Catalogue no. 12-001-X
ISSN 1492-0921

Survey Methodology

How to decompose the non-response variance: A total survey error approach

by Keven Bosa, Serge Godbout, Fraser Mills and Frédéric Picard

Release date: December 20, 2018

Statistics Canada   Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at STATCAN.infostats-infostats.STATCAN@canada.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-514-283-9350

Depository Services Program

- Inquiries line 1-800-635-7943
- Fax line 1-800-565-7757

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "Standards of service to the public."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Standard table symbols

The following symbols are used in Statistics Canada publications:

. not available for any reference period
.. not available for a specific reference period
... not applicable
0 true zero or a value rounded to zero
0^s value rounded to 0 (zero) where there is a meaningful distinction between true zero and the value that was rounded
^{p} preliminary
^{r} revised
x suppressed to meet the confidentiality requirements of the *Statistics Act*
^{E} use with caution
F too unreliable to be published
* significantly different from reference category ($p < 0.05$)

An HTML version is also available.

Cette publication est aussi disponible en français.

Survey Methodology, December 2018
Vol. 44, No. 2, pp. 291-308
Statistics Canada, Catalogue No. 12-001-X

How to decompose the non-response variance: A total survey error approach

Keven Bosa, Serge Godbout, Fraser Mills and Frédéric Picard[1]

Abstract

When a linear imputation method is used to correct non-response based on certain assumptions, total variance can be assigned to non-responding units. Linear imputation is not as limited as it seems, given that the most common methods – ratio, donor, mean and auxiliary value imputation – are all linear imputation methods. We will discuss the inference framework and the unit-level decomposition of variance due to non-response. Simulation results will also be presented. This decomposition can be used to prioritize non-response follow-up or manual corrections, or simply to guide data analysis.

Key Words: Total variance; Adaptive design; Imputation.

1 Introduction

Total survey error is described by Biemer (2010) as the “accumulation of all errors that may arise in the design, collection, processing and analysis of survey data”. He classified survey error components into sampling error and nonsampling errors, such as, non-response, coverage, measurement and data processing errors. These errors may affect variance, bias, or both. The total survey error paradigm aims at maximizing survey quality by minimizing total survey error within prespecified resource constraints like budget, people, or time.

At Statistics Canada, the Corporate Business Architecture initiated the Integrated Business Statistics Program (IBSP) as the standardized platform for more than 140 economic surveys with the objective of achieving efficiency, enhancing quality and improving responsiveness. In particular, reducing collection costs while managing non-response error was identified as one of the program’s pillars. Consequently, an adaptive design where different units may receive different treatments became a keystone for this program. For more details on IBSP, see Statistics Canada (2015). Groves and Heeringa (2006) showed how paradata could be used to increase the response rate. Schouten, Calinescu and Luiten (2013) gave a general framework for an adaptive design and explained how the R-indicator could be used in this context.

A new survey process model called Rolling Estimates has been developed as an attempt to address the IBSP’s pillar mentioned above. The Rolling Estimates model is based on iterative processing and estimation cycles throughout the collection period. Basically, the idea of this model is to compute key estimates with their associated quality indicators at several specific times during the collection period. At the beginning, all units are assigned to the self-response survey treatment which means that the respondents are asked to complete the online questionnaire. Collection efforts like computer-assisted telephone interview non-response follow-ups are then performed on units contributing the most to the estimates where the quality is low based on the preliminary results of the Rolling Estimates. This can be viewed as an adaptive design since the treatments on the units depend on the quality of the estimates produced during the collection period. Most of the work

1. Keven Bosa, Serge Godbout, Fraser Mills and Frédéric Picard, Statistics Canada, 100 Tunney’s Pasture Driveway, Ottawa, Ontario, K1A 0T6. E-mail: serge.godbout@canada.ca.

regarding the development of the IBSP's adaptive design has been done since 2010. Godbout, Beaucage and Turmelle (2011), Turmelle, Godbout and Bosa (2012), Mills, Godbout, Bosa and Turmelle (2013) and Bosa and Godbout (2014) made use of this idea in the context of the IBSP adaptive design to minimize the number of follow-ups in order to reach a targeted quality in terms of coefficient of variation.

