H_s} \frac{n_h}{m_h^{(a)}} \textstyle\sum_{r_h^{(a)}} \dfrac{\mathbf{x}_{k}\mathbf{x}'_{k}}{\sigma^2_{k}\pi_k}\right)^{-1} \displaystyle\sum_{h=1}^{H_s} \frac{n_h}{m_h^{(a)}} \textstyle\sum_{r_h^{(a)}} \dfrac{\mathbf{x}_{k} y_k}{\sigma^2_{k}\pi_k}$$

The weights σ^2_{1k} and σ^2_k are residual variances in assumed, hypothetical, regression models of y on $\mathbf{x}_1$ and $\mathbf{x}$, respectively. Two important special cases of the regression estimator (2) are

Special case 1:
No additional auxiliary information is available at the sample level. We then

have $\mathbf{x}_k = (\mathbf{x}'_{1k}, \mathbf{x}'_{2k})' = \mathbf{x}_{1k}$, so the estimator becomes

$$\hat{t}^{(a)}_{ycreg1} = \textstyle\sum_U \hat{y}^{(a)}_{1k} + \displaystyle\sum_{h=1}^{H_s} \frac{n_h}{m^{(a)}_h} \textstyle\sum_{r^{(a)}_h} \displaystyle\frac{y_k - \hat{y}^{(a)}_{1k}}{\pi_k} \tag{3}$$

Special case 2:
The only available auxiliary information is $\mathbf{x}_k$, known for $k \in s$. Assuming that $\sigma^2_{1k} = \sigma^2_k$ for all k, the predictions $\hat{y}^{(a)}_{1k}$ drop out and the estimator becomes

$$\hat{t}^{(a)}_{ycreg2} = \textstyle\sum_s \displaystyle\frac{\hat{y}^{(a)}_k}{\pi_k} + \sum_{h=1}^{H_s} \frac{n_h}{m^{(a)}_h} \textstyle\sum_{r^{(a)}_h} \displaystyle\frac{y_k - \hat{y}^{(a)}_k}{\pi_k} \tag{4}$$

Estimators (1)–(4) are given in Särndal, Swensson, and Wretman (1992).

In the calibration for nonresponse approach, only special cases 1 and 2 will be studied. The general case with two auxiliary vectors is excluded since the properties of such an estimator have not yet been fully investigated and consequently it is not widely used, although some recent results can be found in Lundström and Särndal (2005). The calibration estimators studied here are:

$$\hat{t}^{(a)}_{ycal1} = \textstyle\sum_{r^{(a)}} v^{(a)}_{1kr} \displaystyle\frac{y_k}{\pi_k} \tag{5}$$

with

$$v^{(a)}_{1kr} = 1 + \left(\textstyle\sum_U \mathbf{x}_{1k} - \sum_{r^{(a)}} \displaystyle\frac{\mathbf{x}_{1k}}{\pi_k} \right)' \left(\textstyle\sum_{r^{(a)}} \displaystyle\frac{c_k \mathbf{x}_{1k} \mathbf{x}'_{1k}}{\pi_k} \right)^{-1} c_k \mathbf{x}_{1k} \tag{6}$$

in special case 1 and

$$\hat{t}^{(a)}_{ycal2} = \textstyle\sum_{r^{(a)}} v^{(a)}_{2kr} \displaystyle\frac{y_k}{\pi_k} \tag{7}$$

with

$$v^{(a)}_{2kr} = 1 + \left(\textstyle\sum_s \displaystyle\frac{\mathbf{x}_k}{\pi_k} - \textstyle\sum_{r^{(a)}} \displaystyle\frac{\mathbf{x}_k}{\pi_k} \right)' \left(\textstyle\sum_{r^{(a)}} \displaystyle\frac{c_k \mathbf{x}_k \mathbf{x}'_k}{\pi_k} \right)^{-1} c_k \mathbf{x}_k \tag{8}$$

in special case 2. The factors c_k are specified by the statistician, usually chosen to be a linear function of the auxiliary vector.

3 Nonresponse bias at time a

Under an assumed true response distribution, Tångdahl (2004) derives general expressions for the nonresponse bias for each of the six estimators given in section 2. Let $\hat{t}_{yc}^{(a)}$ be an estimator for the population total t_y, based on response set $r^{(a)}$, and let $\hat{t}_{ys}$ be the corresponding (approximately) unbiased full response estimator. The bias for $\hat{t}_{yc}^{(a)}$ can then be written

$$B(\hat{t}_{yc}^{(a)}) = E_p\left[E_{RD^{(a)}}(\hat{t}_{yc}^{(a)})\right] - t_y \approx E_p\left[E_{RD^{(a)}}(\hat{t}_{yc}^{(a)}) - \hat{t}_{ys}\right] \tag{9}$$

where E_p denotes expectation with respect to the sampling stage, and $E_{RD^{(a)}}$ is expectation with respect to the response distribution $RD^{(a)}$, given s. Since the estimators are nonlinear functions of the unknown response probabilities, the resulting conditional bias expressions are approximate. Using the defined sequence of response distributions, Tångdahl also derives expressions for the difference in approximate bias between estimators $\hat{t}_{yc}^{(a-1)}$ and $\hat{t}_{yc}^{(a)}$. By using

$$B(\hat{t}_{yc}^{(a)}) - B(\hat{t}_{yc}^{(a-1)}) = E_p\left[E_{RD^{(a)}}(\hat{t}_{yc}^{(a)}) - E_{RD^{(a-1)}}(\hat{t}_{yc}^{(a-1)})\middle|\, s\right],$$

explicit expressions for the bias and the bias difference are derived for each of the given estimators, leaving the expectation with respect to the sampling stage indeterminate. Due to the nonlinearity of the estimators, large sample approximations are used. The consequences for the bias difference is discussed under specific assumptions about the response distributions $RD^{(a-1)}$ and $RD^{(a)}$.

4 Variance at time a, the approach

In deriving the total variance of the point estimators, a common decomposition of the total estimation error will be useful. Let $\hat{t}_{yc}^{(a)}$ be the point estimator to be used at time a in the presence of nonresponse, and let $\hat{t}_{ys}$ be the corresponding full response estimator. The total estimation error can then be written

$$\hat{t}_{yc}^{(a)} - t_y = (\hat{t}_{yc}^{(a)} - \hat{t}_{ys}) - (t_y - \hat{t}_{ys}) \tag{10}$$

Using decomposition (10) and rules for conditional expectations, it is easily seen that the total variance $V\left(\hat{t}_{yc}^{(a)}\right) = E\left[\hat{t}_{yc}^{(a)} - E(\hat{t}_{yc}^{(a)})\right]^2$, can be written as

$$\begin{aligned} V\left(\hat{t}_{yc}^{(a)}\right) =& V_p\left(\hat{t}_{ys}\right) + V_p\left[E_{RD^{(a)}}\left(\hat{t}_{yc}^{(a)} - \hat{t}_{ys}|s\right)\right] \\ &+ E_p\left[V_{RD^{(a)}}\left(\hat{t}_{yc}^{(a)}|s\right)\right] + 2Cov_p\left[E_{RD^{(a)}}\left(\hat{t}_{yc}^{(a)} - \hat{t}_{ys}|s\right), \hat{t}_{ys}\right] \\ =& V_1 + V_2 + V_3 + V_4 \end{aligned} \tag{11}$$

where the first term is the sampling variance, and the last three are due to the nonresponse. The components are

V_1 =sampling variance
V_2 =variance of the conditional nonresponse bias
V_3 =(expected value of) conditional nonresponse variance
V_4 =covariance between conditional nonresponse bias and
the full response estimator

The sampling variance $V_p\left(\hat{t}_{ys}\right)$ is not of primary interest here since focus will be on *change* in total variance, which is a function only of the three nonresponse components in the variance. It will be important however when the cost and effect of the nonresponse rate reduction procedure is weighed against other efforts to increase the total survey quality. Such a discussion is beyond the scope of this paper.

In Särndal et al. (1992), ch. 15, the total variance and variance estimators for all the RHG estimators used here, are presented. However, this is done under the assumption that the response homogeneity groups model holds. Variance estimators, but no expression for the true total variance, have been presented for the calibration estimator, drawing on two-phase sampling theory and the similarity between the calibration estimator and the two-phase regression estimator. What happens when the model assumptions do not hold? In this section we present and discuss components of the total variance in more general terms, while expressions for the variance components for the six studied estimators are derived in section 5, under the assumed true, general, response distribution presented in section 1.2.

4.1 Sampling variance

The sampling variance is the variance of the corresponding full response estimator in each case, i.e. the estimator we would have used if we had no nonresponse, given the auxiliary information (if any) used in the reweighting estimator. The simple RHG estimator and the estimators under special case 2 reduce to the Horvitz-Thompson estimator in the full response case. In the general case and in special case 1, the full response estimator is the regression estimator with only $\mathbf{x}_1$ as auxiliary vector. Consequently, $\mathbf{x}_2$ has no influence on the sampling variance.

Results on the sampling variance of each of the studied estimators are well known. The sampling variance will depend on the strength of the auxiliary vector, $\mathbf{x}_1$, and of course on the design and sample size. A larger correlation between $\mathbf{x}_1$ and y will give smaller residuals, and thus a smaller sampling variance. For the estimators whose corresponding full response estimators do not utilize auxiliary information, the only way to reduce the sampling variance is to increase the sample size or to use a more efficient design.

In section 5, only expressions for a general design are given. Explicit expressions must be worked out for each specific design.

4.2 Variance of the conditional nonresponse bias

Approximate expressions for $E_{RD^{(a)}}(\hat{t}_{yc}^{(a)} - \hat{t}_{ys})$, the conditional nonresponse bias, are derived in Tångdahl (2004) for each of the studied estimators. A small simulation study indicates that the variance of this conditional bias is numerically small compared to the sampling variance and to the expected value of the conditional nonresponse variance. We do not derive explicit expressions for this variance component. It has a complex structure, even in the case of the simple RHG estimator, and will thus not be very informative.