This paper revisits the work done so far for IBSP and presents an approach to decompose non-response variance into an item-level score for a given variable of interest within a domain. This item score is basically an attempt to estimate the contribution to the variance borrowed by a single unit. Units with a large score will contribute the most to reduce the variance and the coefficient of variation which is often used as a quality indicator in surveys. However, there are generally many important variables and domains in a survey. The proposed approach first computes, for a given unit, item-level scores for important variables and domains. Then, item scores can be combined into a single unit-level score in order to rank units. For example, the unit score can be a weighted sum or the maximum of its item scores. The most attractive use of the resulting unit-level score is to prioritize units, the ones with the highest scores, for the most expensive collection operations such as telephone follow-up, computer-assisted telephone interview or computer-assisted personal interview. This paper assumes total and partial non-response are both treated in the adaptive design, but treatments may be different depending on the type of non-response. For instance, telephone follow-ups could be made in the case of total non-response whereas questionnaires with partial non-response could be reviewed by analysts. This type of adaptive design generates strong interactions between collection operations, observed data and measured quality. Bosa and Godbout (2014) showed how this methodology was implemented in IBSP under the Rolling Estimates model.

Emphasis will be placed on the derivation of the item-level score throughout this paper. Therefore, the special case of only one variable of interest within a domain is studied. Also, only one imputation method is used to impute the variable of interest in the case of non-response so as to simplify the notation and to ease comprehension for the reader.

Section 2 describes the inference framework. In Section 3, the decomposition of the variance at the unit-level is expressed. In other words, the contribution of each nonresponding unit to the variance is computed. A simulation study was conducted to evaluate the proposed score. It is described in Section 4. Finally, Section 5 expresses some thoughts and conclusions.

2 Inference framework

Assume a sample s of size n is drawn from a population U of size N. Define the population total by

$$t_d = \sum_{k \in U} d_k y_k \tag{2.1}$$

for a variable, y, and a domain indicator, d_k, which takes the value $d_k = 1$ if unit k belongs to the domain d, and $d_k = 0$ otherwise. In the context of full response, t_d is estimated by $\hat{t}_d^{\,0} = \sum_{k \in s} d_k w_k y_k$ where w_k could be the sampling weight or a calibrated weight if calibration is performed. Because surveys are generally subject to non-response, both unit or item, a sample unit is classified into either a responding or a nonresponding unit with regard to the variable y at any given point during data collection. The subset s_r contains item-responding units whereas s_m contains item-nonresponding units. Note that s_r and s_m, respectively of size n_r and n_m, form a partition of the sample s, $P_s = \{s_r, s_m\}$, with $s_r \cup s_m = s$ and $s_r \cap s_m = \varnothing$.

The approach proposed in this paper assumes that imputation is used in case of non-response, which is the common approach in business surveys. Moreover, this approach can be considered for both item and unit non-response as long as imputation is used. However, since only one variable of interest y is considered here for simplicity, then no distinction is made if the y variable is imputed because of item or unit non-response. Also, the set s_r and s_m are not indexed by an item number for simplicity without loss of generality. However, the action following the calculation of a unit score might be different depending on whether the unit is responding or not.

2.1 Estimation under imputation

The framework requires linear imputation methods. In other words, the imputed value, y_k^*, can be written as a linear combination of the values reported by the other units. This linear combination is given by $y_k^* = \varphi_{0k} + \sum_{l \in s_r} \varphi_{lk} y_l$. The quantities, φ_{0k} and φ_{lk} do not depend on the values of variable of interest, y, but they may depend on s, s_r and auxiliary data from the nonrespondents available on the frame, registers or elsewhere. Linear imputation methods cover most methods used in practice like auxiliary value imputation (Beaumont, Haziza and Bocci, 2011) and linear regression imputation, as well as donor imputation, which is often used to impute categorical variables.

It is common practice to use several imputation methods, referred to as composite imputation, applied sequentially to the same variable. More than one linear imputation method can be used to impute nonresponding units. Section 2 of Beaumont and Bissonnette (2011) defines composite imputation in detail. Briefly, suppose that the set of nonrespondents is broken down into two or more groups and that a different imputation method is used within each group. For example, let $\mathbf{x}_k$ be the complete vector of auxiliary variables for unit k, and suppose regression imputation is used to impute the variable of interest. However, if, for some cases, $\mathbf{x}_k$ were incomplete, another imputation method, based on the available subset of $\mathbf{x}_k$, would be used. The approach presented in our paper can be generalized to include composite imputation as long as linear imputation methods are used. For simplicity of notation, the case of a single linear imputation method is presented.

The estimator of the domain total after imputation is given by

$$\hat{t}_d = \sum_{l \in s_r} w_l d_l y_l + \sum_{k \in s_m} w_k d_k y_k^* \tag{2.2}$$

where w_k is the sampling weight or a calibrated weight. The estimator presented in equation (2.2) can be rewritten as

$$\begin{aligned}
\hat{t}_d &= \sum_{l \in s_r} w_l d_l y_l + \sum_{k \in s_m} w_k d_k y_k^* \\
&= \sum_{l \in s_r} w_l d_l y_l + \sum_{k \in s_m} w_k d_k \left(\varphi_{0k} + \sum_{l \in s_r} \varphi_{lk} y_l \right) \\
&= \sum_{l \in s_r} w_l d_l y_l + \sum_{k \in s_m} w_k d_k \varphi_{0k} + \sum_{l \in s_r} y_l \sum_{k \in s_m} w_k d_k \varphi_{lk} \\
&= W_{0d} + \sum_{l \in s_r} w_l d_l y_l + \sum_{l \in s_r} y_l W_{dl} \\
&= W_{0d} + \sum_{l \in s_r} y_l \left(w_l d_l + W_{dl} \right).
\end{aligned}$$

The quantities W_{dl} and W_{0d} denote the compensatory weights (or adjustment weights) defined as

$$\begin{aligned} W_{dl} &= \sum_{k \in s_m} w_k d_k \varphi_{lk} \\ W_{0d} &= \sum_{k \in s_m} w_k d_k \varphi_{0k}. \end{aligned}$$

They represent the effect of the non-response in the domain, d, carried by the respondent unit, $l \in s_r$, with a reported value, y_l.

2.2 Variance estimation

Consider an imputation model, η, describing the relationship between variable y and the vector of observed auxiliary variables $\mathbf{x}^{\text{obs}}$. Let $E_\eta(.)$, $\text{Var}_\eta(.)$ and $\text{cov}_\eta(.)$ denote respectively the expectation, the variance, and the covariance with respect to the imputation model η. The imputation model is

$$\begin{aligned} E_\eta(y_k \mid \mathbf{X}^{\text{obs}}) &= \mu_k \\ V_\eta(y_k \mid \mathbf{X}^{\text{obs}}) &= \sigma_k^2 \\ \text{cov}_\eta(y_k, y_{k'} \mid \mathbf{X}^{\text{obs}}) &= 0 \end{aligned}$$

where $k, k' \in U$ and $k \neq k'$. The matrix $\mathbf{X}^{\text{obs}}$ contains all observed vectors $\mathbf{x}^{\text{obs}}$. The quantities μ_k and σ_k^2 can be estimated by $\hat{\mu}_k$ and $\hat{\sigma}_k^2$ respectively. We assume that these estimators are unbiased with respect to the imputation model η. These estimators will be useful later for estimating the total variance components and the unit decompositions of those components.

The total error of the estimator (2.2) can be expressed as

$$\hat{t}_d - t_d = \left(\hat{t}_d^0 - t_d\right) + \left(\hat{t}_d - \hat{t}_d^0\right), \tag{2.3}$$

where $\hat{t}_d^0$ is the estimator under complete response given by (2.1). The first term on the right-hand side of (2.3) is usually referred to as the sampling error and the second term is called the non-response error. As proposed in Särndal (1992) and in Beaumont and Bissonnette (2011), the mean square error of $\hat{t}_d$ using (2.3) can be decomposed in three components and is given by

$$\begin{aligned} E_{\eta pq}\left(\hat{t}_d - t_d\right)^2 = {} & E_\eta V_p\left(\hat{t}_d\right) + E_{pq} E_\eta\left[\left(\hat{t}_d - \hat{t}_d^0\right)^2 \mid s, s_r\right] \\ & + 2 E_{pq} E_\eta\left[\left(\hat{t}_d - \hat{t}_d^0\right)\left(\hat{t}_d^0 - t_d\right) \mid s, s_r\right], \end{aligned} \tag{2.4}$$