4.3 Conditional nonresponse variance

The component $E_p \left[V_{RD^{(a)}} \left(\hat{t}_{yc}^{(a)} | s \right) \right]$ in (11) represents the increase in estimator variance caused by nonresponse, if the point estimator has zero nonresponse bias and is often called the *nonresponse variance*. $V_{RD^{(a)}} \left(\hat{t}_{yc}^{(a)} | s \right)$ will be referred to as the *conditional (on s) nonresponse variance.*

All of the estimators used here are complex (nonlinear) functions of the response probabilities, so exact explicit expressions for their conditional nonresponse variances cannot be found. We will instead use Taylor linearization to derive approximate variance expressions, valid for large response homogeneity groups. The linearization techniques for two-phase regression estimators in e.g. Särndal et al. (1992) can not be used since we are not guaranteed unbiased estimation under nonresponse. Expressions for the conditional nonresponse variance are derived in section 5.

4.4 Covariance between the conditional nonresponse bias and the full response estimator

A small simulation study indicates that this component of the total variance is generally numerically small compared to the sampling variance and the expected value of the conditional nonresponse variance. Explicit expressions are not derived.

5 Estimator variances

As shown in section 4, the variance of an estimator $\hat{t}_{yc}^{(a)}$ is the sum of four components. In this section, the total variance is given for each estimator presented in section 2. Explicit unconditional expressions cannot be derived without making specific assumptions about the true response distribution, and, for the RHG estimators, about the response homogeneity groups, so the expectations with respect to the sampling phase are left indeterminate. Furthermore, due to the complexity of the estimators, most of the expressions are approximate.

5.1 The simple RHG estimator

The estimator corresponding to $\hat{t}^{(a)}_{yc\pi^*}$ in the full response case is the ordinary π-estimator $\hat{t}_{y\pi} = \sum_s \check{y}_k$. The variance of $\hat{t}^{(a)}_{yc\pi^*}$ is the sum of the four components

$$V_1 = V_p(\hat{t}_{y\pi}) = \sum\sum_U \Delta_{kl}\check{y}_k\check{y}_l \tag{12}$$

$$\begin{aligned} V_2 &= V_p\left(E_{RD^{(a)}}(\hat{t}^{(a)}_{yc\pi^*} - \hat{t}_{y\pi}|s)\right) \\ &\approx V_p\left(\sum_{h=1}^{H_s}(n_h-1)\frac{1}{\bar{\theta}^{(a)}_{s_h}}S_{\theta^{(a)}\check{y}s_h}\right) \end{aligned} \tag{13}$$

with $\bar{\theta}^{(a)}_{s_h} = \dfrac{1}{n_h}\sum_{s_h}\theta^{(a)}_{k|s}$ and $S_{\theta^{(a)}\check{y}s_h} = \dfrac{1}{n_h-1}\left(\sum_{s_h}\theta^{(a)}_{k|s}\check{y}_k - \dfrac{1}{n_h}\left(\sum_{s_h}\theta^{(a)}_{k|s}\right)\left(\sum_{s_h}\check{y}_k\right)\right)$

$$\begin{aligned} V_3 &= E_p\left(V_{RD^{(a)}}(\hat{t}^{(a)}_{yc\pi^*}|s)\right) \\ &\approx E_p\left(\sum_{h=1}^{H_s}\left(\frac{n_h}{\hat{t}^{(a)}_{\theta h}}\right)^2\sum_{s_h}\theta^{(a)}_{k|s}(1-\theta^{(a)}_{k|s})\left(\check{y}_k - \frac{\hat{t}^{(a)}_{y\theta h}}{\hat{t}^{(a)}_{\theta h}}\right)^2\right) \end{aligned} \tag{14}$$

with $\hat{t}^{(a)}_{y\theta h} = \sum_{s_h}\theta^{(a)}_{k|s}\check{y}_k$ and $\hat{t}^{(a)}_{\theta h} = \sum_{s_h}\theta^{(a)}_{k|s}$

$$\begin{aligned} V_4 &= 2Cov_p\left(E_{RD^{(a)}}(\hat{t}^{(a)}_{yc\pi^*} - \hat{t}_{y\pi}|s), \hat{t}_{y\pi}\right) \\ &= 2Cov_p\left(\sum_{h=1}^{H_s}(n_h-1)\frac{S_{\theta^{(a)}\check{y}s_h}}{\bar{\theta}^{(a)}_{s_h}}, \hat{t}_{y\pi}\right) \end{aligned} \tag{15}$$

The expressions (13) and (15) follow directly from Tångdahl (2004). The third variance component, the expected value of the conditional nonresponse variance, is arrived at by the following reasoning. We note that $\hat{t}_{yc\pi^*}$ can be expressed as the sum of ratios of estimated totals:

$$\hat{t}^{(a)}_{yc\pi^*} = \sum_{h=1}^{H_s}\frac{n_h}{m^{(a)}_h}\sum_{r^{(a)}_h}\check{y}_k = \sum_{h=1}^{H_s} n_h\frac{\hat{t}^{(a)}_{yr_h}}{\hat{t}^{(a)}_{\theta r_h}}$$

where $\hat{t}^{(a)}_{yr_h} = \sum_{s_h} R^{(a)}_{k|s}\check{y}_k$, $\hat{t}^{(a)}_{\theta r_h} = m^{(a)}_h = \sum_{s_h} R^{(a)}_{k|s}$ and $R^{(a)}_{k|s}$ is a response indicator variable, defined as

$$R^{(a)}_{k|s} = \begin{cases} 1 & \text{if } k \in r|s \\ 0 & \text{if not} \end{cases}$$

with expected value $\theta^{(a)}_{k|s}$.

Using first order Taylor expansion results for estimation of a ratio, a linear large sample approximation to $\hat{t}^{(a)}_{yc\pi^*}$ is given by

$$\begin{aligned} \hat{t}^{(a)}_{yc\pi^*} &\approx \sum_{h=1}^{H_s} \frac{n_h \hat{t}^{(a)}_{y\theta h}}{\hat{t}^{(a)}_{\theta h}} + \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta h}} \left(\hat{t}^{(a)}_{ys_h} - \frac{\hat{t}^{(a)}_{y\theta h}}{\hat{t}^{(a)}_{\theta h}} \hat{t}^{(a)}_{\theta s_h} \right) \\ &= \hat{t}^{(a)}_{yc\theta} + \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta h}} \sum\nolimits_{s_h} R^{(a)}_{k|s} \left(\check{y}_k - \frac{\hat{t}^{(a)}_{y\theta h}}{\hat{t}^{(a)}_{\theta h}} \right) \end{aligned} \tag{16}$$

where $\hat{t}^{(a)}_{y\theta h} = \sum_{s_h} \theta^{(a)}_{k|s}\check{y}_k$ and $\hat{t}^{(a)}_{\theta h} = \sum_{s_h} \theta^{(a)}_{k|s}$ are the expected values under $RD^{(a)}$, conditional on s, of $\hat{t}^{(a)}_{yr_h}$ and $\hat{t}^{(a)}_{\theta r_h}$, respectively. For notational convenience, we let $\hat{t}^{(a)}_{yc\theta} = \sum_{h=1}^{H_s} n_h \hat{t}^{(a)}_{y\theta h} / \hat{t}^{(a)}_{\theta h}$.

An approximation to the conditional nonresponse variance of $\hat{t}^{(a)}_{yc\pi^*}$ is then

$$\begin{aligned} AV_{RD^{(a)}}\left(\hat{t}^{(a)}_{yc\pi^*}\right) &= V_{RD^{(a)}}\left(\sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta h}} \sum\nolimits_{s_h} R^{(a)}_{k|s} \left(\check{y}_k - \frac{\hat{t}^{(a)}_{y\theta h}}{\hat{t}^{(a)}_{\theta h}} \right) \right) \\ &= \sum_{h=1}^{H_s} \left(\frac{n_h}{\hat{t}^{(a)}_{\theta h}} \right)^2 \sum\nolimits_{s_h} \theta^{(a)}_{k|s} \left(1 - \theta^{(a)}_{k|s}\right) \left(\check{y}_k - \frac{\hat{t}^{(a)}_{y\theta h}}{\hat{t}^{(a)}_{\theta h}} \right)^2 \end{aligned} \tag{17}$$

which follows directly if we note that $V_{RD^{(a)}}(R^{(a)}_{k|s}) = \theta^{(a)}_{k|s}\left(1 - \theta^{(a)}_{k|s}\right)$ and $Cov_{RD^{(a)}}\left(R^{(a)}_{k|s}, R^{(a)}_{l|s}\right) = 0$. This approximation will work well if the expected sizes of the response homogeneity groups are large. The conditional nonresponse variance will be small if the "residuals" $\check{y}_k - \hat{t}^{(a)}_{y\theta h}/\hat{t}^{(a)}_{\theta h}$ are small, which they will be if $\check{y}_k$ is approximately constant within groups. Another possibility for small conditional nonresponse variance is if $\hat{t}^{(a)}_{\theta h}$ are close to n_h for all h, i.e. if the overall response propensity is high in all groups. This means that efforts that actually increase the response probabilities, will reduce the conditional nonresponse variance. Also, for the purpose of reducing the bias, it was shown in Tångdahl (2004) that the RHGs should be chosen so that the variability of $\breve{y}$ is small within groups. From (17), we see that this also helps reduce the conditional nonresponse variance. Small variability of $\breve{y}$ within RHGs can be aimed at through clever grouping of the sample into RHGs or by using a sampling design so that π_k is proportional to y_k.

5.2 The regression based RHG estimators

General case

The full response estimator corresponding to $\hat{t}^{(a)}_{ycreg}$ is

$$\hat{t}_{yreg} = \textstyle\sum_U \hat{y}_{1k} + \sum_s \frac{y_k - \hat{y}_{1k}}{\pi_k}$$

with predictions

$$\hat{y}_{1k} = \mathbf{x}'_{1k}\hat{\mathbf{B}}_{1s} = \mathbf{x}'_{1k}\left(\textstyle\sum_s \frac{\mathbf{x}_{1k}\mathbf{x}'_{1k}}{\sigma^2_{1k}\pi_k}\right)^{-1}\sum_s \frac{\mathbf{x}_{1k}y_k}{\sigma^2_{1k}\pi_k}$$

Approximate expressions for the four variance components of $V(\hat{t}^{(a)}_{ycreg})$ are:

$$V_1 \approx AV_p(\hat{t}_{yreg}) = \textstyle\sum\sum_U \Delta_{kl}\check{E}_{1k}\check{E}_{1k} \tag{18}$$

where $E_{1k} = y_k - \mathbf{x}'_{1k}\mathbf{B}_1$ with $\mathbf{B}_1 = \left(\sum_U \frac{\mathbf{x}_{1k}\mathbf{x}'_{1k}}{\sigma^2_{1k}}\right)^{-1}\sum_U \frac{\mathbf{x}_{1k}y_k}{\sigma^2_{1k}}$

$$\begin{aligned} V_2 &= V_p\left(E_{RD^{(a)}}(\hat{t}^{(a)}_{ycreg} - \hat{t}_{yreg}|s)\right) \\ &\approx V_p\left[\sum_{h=1}^{H_s}(n_h-1)\frac{1}{\bar{\theta}^{(a)}_{s_h}}\left(S_{\theta^{(a)}\check{e}_s s_h} - S_{\theta^{(a)}\check{\mathbf{x}} s_h}\left(\hat{\mathbf{B}}^{(a)}_\theta - \hat{\mathbf{B}}_s\right)\right)\right] \end{aligned} \tag{19}$$

with $\hat{\mathbf{B}}_s = \left(\sum_s \frac{\mathbf{x}_k\mathbf{x}'_k}{\sigma^2_k\pi_k}\right)^{-1}\sum_s \frac{\mathbf{x}_k y_k}{\sigma^2_k\pi_k}$,
$S_{\theta^{(a)}\check{e}_s s_h} = \frac{1}{n_h - 1}\left(\sum_{s_h}\theta^{(a)}_{k|s}\check{e}_{ks} - \frac{1}{n_h}\left(\sum_{s_h}\theta^{(a)}_{k|s}\right)\left(\sum_{s_h}\check{e}_{ks}\right)\right)$ and
$S_{\theta^{(a)}\check{\mathbf{x}}_s s_h}$ is analogously defined.

$$V_3 = E_p\left(V_{RD^{(a)}}(\hat{t}^{(a)}_{ycreg})\right)$$
$$\approx E_p\left[\sum_{h=1}^{H_s}\left(\frac{n_h}{\hat{t}^{(a)}_{\theta_h}}\right)^2 \sum\nolimits_{s_h}\theta^{(a)}_{k|s}(1-\theta^{(a)}_{k|s})\left(\check{f}^{(a)}_{k\theta;g} - \frac{\sum_{s_h}\theta^{(a)}_{k|s}\check{f}^{(a)}_{k\theta;g}}{\hat{t}^{(a)}_{\theta_h}}\right)^2\right] \quad (20)$$

with $\check{f}^{(a)}_{k\theta;g} = (g^{(a)}_{1k\theta}-1)\check{e}^{(a)}_{1k\theta} + g^{(a)}_{2k\theta}\check{e}^{(a)}_{k\theta}$ and where $g_{1k\theta}$ and $g_{2k\theta}$ are given by (24) and (25) respectively.

$$V_4 = 2Cov_p(E_{RD^{(a)}}(\hat{t}^{(a)}_{ycreg} - \hat{t}_{yreg}|s), \hat{t}^{(a)}_{yreg})$$
$$\approx\ 2Cov_p\left(\sum_{h=1}^{H_s}(n_h-1)\frac{1}{\bar{\theta}^{(a)}_{s_h}}\left(S_{\theta\check{e}_s s_h} - S_{\theta^{(a)}\check{\mathbf{x}}s_h}\left(\hat{\mathbf{B}}^{(a)}_{\theta} - \hat{\mathbf{B}}_s\right)\right), \hat{t}^{(a)}_{yreg}\right) \quad (21)$$

The components (19) and (21) follow from Tångdahl (2004). In the derivation of (20), we apply the same method of linearization as in the simple RHG case in section 5.1.

We note that the general regression based RHG estimator given by (2) is a nonlinear function of estimated totals and can be written as

$$\hat{t}^{(a)}_{ycreg} = \hat{t}^{(a)}_{yc\pi^*} + \left(\mathbf{t}_{\mathbf{x}_1U} - \hat{\mathbf{t}}_{\mathbf{x}_1\pi}\right)'\hat{\mathbf{B}}^{(a)}_{1r} + \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\pi^*}\right)'\hat{\mathbf{B}}^{(a)}_{r}$$
$$= \sum_{h=1}^{H_s}\frac{n_h}{\hat{t}^{(a)}_{\theta r_h}}\hat{t}^{(a)}_{yr_h} + \left(\mathbf{t}_{\mathbf{x}_1U} - \hat{\mathbf{t}}_{\mathbf{x}_1\pi}\right)'\left(\sum_{h=1}^{H_s}\frac{n_h}{\hat{t}^{(a)}_{\theta r_h}}\hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1r_h}\right)^{-1}\sum_{h=1}^{H_s}\frac{n_h}{\hat{t}^{(a)}_{\theta r_h}}\hat{\mathbf{t}}^{(a)}_{\mathbf{x}_1yr_h}$$
$$+ \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\pi^*}\right)'\left(\sum_{h=1}^{H_s}\frac{n_h}{\hat{t}^{(a)}_{\theta r_h}}\hat{\mathbf{T}}^{(a)}_{\mathbf{xx}r_h}\right)^{-1}\sum_{h=1}^{H_s}\frac{n_h}{\hat{t}^{(a)}_{\theta r_h}}\hat{\mathbf{t}}^{(a)}_{\mathbf{x}yr_h} \quad (22)$$

where $\mathbf{t}_{\mathbf{x}_1U} = \sum_U \mathbf{x}_{1k}$, $\hat{\mathbf{t}}_{\mathbf{x}_1\pi} = \sum_s \check{\mathbf{x}}_{1k}$ and $\hat{\mathbf{t}}_{\mathbf{x}\pi} = \sum_s \check{\mathbf{x}}_k$, and

$$\hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1r_h} = \sum\nolimits_{r^{(a)}_h}\frac{\mathbf{x}_{1k}\mathbf{x}'_{1k}}{\sigma^2_{1k}\pi_k}$$
$$\hat{\mathbf{t}}^{(a)}_{\mathbf{x}_1yr_h} = \sum\nolimits_{r^{(a)}_h}\frac{\mathbf{x}_{1k}y_k}{\sigma^2_{1k}\pi_k}$$
$$\hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\pi^*} = \sum_{h=1}^{H_s} n_h\frac{\hat{\mathbf{t}}^{(a)}_{\mathbf{x}r_h}}{\hat{t}^{(a)}_{\theta r_h}}$$

and analogously for $\hat{\mathbf{T}}^{(a)}_{\mathbf{xx}r_h}$ and $\hat{\mathbf{t}}^{(a)}_{\mathbf{x}yr_h}$.

By Taylor linearization, it is shown in Appendix A.1 that a linear large sample approximation to $\hat{t}^{(a)}_{ycreg}$ is given by

$$\begin{aligned}\hat{t}^{(a)}_{ycreg} \approx\ & \hat{t}^{(a)}_{yc\theta} + \left(\mathbf{t}_{\mathbf{x}_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1\pi}\right)' \hat{\mathbf{B}}^{(a)}_{1\theta} + \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\theta}\right)' \hat{\mathbf{B}}^{(a)}_{\theta} \\ & + \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta h}} \textstyle\sum_{s_h} R^{(a)}_{k|s} \left((g^{(a)}_{1k\theta} - 1)\check{e}^{(a)}_{1k\theta} + g^{(a)}_{2k\theta}\check{e}^{(a)}_{k\theta}\right) \\ & - \sum_{h=1}^{H_s} \frac{n_h}{(\hat{t}^{(a)}_{\theta h})^2} \textstyle\sum_{s_h} \theta^{(a)}_{k|s} \left((g^{(a)}_{1k\theta} - 1)\check{e}^{(a)}_{1k\theta} + g^{(a)}_{2k\theta}\check{e}^{(a)}_{k\theta}\right) \textstyle\sum_{s_h} R^{(a)}_{k|s}\end{aligned} \tag{23}$$

where $e^{(a)}_{1k\theta} = y_k - \mathbf{x}'_{1k}\hat{\mathbf{B}}^{(a)}_{1\theta}$, $e^{(a)}_{k\theta} = y_k - \mathbf{x}'_k\hat{\mathbf{B}}^{(a)}_{\theta}$,

$$g^{(a)}_{1k\theta} = 1 + \left(\mathbf{t}_{\mathbf{x}_1} - \hat{\mathbf{t}}_{\mathbf{x}_1\pi}\right)' \left(\sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta_h}} \hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta_h} \right)^{-1} \frac{\mathbf{x}_{1k}}{\sigma^2_{1k}} \tag{24}$$

and

$$g^{(a)}_{2k\theta} = 1 + \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\theta}\right)' \left(\sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta_h}} \hat{\mathbf{T}}^{(a)}_{\mathbf{xx}\theta_h} \right)^{-1} \frac{\mathbf{x}_{k}}{\sigma^2_{k}} \tag{25}$$

Furthermore,

$$E_{RD^{(a)}}\left(\hat{\mathbf{B}}^{(a)}_{1r}\right) \approx \hat{\mathbf{B}}^{(a)}_{1\theta} = \left(\sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta_h}} \hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta_h} \right)^{-1} \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta_h}} \hat{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y\theta_h} \tag{26}$$

and

$$E_{RD^{(a)}}\left(\hat{\mathbf{B}}^{(a)}_{r}\right) \approx \hat{\mathbf{B}}^{(a)}_{\theta} = \left(\sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta_h}} \hat{\mathbf{T}}^{(a)}_{\mathbf{xx}\theta_h} \right)^{-1} \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta_h}} \hat{\mathbf{t}}^{(a)}_{\mathbf{x}y\theta_h} \tag{27}$$