under imputation model, η, sampling design, p, and response mechanism, q. $E_{\eta pq}\left(\hat{t}_d - t_d\right)^2$ is approximately equivalent to the variance $V_{\eta pq}\left(\hat{t}_d - t_d\right)$ assuming that the overall bias is negligible. Thus, the equation (2.4) is equivalent to $V_{\eta pq}\left(\hat{t}_d - t_d\right) \equiv V_{\text{TOT}}\left(\hat{t}_d\right) = V_{\text{SAM}}\left(\hat{t}_d\right) + V_{\text{NR}}\left(\hat{t}_d\right) + V_{\text{MIX}}\left(\hat{t}_d\right)$, where:

- $V_{\text{SAM}}\left(\hat{t}_d\right) \equiv E_\eta V_p\left(\hat{t}_d\right)$ is the sampling variance;
- $V_{\text{NR}}\left(\hat{t}_d\right) \equiv E_{pq} E_\eta\left[\left(\hat{t}_d - \hat{t}_d^0\right)^2 \mid s, s_r\right]$ is the non-response variance;
- $V_{\text{MIX}}\left(\hat{t}_d\right) \equiv 2 E_{pq} E_\eta\left[\left(\hat{t}_d - \hat{t}_d^0\right)\left(\hat{t}_d^0 - t_d\right) \mid s, s_r\right]$ is the covariance between sampling and non-response error terms, also called the mixed variance component.

Beaumont and Bissonnette (2011) proposed the following estimators for $V_{\text{SAM}}\left(\hat{t}_d\right)$, $V_{\text{NR}}\left(\hat{t}_d\right)$ and $V_{\text{MIX}}\left(\hat{t}_d\right)$.

1. $\hat{V}_{\text{SAM}}\left(\hat{t}_d\right) = \hat{V}_{\text{ORD}}\left(\hat{t}_d\right) + \hat{V}_{\text{DIF}}\left(\hat{t}_d\right)$ where:
   - o $\hat{V}_{\text{ORD}}\left(\hat{t}_d\right)$ is the naive sampling variance estimator using the imputed values as though they were reported values.
   - o $\hat{V}_{\text{DIF}}\left(\hat{t}_d\right) = \sum_{k\in s_m}(1-\pi_k)\,w_k^2 d_k \hat{\sigma}_k^2$ is a correction to $\hat{V}_{\text{ORD}}\left(\hat{t}_d\right)$ in order to reduce the bias of $\hat{V}_{\text{ORD}}\left(\hat{t}_d\right)$, as proposed by Beaumont and Bocci (2009), since the variance component $\hat{V}_{\text{ORD}}\left(\hat{t}_d\right)$ relies on the use of imputed values, usually more homogeneous than the reported values.
2. $\hat{V}_{\text{NR}}\left(\hat{t}_d\right) = \sum_{l\in s_r} W_{dl}^2 \hat{\sigma}_l^2 + \sum_{k\in s_m} w_k^2 d_k \hat{\sigma}_k^2$ is the estimator of the non-response component of variance.
3. $\hat{V}_{\text{MIX}}\left(\hat{t}_d\right) = 2\sum_{l\in s_r} W_{dl}\,(w_l - 1)\,d_l \hat{\sigma}_l^2 - 2\sum_{k\in s_m} w_k\,(w_k - 1)\,d_k \hat{\sigma}_k^2$ is the estimator of the mixed variance component.