If we let $\check{f}^{(a)}_{k\theta;g} = (g^{(a)}_{1k\theta} - 1)\check{e}^{(a)}_{1k\theta} + g^{(a)}_{2k\theta}\check{e}^{(a)}_{k\theta}$, we can write (23) as

$$\begin{aligned}\hat{t}^{(a)}_{ycreg} \approx\ & \hat{t}^{(a)}_{yc\theta} + \left(\mathbf{t}_{\mathbf{x}_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1\pi}\right)' \hat{\mathbf{B}}^{(a)}_{1\theta} + \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\theta}\right)' \hat{\mathbf{B}}^{(a)}_{\theta} \\ & + \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta h}} \textstyle\sum_{s_h} R^{(a)}_{k|s} \left(\check{f}^{(a)}_{k\theta;g} - \dfrac{\sum_{s_h} \theta^{(a)}_{k|s} \check{f}^{(a)}_{k\theta;g}}{\hat{t}^{(a)}_{\theta h}} \right)\end{aligned}$$

We can then easily use the analogy with the linearized expression for the simple RHG estimator (16). From (23), and by applying the same reasoning as in the simple RHG case, we thus get an approximate conditional nonresponse variance of $\hat{t}^{(a)}_{ycreg}$ as

$$AV_{RD^{(a)}}(\hat{t}^{(a)}_{ycreg}) = \sum_{h=1}^{H_s} \left(\frac{n_h}{\hat{t}^{(a)}_{\theta h}} \right)^2 \sum\nolimits_{s_h} \theta^{(a)}_{k|s}(1-\theta^{(a)}_{k|s}) \left(\check{f}^{(a)}_{k\theta;g} - \frac{\sum_{s_h} \theta^{(a)}_{k|s} \check{f}^{(a)}_{k\theta;g}}{\hat{t}^{(a)}_{\theta h}} \right)^2 \tag{28}$$

The conditional nonresponse variance will thus be small if $\check{f}^{(a)}_{k\theta;g} - \frac{\sum_{s_h} \theta^{(a)}_{k|s} \check{f}^{(a)}_{k\theta;g}}{\hat{t}^{(a)}_{\theta h}}$ are small, i.e. if $\check{f}^{(a)}_{k\theta;g}$ are approximately constant within groups, or if the overall response propensities are high in all groups.

Special case 1

The corresponding estimator with full response is $\hat{t}_{yreg}$, the same as in the general case. Thus, the first variance component, the sampling variance, is given by (18). Moreover, V_2 and V_4 are given by (19) and (21) respectively, but with $\mathbf{x}$ instead of $\mathbf{x}_1$.

The third variance component is

$$\begin{aligned} V_3 =& E_p\left(V_{RD^{(a)}}(\hat{t}^{(a)}_{ycreg1}|s) \right) \\ \approx& E_p\left(\sum_{h=1}^{H_s} \left(\frac{n_h}{\hat{t}^{(a)}_{\theta_h}} \right)^2 \sum\nolimits_{s_h} \theta^{(a)}_{k|s}(1-\theta^{(a)}_{k|s}) \left(g^{*(a)}_{1k\theta} \check{e}^{(a)}_{1k\theta} - \frac{\sum_{s_h} \theta^{(a)}_{k|s} g^{*(a)}_{1k\theta} \check{e}^{(a)}_{1k\theta}}{\hat{t}^{(a)}_{\theta_h}} \right)^2 \right) \end{aligned}$$

The conditional nonresponse variance in special case 1 follows easily from (28) if we note that $e^{(a)}_{k\theta} = e^{(a)}_{1k\theta}$, so that $(g^{(a)}_{1k\theta} - 1)\check{e}^{(a)}_{1k\theta} + g^{(a)}_{2k\theta}\check{e}^{(a)}_{k\theta} = (g^{(a)}_{1k\theta} + g^{(a)}_{2k\theta} - 1)\check{e}^{(a)}_{1k\theta}$. If we define

$$g^{*(a)}_{1k\theta} = (g^{(a)}_{1k\theta} + g^{(a)}_{2k\theta} - 1) = 1 + \left(\mathbf{t}_{x_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1 c\theta} \right)' \left(\sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta h}} \hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta h} \right)^{-1} \frac{\mathbf{x}_{1k}}{\sigma^2_{1k}} \tag{29}$$

we can write the approximate conditional nonresponse variance as

$$\begin{aligned} & AV_{RD^{(a)}}(\hat{t}^{(a)}_{ycreg1}) = \\ & \quad = \sum_{h=1}^{H_s} \left(\frac{n_h}{\hat{t}^{(a)}_{\theta_h}} \right)^2 \sum\nolimits_{s_h} \theta^{(a)}_{k|s}(1-\theta^{(a)}_{k|s}) \left(g^{*(a)}_{1k\theta} \check{e}^{(a)}_{1k\theta} - \frac{\sum_{s_h} \theta^{(a)}_{k|s} g^{*(a)}_{1k\theta} \check{e}^{(a)}_{1k\theta}}{\hat{t}^{(a)}_{\theta_h}} \right)^2 \end{aligned} \tag{30}$$

Special case 2

In this case, the full response estimator is $\hat{t}_{y\pi} = \sum_s \check{y}_k$, so the sampling variance is given by (12). The variance components V_2 and V_4 are given by (19) and (21) respectively, while

$$\begin{aligned} V_3 &= E_p\left(V_{RD^{(a)}}(\hat{t}^{(a)}_{ycreg2}|s)\right) \\ &\approx E_p\left(\sum_{h=1}^{H_s}\left(\frac{n_h}{\hat{t}^{(a)}_{\theta_h}}\right)^2 \sum\nolimits_{s_h}\theta^{(a)}_{k|s}(1-\theta^{(a)}_{k|s})\left(g^{(a)}_{2k\theta}\check{e}^{(a)}_{k\theta} - \frac{\sum_{s_h}\theta^{(a)}_{k|s}g^{(a)}_{2k\theta}\check{e}^{(a)}_{k\theta}}{\hat{t}^{(a)}_{\theta_h}}\right)^2\right) \end{aligned} \tag{31}$$

From (28), we obtain the conditional nonresponse variance of $\hat{t}^{(a)}_{ycreg2}$. Since the predictions $\hat{y}_{1k}$ drop out of (4), we have no residuals $e^{(a)}_{1kr}$. The linear large sample approximation to (4) then follows directly as

$$\begin{aligned} \hat{t}^{(a)}_{ycreg2} \approx \hat{t}^{(a)}_{yc\theta} + \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\theta}\right)' \hat{\mathbf{B}}^{(a)}_{\theta} + \sum_{h=1}^{H_s}\frac{n_h}{\hat{t}^{(a)}_{\theta h}}\sum\nolimits_{s_h} R^{(a)}_{k|s}g^{(a)}_{2k\theta}\check{e}^{(a)}_{k\theta} \\ - \sum_{h=1}^{H_s}\frac{n_h}{(\hat{t}^{(a)}_{\theta h})^2}\sum\nolimits_{s_h}\theta^{(a)}_{k|s}g^{(a)}_{2k\theta}\check{e}^{(a)}_{k\theta}\sum\nolimits_{s_h}R^{(a)}_{k|s} \end{aligned} \tag{32}$$

Again, we can use the same reasoning as in the case of the simple RHG estimator. Thus, the approximate conditional nonresponse variance of $\hat{t}^{(a)}_{ycreg2}$ is

$$\begin{aligned} &AV_{RD^{(a)}}(\hat{t}^{(a)}_{ycreg2}) = \\ &= \sum_{h=1}^{H_s}\left(\frac{n_h}{\hat{t}^{(a)}_{\theta h}}\right)^2 \sum\nolimits_{s_h}\theta^{(a)}_{k|s}(1-\theta^{(a)}_{k|s})\left(g^{(a)}_{2k\theta}\check{e}^{(a)}_{k\theta} - \frac{\sum_{s_h}\theta^{(a)}_{k|s}g^{(a)}_{2k\theta}\check{e}^{(a)}_{k\theta}}{\hat{t}^{(a)}_{\theta h}}\right)^2 \end{aligned} \tag{33}$$

5.3 The calibration estimators

Special case 1

Since the full response estimator corresponding to $\hat{t}^{(a)}_{ycal1}$ is $\hat{t}_{yreg}$, the same as for the regression based RHG estimator in special case 1, the sampling variance of $\hat{t}^{(a)}_{ycal1}$ is given by (18). Furthermore, we have

$$V_2 = V_p\left(E_{RD^{(a)}}(\hat{t}^{(a)}_{ycal1} - \hat{t}_{yreg}|s)\right)$$
$$\approx V_p\left(\sum\nolimits_s\left(1-\theta^{(a)}_{k|s}\right)\left(\frac{\mathbf{x}'_{1k}}{\pi_k}\left(\tilde{\mathbf{B}}^{(a)}_{1\theta} - \hat{\mathbf{B}}_{1s}\right) - \frac{e_{1ks}}{\pi_k}\right)\right) \tag{34}$$

with $e_{1ks} = y_k - \mathbf{x}'_{1k}\hat{\mathbf{B}}_{1s}$

$$V_3 = E_p(V_{RD^{(a)}}(\hat{t}^{(a)}_{ycal1})) \approx \sum\nolimits_s \theta^{(a)}_{k|s}(1-\theta^{(a)}_{k|s})\left(v^{(a)}_{1k\theta}\check{e}^{(a)}_{1k\theta}\right)^2 \tag{35}$$

where $e^{(a)}_{1k\theta} = y_k - \mathbf{x}'_{1k}\tilde{\mathbf{B}}^{(a)}_{1\theta}$ and $v^{(a)}_{1k\theta}$ is given by (44)

$$V_4 = 2Cov_p(E_{RD^{(a)}}\left(\hat{t}^{(a)}_{ycal1} - \hat{t}_{yreg}|s), \hat{t}^{(a)}_{yreg}\right)$$
$$\approx 2Cov_p\left(\sum\nolimits_s\left(1-\theta^{(a)}_{k|s}\right)\left(\frac{\mathbf{x}'_{1k}}{\pi_k}\left(\tilde{\mathbf{B}}^{(a)}_{1\theta} - \hat{\mathbf{B}}_{1s}\right) - \frac{e_{1ks}}{\pi_k}\right), \hat{t}^{(a)}_{yreg}\right) \tag{36}$$

where $\tilde{\mathbf{B}}^{(a)}_{1\theta}$ is given by (39).