Under complete response, $s_m = \varnothing$, the compensation weights are $W_{dl} = 0$, and the variance components, $\hat{V}_{\text{DIF}}\left(\hat{t}_d\right), \hat{V}_{\text{NR}}\left(\hat{t}_d\right)$, and $\hat{V}_{\text{MIX}}\left(\hat{t}_d\right)$, are also equal to 0, leaving the total variance as $\hat{V}_{\text{TOT}}\left(\hat{t}_d\right) = \hat{V}_{\text{ORD}}\left(\hat{t}_d\right)$. Under a census, $s = U$, the variance components, $\hat{V}_{\text{DIF}}\left(\hat{t}_d\right), \hat{V}_{\text{ORD}}\left(\hat{t}_d\right)$, and $\hat{V}_{\text{MIX}}\left(\hat{t}_d\right)$, are equal to 0, leaving the total variance as $\hat{V}_{\text{TOT}}\left(\hat{t}_d\right) = \hat{V}_{\text{NR}}\left(\hat{t}_d\right)$.

2.3 Non-response bias

The reduction of non-response bias is always a desirable goal. It can be achieved through an adaptive design and/or through an appropriate method of dealing with missing values. Our framework assumes that the non-response bias is removed through imputation methods that use relevant auxiliary information. In practice, it is likely that imputation will only reduce non-response bias, not eliminate it. We may then wonder whether adaptive designs could be used to reduce further the bias. In the context of non-response weighting, Beaumont, Bocci and Haziza (2014) argued that auxiliary information used in an adaptive design to reduce non-response bias can also be used in non-response weighting to reduce the same amount of bias. Their argument can also be made in the context of imputation. This justifies our focus on variance reduction rather than bias reduction. We acknowledge that some bias may remain after imputation but ignore this bias because it may not be possible to reduce it further through an adaptive design without the availability of additional auxiliary information. However, it is possible to reduce the variance through an adaptive design.

3 Unit-level error decomposition of variance components

This section describes the approach used to evaluate the contribution of a given nonresponding unit, $\lambda \in s_m$, to the estimated total variance for the estimation of a total for a given variable.

The unit-level error decomposition, δ_λ, of the total variance for a given unit, λ, is defined as the difference between the estimated total variance, and the projected total variance, i.e., $\delta_\lambda\left(\hat{V}_{\text{TOT}}\left(\hat{t}_d\right)\right) \equiv \hat{V}_{\text{TOT}}\left(\hat{t}_d\right) - \hat{V}_{\text{TOT}}^{(\lambda)}\left(\hat{t}_d\right)$. The superscript (λ) is used to indicate projected quantities when unit λ is converted to a respondent. So, $\delta_\lambda\left(\hat{V}_{\text{TOT}}\left(\hat{t}_d\right)\right)$ can be seen as the expected gain, in terms of total variance, of converting a nonrespondent unit λ to a respondent.

In order to get $\delta_\lambda\left(\hat{V}_{\text{TOT}}\left(\hat{t}_d\right)\right)$, λ is moved from s_m to s_r, generating the new partition $P_s^{(\lambda)}$ of the sample from P_s where $P_s^{(\lambda)} = \left\{s_r^{(\lambda)}, s_m^{(\lambda)}\right\}$, $s_r^{(\lambda)} = s_r \cup \{\lambda\}$ and $s_m^{(\lambda)} = s_m \setminus \{\lambda\}$, as illustrated in Figure 3.1.

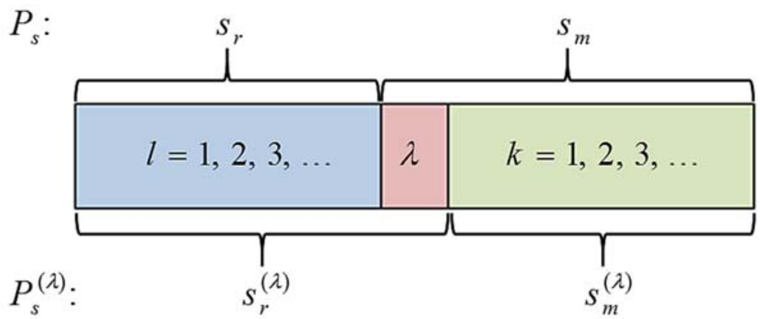

Figure 3.1 Sample partitions.