The components V_2 and V_4 follow directly from Tångdahl (2004). To derive V_3, we again use Taylor expansion. The calibration estimators can be expressed as nonlinear functions of estimated totals. In special case 1, we have

$$\begin{aligned}\hat{t}^{(a)}_{ycal1} &= \sum\nolimits_{r^{(a)}}\check{y}_k + \left(\sum\nolimits_U \mathbf{x}_{1k} + \sum\nolimits_{r^{(a)}}\check{\mathbf{x}}_{1k}\right)'\tilde{\mathbf{B}}^{(a)}_{1r}\\ &= \tilde{t}^{(a)}_{yr} + \left(\mathbf{t}_{\mathbf{x}_{1U}} - \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 r}\right)'\left(\tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r}\right)^{-1}\tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 yr}\\ &= f^{(a)}(\tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r}, \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 yr}, \tilde{t}^{(a)}_{yr}, \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 r})\end{aligned}$$

where $\tilde{t}^{(a)}_{yr} = \sum_{r^{(a)}}\check{y}_k$, $\tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 r} = \sum_{r^{(a)}}\check{\mathbf{x}}_{1k}$,

$$\tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r} = \sum\nolimits_{r^{(a)}}\frac{c_k\mathbf{x}_{1k}\mathbf{x}'_{1k}}{\pi_k} \qquad \text{and} \qquad \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 yr} = \sum\nolimits_{r^{(a)}}\frac{c_k\mathbf{x}_{1k}y_k}{\pi_k}.$$

The expected values under $RD^{(a)}$, given s, are $\tilde{t}^{(a)}_{y\theta} = \sum_s\theta^{(a)}_{k|s}\check{y}_k$, $\tilde{\mathbf{t}}_{\mathbf{x}_{1\theta}} = \sum_s\theta^{(a)}_{k|s}\check{\mathbf{x}}_{1k}$,

$$\tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta} = \sum\nolimits_s\frac{\theta^{(a)}_{k|s}c_k\mathbf{x}_{1k}\mathbf{x}'_{1k}}{\pi_k} \tag{37}$$

$$\tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y\theta} = \sum_{r^{(a)}} \frac{\theta^{(a)}_{k|s} c_k \mathbf{x}_{1k} y_k}{\pi_k} \tag{38}$$

and

$$E_{RD^{(a)}}\left(\tilde{\mathbf{B}}^{(a)}_{1r}\right) \approx \tilde{\mathbf{B}}^{(a)}_{1\theta} = \left(\tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta}\right)^{-1} \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y\theta} \tag{39}$$

A large sample linear approximation to $\hat{t}^{(a)}_{ycal1}$, using Taylor expansion, is then

$$\hat{t}^{(a)}_{ycal1} \approx \mathbf{t}'_{\mathbf{x}_1 U}\tilde{\mathbf{B}}^{(a)}_{1\theta} + \sum_s R^{(a)}_{k|s} v^{(a)}_{1k\theta} \check{e}^{(a)}_{1k\theta} \tag{40}$$

with $e^{(a)}_{1k\theta} = y_k - \mathbf{x}_{1k}\tilde{\mathbf{B}}^{(a)}_{1\theta}$ and

$$v^{(a)}_{1k\theta} = 1 + \left(\mathbf{t}_{\mathbf{x}_{1U}} - \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1\theta}\right)' \left(\tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta}\right)^{-1} c_k \mathbf{x}_{1k} \tag{41}$$

The details of the derivation are given in Appendix A.2. From the linear approximation, we get the approximate conditional nonresponse variance of $\hat{t}^{(a)}_{ycal1}$ as

$$AV_{RD^{(a)}}(\hat{t}^{(a)}_{ycal1}) = \sum_s \theta^{(a)}_{k|s}(1-\theta^{(a)}_{k|s})(v^{(a)}_{1k\theta}\check{e}^{(a)}_{1k\theta})^2 \tag{42}$$

Special case 2

The sampling variance of $\hat{t}^{(a)}_{ycal2}$ is given by (12) while V_2 and V_4 are given by (34) and (36) respectively, but with $\mathbf{x}$ instead of $\mathbf{x}_1$. Furthermore,

$$V_3 = E_p(V_{RD^{(a)}}(\hat{t}^{(a)}_{ycal2})) \approx E_p\left(\sum_s \theta^{(a)}_{k|s}(1-\theta^{(a)}_{k|s})(v^{(a)}_{2k\theta}\check{e}^{(a)}_{k\theta})^2\right) \tag{43}$$

with

$$v^{(a)}_{2k\theta} = 1 + \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}\theta}\right)' \left(\tilde{\mathbf{T}}^{(a)}_{\mathbf{x}\mathbf{x}\theta}\right)^{-1} c_k \mathbf{x}_k \tag{44}$$

When auxiliary information is available only for $k \in s$, we note that a linear approximation to $\hat{t}^{(a)}_{ycal2}$ is given by

$$\hat{t}^{(a)}_{ycal2} \approx \hat{\mathbf{t}}'_{\mathbf{x}\pi}\tilde{\mathbf{B}}^{(a)}_{\theta} + \sum_s R^{(a)}_{k|s} v^{(a)}_{2k\theta} \check{e}^{(a)}_{k\theta} \tag{45}$$

with $e^{(a)}_{k\theta} = y_k - \mathbf{x}'_k\tilde{\mathbf{B}}^{(a)}_{\theta}$. Hence, the conditional nonresponse variance of $\hat{t}^{(a)}_{cal2}$ is

$$AV_{RD^{(a)}}(\hat{t}^{(a)}_{cal2}) = \sum_s \theta^{(a)}_{k|s}(1-\theta^{(a)}_{k|s})(v^{(a)}_{2k\theta}\check{e}^{(a)}_{k\theta})^2$$

6 Measures of error

In a survey with nonresponse, where estimators may be greatly biased, the variance is not a satisfactory measure of the error. Instead, measures taking both variance and bias into account should be used. We believe that in the presence of nonresponse, the risk of bias overwhelms any requirement of precision. Focus lies on accuracy rather than precision and this should be reflected in the measure of error. The *mean square error*, which is the sum of variance and squared bias, is a common measure of quality in estimates. Other possibilities are the bias ratio and the mean absolute error. In practical applications, regardless of which error measure we choose, the problem of estimating the bias still remains. Unless special efforts are made to estimate the bias, one has to settle for an estimate of the (sampling) variance.

7 Concluding remarks

In a survey with nonresponse, decisions must be made on how much resources should be spent on data collection in order to reduce the nonresponse rate. To be able to make an informed decision, the effects of nonresponse rate reduction on estimator properties must be studied. In this paper, as part of a project on balancing nonresponse rate reduction efforts and costs of data collection, we present expressions for the variance and its components for some common estimators under nonresponse. This is done for an arbitrary point of time during the data collection process under a general, assumed true, response distribution. The estimators that are used are chosen to represent two widely used classes of estimators that incorporate auxiliary information. The variances are, even for the simplest of the estimators, complex functions of the response probabilities.

The variance, and how it changes during the data collection process, becomes important when the nonresponse rate reduction efforts are evaluated and a possible truncation of field efforts is considered, since the variance may be prohibitively large when the nonresponse rate is high.

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*, 413–428.

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*, 125–148.

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

Särndal, C. E., B. Swensson, and J. Wretman (1992). *Model Assisted Survey Sampling*. New York: Springer-Verlag.

Tångdahl, S. (2004). Nonresponse bias for some common estimators and its change over time in the data collection process. ESI Working Paper Series 2004:13, Örebro University.

A Derivations

A.1 Taylor linearization of $\hat{t}^{(a)}_{ycreg}$

The regression based RHG estimator in the general case is

$$\hat{t}^{(a)}_{ycreg} = \sum_{h=1}^{H_s} \frac{n_h}{m_h^{(a)}} \sum\nolimits_{r_h^{(a)}} \check{y}_k + \left(\mathbf{t}_{\mathbf{x}_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1 \pi}\right)' \hat{\mathbf{B}}^{(a)}_{1r} + \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\pi^*}\right)' \hat{\mathbf{B}}^{(a)}_{r}$$

with $\hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\pi^*} = \sum_{h=1}^{H_s} n_h \frac{\hat{\mathbf{t}}^{(a)}_{\mathbf{x}r_h}}{\hat{t}^{(a)}_{\theta r_h}}$, where $\hat{\mathbf{t}}^{(a)}_{\mathbf{x}r_h} = \sum_{r_h^{(a)}} \check{\mathbf{x}}_k$ and $\hat{t}^{(a)}_{\theta r_h} = m_h^{(a)} = \sum_{s_h} R^{(a)}_{k|s}$.

The estimator can be written as

$$\begin{aligned}
\hat{t}^{(a)}_{ycreg} =& \hat{t}^{(a)}_{yc\pi^*} + \left(\mathbf{t}_{\mathbf{x}_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1 \pi}\right)' \left(\sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta r_h}} \hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r_h}\right)^{-1} \left(\sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta r_h}} \hat{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y r_h}\right) \\
&+ \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta r_h}} \hat{\mathbf{t}}^{(a)}_{\mathbf{x} r_h}\right)' \left(\sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta r_h}} \hat{\mathbf{T}}^{(a)}_{\mathbf{x}\mathbf{x} r_h}\right)^{-1} \left(\sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta r_h}} \hat{\mathbf{t}}^{(a)}_{\mathbf{x} y r_h}\right) \\
=& f^{(a)}(\hat{t}^{(a)}_{\theta r_h}, \hat{\mathbf{t}}^{(a)}_{\mathbf{x} r_h}, \hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r_h}, \hat{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y r_h}, \hat{\mathbf{T}}^{(a)}_{\mathbf{x}\mathbf{x} r_h}, \hat{\mathbf{t}}^{(a)}_{\mathbf{x} y r_h}, \hat{t}^{(a)}_{y r_h}; h = 1, \ldots, H_s)
\end{aligned} \tag{A.1}$$

where $\hat{t}^{(a)}_{yr_h} = \sum_{r_h^{(a)}} \check{y}_k$ and we define the totals

$$\hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r_h} = \sum\nolimits_{r_h^{(a)}} \frac{\mathbf{x}_{1k}\mathbf{x}'_{1k}}{\sigma^2_{1k}\pi_k} \quad \text{and} \quad \hat{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y r_h} = \sum\nolimits_{r_h^{(a)}} \frac{\mathbf{x}_{1k} y_k}{\sigma^2_{1k}\pi_k}$$

with typical elements

$$\hat{t}^{(a)}_{j_1 j'_1 r_h} = \sum\nolimits_{r_h^{(a)}} \frac{x_{j_1 k} x_{j'_1 k}}{\sigma^{(a)}_{1k}\pi_k} \quad \text{and} \quad \hat{t}^{(a)}_{j_1 y r_h} = \sum\nolimits_{r_h^{(a)}} \frac{x_{j_1 k} y_k}{\sigma^{(a)}_{1k}\pi_k}.$$

The totals $\hat{\mathbf{T}}^{(a)}_{\mathbf{x}\mathbf{x} r_h}$ and $\hat{\mathbf{t}}^{(a)}_{\mathbf{x} y r_h}$ are analogously defined.

Thus, conditional on s, $\hat{t}^{(a)}_{ycreg}$ is a nonlinear function of the estimated totals $\hat{t}^{(a)}_{\theta r_h}, \hat{\mathbf{t}}^{(a)}_{\mathbf{x} r_h}, \hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r_h}, \hat{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y r_h}, \hat{\mathbf{T}}^{(a)}_{\mathbf{x}\mathbf{x} r_h}, \hat{\mathbf{t}}^{(a)}_{\mathbf{x} y r_h}, \hat{t}^{(a)}_{y r_h}; h = 1, \ldots, H_s$. Using first order Taylor expansion, this estimator can be approximated by a linear pseudoestimator through

$$\hat{t} \approx \hat{t}_0 = t + \sum_{q=1}^{Q} a_q \left(\hat{t}_q - t_q\right) \tag{A.2}$$

where

$$a_q = \left.\frac{\delta f}{\delta \hat{t}_q}\right|_{(\hat{t}_1,\ldots,\hat{t}_Q)=E(\hat{t}_1,\ldots,\hat{t}_Q)}$$

In this case, the expectation will be taken under $RD^{(a)}$, given s. We will need the following partial derivatives:

$$\begin{aligned}
\frac{\delta f^{(a)}}{\delta \hat{t}^{(a)}_{\theta r_h}} = & -\frac{n_h}{(\hat{t}^{(a)}_{\theta r_h})^2}\hat{t}^{(a)}_{yr_h} + \frac{n_h}{(\hat{t}^{(a)}_{\theta r_h})^2}\hat{\mathbf{t}}^{(a)}_{\mathbf{x}r_h}\hat{\mathbf{B}}^{(a)}_r \\
& - \left(\mathbf{t}_{\mathbf{x}_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1\pi}\right)' \left(\sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta r_h}} \hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r_h}\right)^{-1} \frac{n_h}{(\hat{t}^{(a)}_{\theta r_h})^2}\left[\hat{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y r_h} - \hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r_h}\hat{\mathbf{B}}^{(a)}_{1r}\right] \\
& - \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\pi^*}\right)' \left(\sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta r_h}} \hat{\mathbf{T}}^{(a)}_{\mathbf{x}\mathbf{x} r_h}\right)^{-1} \frac{n_h}{(\hat{t}^{(a)}_{\theta r_h})^2}\left[\hat{\mathbf{t}}^{(a)}_{\mathbf{x} y r_h} - \hat{\mathbf{T}}^{(a)}_{\mathbf{x}\mathbf{x} r_h}\hat{\mathbf{B}}^{(a)}_{r}\right]
\end{aligned}$$

$$\frac{\delta f^{(a)}}{\delta \hat{t}^{(a)}_{j_1 j_1' r_h}} = -\left(\mathbf{t}_{\mathbf{x}_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1\pi}\right)' \left(\sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta r_h}} \hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r_h}\right)^{-1} \hat{\mathbf{\Lambda}}^{(a)}_{j_1 j_1' r_h}\hat{\mathbf{B}}^{(a)}_{1r}$$

$$\frac{\delta f^{(a)}}{\delta \hat{t}^{(a)}_{j j' r_h}} = -\left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\pi^*}\right)' \left(\sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta r_h}} \hat{\mathbf{T}}^{(a)}_{\mathbf{x}\mathbf{x} r_h}\right)^{-1} \hat{\mathbf{\Lambda}}^{(a)}_{j j' r_h}\hat{\mathbf{B}}^{(a)}_{r}$$

$$\frac{\delta f^{(a)}}{\delta \hat{t}^{(a)}_{j_1 y r_h}} = \left(\mathbf{t}_{\mathbf{x}_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1\pi}\right)' \left(\sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta r_h}} \hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r_h}\right)^{-1} \hat{\boldsymbol{\lambda}}^{(a)}_{j_1 r_h}$$

$$\frac{\delta f^{(a)}}{\delta \hat{t}^{(a)}_{j y r_h}} = \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\pi^*}\right)' \left(\sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta r_h}} \hat{\mathbf{T}}^{(a)}_{\mathbf{x}\mathbf{x} r_h}\right)^{-1} \hat{\boldsymbol{\lambda}}^{(a)}_{j r_h}$$

$$\frac{\delta f^{(a)}}{\delta \hat{t}^{(a)}_{j r_h}} = -\hat{\boldsymbol{\lambda}}^{(a)'}_{j r_h}\hat{\mathbf{B}}^{(a)}_{r}$$

$$\frac{\delta f^{(a)}}{\delta \hat{t}^{(a)}_{y r_h}} = \frac{n_h}{\hat{t}^{(a)}_{\theta r_h}}$$

where $\hat{\mathbf{\Lambda}}_{j_1 j_1' r_h}$ is a $J_1 \times J_1$ matrix with the value $n_h/\hat{t}^{(a)}_{\theta r_h}$ in positions (j_1, j_1') and (j_1', j_1) and the value zero elsewhere, $\hat{\mathbf{\Lambda}}_{jj' r_h}$ is a $J \times J$ matrix with the value $n_h/\hat{t}^{(a)}_{\theta r_h}$ in positions (j, j') and (j', j) and the value zero elsewhere. The

J_1-vector $\hat{\boldsymbol{\lambda}}_{j_1 r_h}$ and the J-vector $\hat{\boldsymbol{\lambda}}_{j r_h}$ have the value $n_h/\hat{t}^{(a)}_{\theta r_h}$ in positions j_1 and j, respectively, and zeros elsewhere.

Let $\hat{t}^{(a)}_{y\theta h} = \sum_{s_h} \theta^{(a)}_{k|s} \breve{y}_k$, $\hat{t}^{(a)}_{\theta h} = \sum_{s_h} \theta^{(a)}_{k|s}$ and

$$\hat{\mathbf{t}}^{(a)}_{\mathbf{x}_1\theta h} = \sum\nolimits_{s_h} \theta^{(a)}_{k|s} \breve{\mathbf{x}}_{1k}, \quad \hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta h} = \sum\nolimits_{s_h} \frac{\theta^{(a)}_{k|s} \mathbf{x}_{1k}\mathbf{x}'_{1k}}{\sigma^2_{1k}\pi_k},$$
$$\hat{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y\theta h} = \sum\nolimits_{s_h} \frac{\theta^{(a)}_{k|s} \mathbf{x}_{1k} y_k}{\sigma^2_{1k}\pi_k}, \quad \hat{\mathbf{T}}^{(a)}_{\mathbf{x}\mathbf{x}\theta h} = \sum\nolimits_{s_h} \frac{\theta^{(a)}_{k|s} \mathbf{x}_{k}\mathbf{x}'_{k}}{\sigma^2_{k}\pi_k} \text{ and}$$
$$\hat{\mathbf{t}}^{(a)}_{\mathbf{x} y\theta h} = \sum\nolimits_{s_h} \frac{\theta^{(a)}_{k|s} \mathbf{x}_{k} y_k}{\sigma^2_{k}\pi_k}$$

be the respective expected values, under $RD^{(a)}$ given s, of the totals in (A.1). Also, for notational convenience, let $\hat{t}^{(a)}_{yc\theta} = \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta h}} \hat{t}^{(a)}_{y\theta h}$, $\hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\theta} = \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta h}} \hat{\mathbf{t}}^{(a)}_{\mathbf{x}\theta h}$, $\hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta} = \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta h}} \hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta h}$ and $\hat{\mathbf{T}}^{(a)}_{\mathbf{x}\mathbf{x}\theta} = \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta h}} \hat{\mathbf{T}}^{(a)}_{\mathbf{x}\mathbf{x}\theta h}$. Furthermore, let

$$E_{RD^{(a)}}\left(\hat{\mathbf{B}}^{(a)}_{1r}\right) \approx \hat{\mathbf{B}}^{(a)}_{1\theta} = \left(\sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta_h}} \hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta_h}\right)^{-1} \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta_h}} \hat{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y\theta_h} \tag{A.3}$$

and

$$E_{RD^{(a)}}\left(\hat{\mathbf{B}}^{(a)}_{r}\right) \approx \hat{\mathbf{B}}^{(a)}_{\theta} = \left(\sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta_h}} \hat{\mathbf{T}}^{(a)}_{\mathbf{x}\mathbf{x}\theta_h}\right)^{-1} \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta_h}} \hat{\mathbf{t}}^{(a)}_{\mathbf{x} y\theta_h} \tag{A.4}$$