Some assumptions are necessary to decompose the variance components. It is recognized that these assumptions may not perfectly hold in reality. However, they can be used to generate accurate results, as shown in the simulation in Section 4. The required assumptions are:

1. Projected reported value: let $\lambda \in s_m$ be converted to a response and let $y_\lambda^{(\lambda)} = y_\lambda^*$.
2. Projected imputation parameters: $\forall k \in s_m$, $\hat{\mu}_k^{(\lambda)} = \hat{\mu}_k$ and $\hat{\sigma}_k^{(\lambda)} = \hat{\sigma}_k$.
3. Projected imputation relationship matrix: $\forall k \in s_m$ and $\forall l \in s_r$, $\varphi_{lk}^{(\lambda)} = 0$ if $l = \lambda$ or if $k = \lambda$ or $\varphi_{lk}^{(\lambda)} = \varphi_{lk}$ otherwise. Similarly, $\varphi_{0k}^{(\lambda)} = 0$ if $k = \lambda$ or $\varphi_{0k}^{(\lambda)} = \varphi_{0k}$ otherwise.

Assumption 1 implies that if a nonresponding unit, λ, would have been converted to a respondent, its reported value is equal to its imputed value. This is not true generally, but the imputed value is our best estimate. The expectation is that this imputed value is close enough to the reported value to estimate the error on the variance components. This assumption will have an impact when the sampling variance is decomposed.

Assumption 2 states that the estimated parameters of the imputation model would remain unchanged if λ were a respondent. In the case of a consistent imputation model parameter estimator, this assumption becomes more realistic when s_r is larger.

Finally, assumption 3 means that the imputation relationship between nonrespondents and respondents remains unchanged, except when unit λ is involved. In other words, the converted unit, λ, is no longer imputed from respondents, but will not be used to impute other nonresponding units. Figure 3.2 shows how assumption 3 is reflected in terms of the phi matrix.

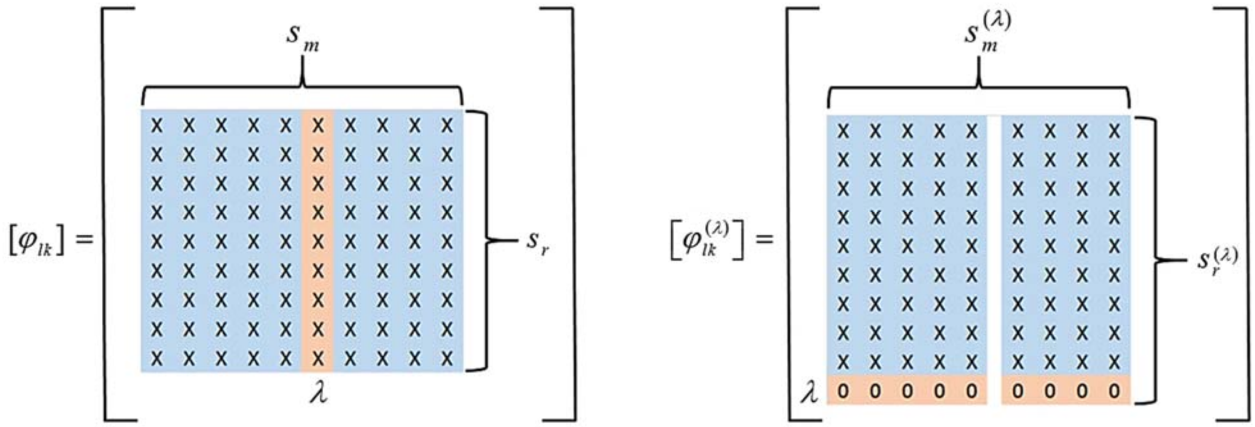

Figure 3.2 Initial and projected imputation relationship phi matrix.

Therefore, the compensation weight, $W_{dl}^{(\lambda)}$, of a responding unit, $\forall l \in s_r$, is projected as

$$\begin{aligned} W_{dl}^{(\lambda)} &= \sum_{k \in s_m^{(\lambda)}} w_k d_k \varphi_{lk}^{(\lambda)} \\ &= \sum_{k \in s_m} w_k d_k \varphi_{lk} - w_\lambda d_\lambda \varphi_{l\lambda} \\ &= W_{dl} - w_\lambda d_\lambda \varphi_{l\lambda}. \end{aligned} \tag{3.1}$$

The marginal weight from the converted unit λ is withdrawn from the original compensation weight, W_{dl}, to obtain the new $W_{dl}^{(\lambda)}$. Note that $W_{d\lambda}^{(\lambda)} = \sum_{k \in s_m^{(\lambda)}} w_k d_k \varphi_{\lambda k}^{(\lambda)} = 0$ because $\varphi_{\lambda k}^{(\lambda)} = 0$ under assumption 3. As mentioned above, it means that λ isn't used to impute nonrespondents.