Evaluating the partial derivatives at the expected value point

$$(\hat{t}^{(a)}_{\theta h}, \hat{\mathbf{t}}^{(a)}_{\mathbf{x}\theta h}, \hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta h}, \hat{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y\theta h}, \hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta h}, \hat{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y\theta h}, \hat{t}^{(a)}_{y\theta h}; h = 1, \ldots, H_s)$$

under $RD^{(a)}$ given s, and inserting into (A.2), we obtain

$$
\begin{aligned}
\hat{t}_{ycreg}^{(a)} \approx\ & \hat{t}_{yc\theta}^{(a)} + \left(\mathbf{t}_{\mathbf{x}_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1 \pi}\right)' \hat{\mathbf{B}}_{1\theta}^{(a)} + \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}_{\mathbf{x}c\theta}^{(a)}\right)' \hat{\mathbf{B}}_{\theta}^{(a)} \\
& - \sum_{h=1}^{H_s} \frac{n_h}{(\hat{t}_{\theta h}^{(a)})^2} \hat{t}_{y\theta h}^{(a)} (\hat{t}_{\theta r_h}^{(a)} - \hat{t}_{\theta h}^{(a)}) \\
& - \left(\mathbf{t}_{\mathbf{x}_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1 \pi}\right)' \left(\hat{\mathbf{T}}_{\mathbf{x}_1 \mathbf{x}_1 \theta}^{(a)}\right)^{-1} \sum_{h=1}^{H_s} \frac{n_h}{(\hat{t}_{\theta h}^{(a)})^2} \left[\hat{\mathbf{t}}_{\mathbf{x}_1 y\theta h}^{(a)} - \hat{\mathbf{T}}_{\mathbf{x}_1 \mathbf{x}_1 \theta h}^{(a)} \hat{\mathbf{B}}_{1\theta}^{(a)}\right] (\hat{t}_{\theta r_h}^{(a)} - \hat{t}_{\theta h}^{(a)}) \\
& - \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}_{\mathbf{x}c\theta}^{(a)}\right)' \left(\hat{\mathbf{T}}_{\mathbf{x}\mathbf{x}\theta}^{(a)}\right)^{-1} \sum_{h=1}^{H_s} \frac{n_h}{(\hat{t}_{\theta h}^{(a)})^2} \left[\hat{\mathbf{t}}_{\mathbf{x}y\theta h}^{(a)} - \hat{\mathbf{T}}_{\mathbf{x}\mathbf{x}\theta h}^{(a)} \hat{\mathbf{B}}_{\theta}^{(a)}\right] (\hat{t}_{\theta r_h}^{(a)} - \hat{t}_{\theta h}^{(a)}) \\
& + \sum_{h=1}^{H_s} \frac{n_h}{(\hat{t}_{\theta h}^{(a)})^2} \hat{\mathbf{t}}_{\mathbf{x}\theta h}^{(a)\prime} \hat{\mathbf{B}}_{\theta}^{(a)} (\hat{t}_{\theta r_h}^{(a)} - \hat{t}_{\theta h}^{(a)}) \\
& - \left(\mathbf{t}_{\mathbf{x}_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1 \pi}\right)' \left(\hat{\mathbf{T}}_{\mathbf{x}_1 \mathbf{x}_1 \theta}^{(a)}\right)^{-1} \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}_{\theta h}^{(a)}} \left(\hat{\mathbf{T}}_{\mathbf{x}_1 \mathbf{x}_1 r_h}^{(a)} - \hat{\mathbf{T}}_{\mathbf{x}_1 \mathbf{x}_1 \theta h}^{(a)}\right) \hat{\mathbf{B}}_{1\theta}^{(a)} \\
& - \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}_{\mathbf{x}c\theta}^{(a)}\right)' \left(\hat{\mathbf{T}}_{\mathbf{x}\mathbf{x}\theta}^{(a)}\right)^{-1} \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}_{\theta h}^{(a)}} \left(\hat{\mathbf{T}}_{\mathbf{x}\mathbf{x}r_h}^{(a)} - \hat{\mathbf{T}}_{\mathbf{x}\mathbf{x}\theta h}^{(a)}\right) \hat{\mathbf{B}}_{\theta}^{(a)} \\
& + \left(\mathbf{t}_{\mathbf{x}_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1 \pi}\right)' \left(\hat{\mathbf{T}}_{\mathbf{x}_1 \mathbf{x}_1 \theta}^{(a)}\right)^{-1} \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}_{\theta h}^{(a)}} \left(\hat{\mathbf{t}}_{\mathbf{x}_1 y r_h}^{(a)} - \hat{\mathbf{t}}_{\mathbf{x}_1 y\theta h}^{(a)}\right) \\
& + \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}_{\mathbf{x}c\theta}^{(a)}\right)' \left(\hat{\mathbf{T}}_{\mathbf{x}\mathbf{x}\theta}^{(a)}\right)^{-1} \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}_{\theta h}^{(a)}} \left(\hat{\mathbf{t}}_{\mathbf{x}y r_h}^{(a)} - \hat{\mathbf{t}}_{\mathbf{x}y\theta h}^{(a)}\right) \\
& - \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}_{\theta h}^{(a)}} \left(\hat{\mathbf{t}}_{\mathbf{x}r_h}^{(a)} - \hat{\mathbf{t}}_{\mathbf{x}\theta h}^{(a)}\right)' \hat{\mathbf{B}}_{\theta}^{(a)} + \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}_{\theta h}^{(a)}} \left(\hat{t}_{y r_h}^{(a)} - \hat{t}_{y\theta h}^{(a)}\right)
\end{aligned}
$$

With some simplifications we then have

$$\hat{t}^{(a)}_{ycreg} \approx \hat{t}^{(a)}_{yc\theta} + \left(\mathbf{t}_{\mathbf{x}_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1\pi}\right)' \hat{\mathbf{B}}^{(a)}_{1\theta} + \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\theta}\right)' \hat{\mathbf{B}}^{(a)}_{\theta}$$
$$- \left(\mathbf{t}_{\mathbf{x}_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1\pi}\right)' \left(\hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta}\right)^{-1} \sum_{h=1}^{H_s} \frac{n_h}{(\hat{t}^{(a)}_{\theta h})^2} \left(\hat{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y\theta h} - \hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta h}\hat{\mathbf{B}}^{(a)}_{1\theta}\right) \hat{t}^{(a)}_{\theta r_h}$$
$$- \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\theta}\right)' \left(\hat{\mathbf{T}}^{(a)}_{\mathbf{x}\mathbf{x}\theta}\right)^{-1} \sum_{h=1}^{H_s} \frac{n_h}{(\hat{t}^{(a)}_{\theta h})^2} \left(\hat{\mathbf{t}}^{(a)}_{\mathbf{x}y\theta h} - \hat{\mathbf{T}}^{(a)}_{\mathbf{x}\mathbf{x}\theta h}\hat{\mathbf{B}}^{(a)}_{\theta}\right) \hat{t}^{(a)}_{\theta r_h}$$
$$+ \sum_{h=1}^{H_s} \frac{n_h}{(\hat{t}^{(a)}_{\theta h})^2} \hat{\mathbf{t}}^{(a)\prime}_{\mathbf{x}\theta h}\hat{\mathbf{B}}^{(a)}_{\theta}\hat{t}^{(a)}_{\theta r_h} - \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta h}} \hat{\mathbf{t}}^{(a)\prime}_{\mathbf{x}r_h}\hat{\mathbf{B}}^{(a)}_{\theta} - \sum_{h=1}^{H_s} \frac{n_h}{(\hat{t}^{(a)}_{\theta h})^2} \hat{t}^{(a)}_{y\theta h}\hat{t}^{(a)}_{\theta r_h}$$
$$+ \left(\mathbf{t}_{\mathbf{x}_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1\pi}\right)' \left(\hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta}\right)^{-1} \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta h}} \left(\hat{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y r_h} - \hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r_h}\hat{\mathbf{B}}^{(a)}_{1\theta}\right)$$
$$+ \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\theta}\right)' \left(\hat{\mathbf{T}}^{(a)}_{\mathbf{x}\mathbf{x}\theta}\right)^{-1} \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta h}} \left(\hat{\mathbf{t}}^{(a)}_{\mathbf{x}y r_h} - \hat{\mathbf{T}}^{(a)}_{\mathbf{x}\mathbf{x}r_h}\hat{\mathbf{B}}^{(a)}_{\theta}\right) + \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta h}} \hat{t}^{(a)}_{y r_h}$$

Rewriting this expression leads to

$$\hat{t}^{(a)}_{ycreg} \approx \hat{t}^{(a)}_{yc\theta} + \left(\mathbf{t}_{\mathbf{x}_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1\pi}\right)' \hat{\mathbf{B}}^{(a)}_{1\theta} + \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\theta}\right)' \hat{\mathbf{B}}^{(a)}_{\theta}$$
$$+ \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta h}} \textstyle\sum_{s_h} R^{(a)}_{k|s} \left(\left(\mathbf{t}_{\mathbf{x}_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1\pi}\right)' \left(\hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta}\right)^{-1} \mathbf{x}_{1k}\right) \check{e}^{(a)}_{1\theta k}$$
$$- \sum_{h=1}^{H_s} \left[\frac{n_h}{(\hat{t}^{(a)}_{\theta h})^2} \textstyle\sum_{s_h} \theta^{(a)}_{k|s} \left(\left(\mathbf{t}_{\mathbf{x}_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1\pi}\right)' \left(\hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta}\right)^{-1} \mathbf{x}_{1k}\right) \check{e}^{(a)}_{1\theta k}\right] \textstyle\sum_{s_h} R^{(a)}_{k|s}$$
$$+ \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta h}} \textstyle\sum_{s_h} R^{(a)}_{k|s} \left(1 + \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\theta}\right)' \left(\hat{\mathbf{T}}^{(a)}_{\mathbf{x}\mathbf{x}\theta}\right)^{-1} \mathbf{x}_{k}\right) \check{e}^{(a)}_{\theta k}$$
$$- \sum_{h=1}^{H_s} \left[\frac{n_h}{(\hat{t}^{(a)}_{\theta h})^2} \textstyle\sum_{s_h} \theta^{(a)}_{k|s} \left(1 + \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\theta}\right)' \left(\hat{\mathbf{T}}^{(a)}_{\mathbf{x}\mathbf{x}\theta}\right)^{-1} \mathbf{x}_{k}\right) \check{e}^{(a)}_{\theta k}\right] \textstyle\sum_{s_h} R^{(a)}_{k|s}$$