In the next subsections, the unit-level error decomposition for unit λ is computed for the four variance components, as described in Section 2.3.

3.1 Unit-level error decomposition of the naive sampling variance

The quantity $\hat{V}_{\text{ORD}}(\hat{t}_d)$ depends on the y – values, the final weights and the first-order and second-order selection probabilities. The unit-level error decomposition of the naive sampling variance component $\hat{V}_{\text{ORD}}(\hat{t}_d)$ is trivial since the assumption that unit λ goes from s_m to s_r does not change weights and selection probabilities. Under assumption 1, the projected reported value $y_\lambda^{(\lambda)}$ is set to y_λ^* so that $\hat{V}_{\text{ORD}}^{(\lambda)}(\hat{t}_d) = \hat{V}_{\text{ORD}}(\hat{t}_d)$ when λ is converted to a responding unit. Consequently, the decomposition of $\hat{V}_{\text{ORD}}(\hat{t}_d)$ is given by

$$\delta_\lambda\left(\hat{V}_{\text{ORD}}(\hat{t}_d)\right) \equiv \hat{V}_{\text{ORD}}(\hat{t}_d) - \hat{V}_{\text{ORD}}^{(\lambda)}(\hat{t}_d) = 0. \tag{3.2}$$

This result is consistent with the idea that the naive sampling variance point estimate will likely change, but it is not expected to decrease with an extra responding unit.

3.2 Unit-level decomposition of the correction to the sampling variance component

The unit-level error decomposition for unit λ of the correction to the sampling variance component, $\hat{V}_{\text{DIF}}\left(\hat{t}_d\right)$, is given by

$$\begin{aligned}\delta_\lambda\left(\hat{V}_{\text{DIF}}\left(\hat{t}_d\right)\right) &\equiv \hat{V}_{\text{DIF}}\left(\hat{t}_d\right) - \hat{V}_{\text{DIF}}^{(\lambda)}\left(\hat{t}_d\right)\\ &= \sum_{k\in s_m}(1-\pi_k)\,d_k w_k^2 \hat{\sigma}_k^2 - \sum_{\lambda\in s_m^{(\lambda)}}(1-\pi_k)\,d_k w_k^2\left(\hat{\sigma}_k^{(\lambda)}\right)^2.\end{aligned}$$

Under assumption 2, $\hat{\sigma}_k^{(\lambda)} = \hat{\sigma}_k$, so that

$$\delta_\lambda\left(\hat{V}_{\text{DIF}}\left(\hat{t}_d\right)\right) = (1-\pi_\lambda)\,d_\lambda w_\lambda^2 \hat{\sigma}_\lambda^2. \tag{3.3}$$

The astute reader will notice that the actual sampling variance (not its estimation) should not be impacted by whether or not a unit is a respondent. However, we decided to include the impact of a unit on the sampling variance estimation in order to be coherent in the way we treat the three components $V_{\text{SAM}}\left(\hat{t}_d\right)$, $V_{\text{NR}}\left(\hat{t}_d\right)$ and $V_{\text{MIX}}\left(\hat{t}_d\right)$.

3.3 Unit-level decomposition of the non-response variance component

The unit-level error decomposition for unit λ of the non-response variance component $\hat{V}_{\text{NR}}\left(\hat{t}_d\right)$ is given by

$$\begin{aligned}\delta_\lambda\left(\hat{V}_{\text{NR}}\left(\hat{t}_d\right)\right) &\equiv \hat{V}_{\text{NR}}\left(\hat{t}_d\right) - \hat{V}_{\text{NR}}^{(\lambda)}\left(\hat{t}_d\right)\\ &= \left(\sum_{l\in s_r} W_{dl}^2 \hat{\sigma}_l^2 + \sum_{k\in s_m} w_k^2 d_k \hat{\sigma}_k^2\right) - \left(\sum_{l\in s_r^{(\lambda)}}\left(W_{dl}^{(\lambda)}\right)^2\left(\hat{\sigma}_l^{(\lambda)}\right)^2 + \sum_{k\in s_m^{(\lambda)}} w_k^2 d_k\left(\hat{\sigma}_k^{(\lambda)}\right)^2\right).\end{aligned}$$