which can be further simplified into

$$\hat{t}^{(a)}_{ycreg} \approx \hat{t}^{(a)}_{yc\theta} + \left(\mathbf{t}_{\mathbf{x}_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1\pi}\right)' \hat{\mathbf{B}}^{(a)}_{1\theta} + \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\theta}\right)' \hat{\mathbf{B}}^{(a)}_{\theta}$$
$$+ \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta h}} \textstyle\sum_{s_h} R^{(a)}_{k|s} \left((g^{(a)}_{1k\theta} - 1)\check{e}^{(a)}_{1k\theta} + g^{(a)}_{2k\theta}\check{e}^{(a)}_{k\theta}\right)$$
$$- \sum_{h=1}^{H_s} \frac{n_h}{(\hat{t}^{(a)}_{\theta h})^2} \textstyle\sum_{s_h} \theta^{(a)}_{k|s} \left((g^{(a)}_{1k\theta} - 1)\check{e}^{(a)}_{1k\theta} + g^{(a)}_{2k\theta}\check{e}^{(a)}_{k\theta}\right) \textstyle\sum_{s_h} R^{(a)}_{k|s}$$

where $e^{(a)}_{1\theta k} = y_k - \mathbf{x}'_{1k}\hat{\mathbf{B}}^{(a)}_{1\theta}$, $e^{(a)}_{\theta k} = y_k - \mathbf{x}'_k\hat{\mathbf{B}}^{(a)}_{\theta}$ and

$$g^{(a)}_{1k\theta} = 1 + \left(\mathbf{t}_{\mathbf{x}_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1\pi}\right)' \left(\sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta h}} \hat{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta h}\right)^{-1} \frac{\mathbf{x}_{1k}}{\sigma^2_{1k}}$$
$$g^{(a)}_{2k\theta} = 1 + \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\theta}\right)' \left(\sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta h}} \hat{\mathbf{T}}^{(a)}_{\mathbf{x}\mathbf{x}\theta h}\right)^{-1} \frac{\mathbf{x}_{k}}{\sigma^2_{k}}$$

If we let $\check{f}^{(a)}_{k\theta;g} = (g^{(a)}_{1k\theta} - 1)\check{e}^{(a)}_{1k\theta} + g^{(a)}_{2k\theta}\check{e}^{(a)}_{k\theta}$, we can simplify further and write $\hat{t}^{(a)}_{ycreg}$ as

$$\hat{t}^{(a)}_{ycreg} \approx \hat{t}^{(a)}_{yc\theta} + \left(\mathbf{t}_{\mathbf{x}_1 U} - \hat{\mathbf{t}}_{\mathbf{x}_1\pi}\right)' \hat{\mathbf{B}}^{(a)}_{1\theta} + \left(\hat{\mathbf{t}}_{\mathbf{x}\pi} - \hat{\mathbf{t}}^{(a)}_{\mathbf{x}c\theta}\right)' \hat{\mathbf{B}}^{(a)}_{\theta}$$
$$+ \sum_{h=1}^{H_s} \frac{n_h}{\hat{t}^{(a)}_{\theta h}} \textstyle\sum_{s_h} R^{(a)}_{k|s} \left(\check{f}^{(a)}_{k\theta;g} - \dfrac{\sum_{s_h} \theta^{(a)}_{k|s} \check{f}^{(a)}_{k\theta;g}}{\hat{t}^{(a)}_{\theta h}}\right)$$

A.2 Taylor linearization of $\hat{t}^{(a)}_{ycal1}$

The calibration estimator is

$$\hat{t}^{(a)}_{ycal1} = \sum\nolimits_{r^{(a)}} \check{y}_k + (\mathbf{t}_{\mathbf{x}_1 U} - \sum\nolimits_{r^{(a)}} \check{\mathbf{x}}_{1k})' \, \tilde{\mathbf{B}}^{(a)}_{1r} \tag{A.5}$$

with

$$\begin{aligned} \tilde{\mathbf{B}}^{(a)}_{1r} &= \left(\sum\nolimits_{r^{(a)}} \frac{c_k \mathbf{x}_{1k} \mathbf{x}'_{1k}}{\pi_k} \right)^{-1} \sum\nolimits_{r^{(a)}} \frac{c_k \mathbf{x}_{1k} y_k}{\pi_k} \\ &= \left(\tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r} \right)^{-1} \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y r} \end{aligned}$$

where $\tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r}$ and $\tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y r}$ have the typical elements

$$\tilde{t}^{(a)}_{j_1 j'_1 r} = \sum\nolimits_{r^{(a)}} \frac{c_k x_{j_1 k} x'_{j_1 k}}{\pi_k} \qquad \text{and} \qquad \tilde{t}^{(a)}_{j_1 y r} = \sum\nolimits_{r^{(a)}} \frac{c_k x_{j_1 k} y_k}{\pi_k}.$$

Let $\tilde{t}^{(a)}_{yr} = \sum_{r^{(a)}} \check{y}_k$ and $\tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 r} = \sum_{r^{(a)}} \check{\mathbf{x}}_{1k}$. Then $\hat{t}^{(a)}_{ycal1}$ can be written as a nonlinear function of estimated totals:

$$\begin{aligned} \hat{t}^{(a)}_{ycal1} &= \tilde{t}^{(a)}_{yr} + \left(\mathbf{t}_{\mathbf{x}_1 U} - \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 r} \right)' \left(\tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r} \right)^{-1} \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y r} \\ &= f(\tilde{t}^{(a)}_{yr}, \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 r}, \tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r}, \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y r}) \end{aligned} \tag{A.6}$$

We will need the following partial derivatives:

$$\begin{aligned} \frac{\delta f}{\delta \tilde{t}^{(a)}_{j_1 j'_1 r}} &= -(\mathbf{t}_{\mathbf{x}_1 U} - \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 r})' (\tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r})^{-1} \tilde{\mathbf{\Lambda}}_{j_1 j'_1} \tilde{\mathbf{B}}^{(a)}_{1r} \\ \frac{\delta f}{\delta \tilde{t}^{(a)}_{j_1 y r}} &= (\mathbf{t}_{\mathbf{x}_1 U} - \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 r})' (\tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r})^{-1} \tilde{\boldsymbol{\lambda}}_{j_1} \\ \frac{\delta f}{\delta \tilde{t}^{(a)}_{j_1}} &= -\tilde{\boldsymbol{\lambda}}'_{j_1} (\tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r})^{-1} \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y r} = -\tilde{\boldsymbol{\lambda}}'_{j_1} \tilde{\mathbf{B}}^{(a)}_{1r} \\ \frac{\delta f}{\delta \tilde{t}^{(a)}_{yr}} &= 1 \end{aligned}$$

where $\tilde{\mathbf{\Lambda}}_{j_1 j'_1}$ is a $J_1 \times J_1$ matrix with the value 1 in positions (j_1, j'_1) and (j'_1, j_1) and the value 0 elsewhere and $\tilde{\boldsymbol{\lambda}}_j$ is a J_1-vector with the value 1 in position j_1 and zeros elsewhere.

The expected values, under $RD^{(a)}$, given s, of the totals in (A.6) are $\tilde{t}^{(a)}_{y\theta} = \sum_s \theta^{(a)}_{k|s} \check{y}_k$, $\tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1\theta} = \sum_s \theta^{(a)}_{k|s} \check{\mathbf{x}}_{1k}$, $\tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta} = \sum_s \dfrac{\theta^{(a)}_{k|s} c_k \mathbf{x}_{1k}\mathbf{x}'_{1k}}{\pi_k}$ and $\tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y\theta} = \sum_s \theta^{(a)}_{k|s} \mathbf{x}_{1k} y_k$ respectively.

Also, let $E_{RD^{(a)}}(\tilde{\mathbf{B}}^{(a)}_{1r}) \approx \tilde{\mathbf{B}}^{(a)}_{1\theta} = (\tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta})^{-1} \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y\theta}$. Evaluating the partial derivatives at the expected value point and inserting into (A.2) leads to

$$
\begin{aligned}
\hat{t}^{(a)}_{ycal1} &\approx \tilde{t}^{(a)}_{y\theta} + (\mathbf{t}_{\mathbf{x}_1 U} - \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1\theta})' \tilde{\mathbf{B}}^{(a)}_{1\theta} + \textstyle\sum_s R^{(a)}_{k|s} \check{y}_k - \sum_s \theta^{(a)}_{k|s} \check{y}_k + (\tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1\theta} - \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 r})' \tilde{\mathbf{B}}^{(a)}_{1\theta} \\
&\quad - (\mathbf{t}_{\mathbf{x}_1 U} - \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1\theta})' (\tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta})^{-1} \left(\tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r} - \tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta} \right) \tilde{\mathbf{B}}^{(a)}_{1\theta} \\
&\quad + (\mathbf{t}_{\mathbf{x}_1 U} - \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1\theta})' (\tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta})^{-1} (\tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 yr} - \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 y\theta}) \\
&= \mathbf{t}'_{\mathbf{x}_1 U} \tilde{\mathbf{B}}^{(a)}_{1\theta} + \textstyle\sum_s \dfrac{R^{(a)}_{k|s} e^{(a)}_{1\theta k}}{\pi_k} \\
&\quad + (\mathbf{t}_{\mathbf{x}_1 U} - \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1\theta})' (\tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta})^{-1} (\tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1 yr} - \tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1 r} \tilde{\mathbf{B}}^{(a)}_{1\theta}) \\
&= \mathbf{t}'_{\mathbf{x}_1 U} \tilde{\mathbf{B}}^{(a)}_{1\theta} + \textstyle\sum_s R^{(a)}_{k|s} v^{(a)}_{1k\theta} \check{e}^{(a)}_{1k\theta}
\end{aligned}
$$

with $e^{(a)}_{1k\theta} = y_k - \mathbf{x}'_{1k} \tilde{\mathbf{B}}^{(a)}_{1\theta}$ and $v^{(a)}_{1k\theta} = 1 + (\mathbf{t}_{\mathbf{x}_1 U} - \tilde{\mathbf{t}}^{(a)}_{\mathbf{x}_1\theta})' \left(\tilde{\mathbf{T}}^{(a)}_{\mathbf{x}_1\mathbf{x}_1\theta} \right)^{-1} c_k \mathbf{x}_{1k}$.