Under assumptions 2 and 3, $\hat{\sigma}_k^{(\lambda)} = \hat{\sigma}_k$ and $W_{d\lambda}^{(\lambda)} = 0$. This can be rewritten as

$$\delta_\lambda\left(\hat{V}_{\text{NR}}\left(\hat{t}_d\right)\right) = \left(\sum_{l\in s_r} W_{dl}^2 \hat{\sigma}_l^2 - \sum_{l\in s_r}\left(W_{dl}^{(\lambda)}\right)^2 \hat{\sigma}_l^2\right) + w_\lambda^2 d_\lambda \hat{\sigma}_\lambda^2.$$

Using formula (3.1), this becomes

$$\begin{aligned}\delta_\lambda\left(\hat{V}_{\text{NR}}\left(\hat{t}_d\right)\right) &= \left(\sum_{l\in s_r} W_{dl}^2 \hat{\sigma}_l^2 - \sum_{l\in s_r}(W_{dl} - w_\lambda d_\lambda \varphi_{l\lambda})^2\,\hat{\sigma}_l^2\right) + w_\lambda^2 d_\lambda \hat{\sigma}_\lambda^2\\ &= \left(\sum_{l\in s_r} W_{dl}^2 \hat{\sigma}_l^2 - (W_{dl}^2 - 2W_{dl} w_\lambda d_\lambda \varphi_{l\lambda} + w_\lambda^2 d_\lambda \varphi_{l\lambda}^2)\,\hat{\sigma}_l^2\right) + w_\lambda^2 d_\lambda \hat{\sigma}_\lambda^2\\ &= \sum_{l\in s_r}(2W_{dl} w_\lambda d_\lambda \varphi_{l\lambda} - w_\lambda^2 d_\lambda \varphi_{l\lambda}^2)\,\hat{\sigma}_l^2 + w_\lambda^2 d_\lambda \hat{\sigma}_\lambda^2. \end{aligned} \tag{3.4}$$

3.4 Unit-level decomposition of the mixed variance component

Finally, the impact of unit λ on the variance component term, $\hat{V}_{\text{MIX}}\left(\hat{t}_d\right)$, is given by

$$\begin{aligned}\delta_\lambda\left(\hat{V}_{\text{MIX}}\left(\hat{t}_d\right)\right) &\equiv \hat{V}_{\text{MIX}}\left(\hat{t}_d\right) - \hat{V}_{\text{MIX}}^{(\lambda)}\left(\hat{t}_d\right) \\ &= \left(2\sum_{l\in s_r} W_{dl}\left(w_l - 1\right) d_l \hat{\sigma}_l^2 - 2\sum_{k\in s_m} w_k\left(w_k - 1\right) d_k \hat{\sigma}_k^2\right) \\ &\quad - \left(2\sum_{l\in s_r^{(\lambda)}} W_{dl}^{(\lambda)}\left(w_l - 1\right) d_l \left(\hat{\sigma}_l^{(\lambda)}\right)^2 - 2\sum_{k\in s_m^{(\lambda)}} w_k\left(w_k - 1\right) d_k \left(\hat{\sigma}_k^{(\lambda)}\right)^2\right).\end{aligned}$$

This equation can be rewritten as follows, under assumptions 2 and 3 and equation (3.1)

$$\begin{aligned}\delta_\lambda\left(\hat{V}_{\text{MIX}}\left(\hat{t}_d\right)\right) &= \left(2\sum_{l\in s_r} W_{dl}\left(w_l - 1\right) d_l \hat{\sigma}_l^2 - 2\sum_{k\in s_m} w_k\left(w_k - 1\right) d_k \hat{\sigma}_k^2\right) \\ &\quad - \left(2\sum_{l\in s_r} \left(W_{dl} - w_\lambda d_\lambda \varphi_{l\lambda}\right)\left(w_l - 1\right) d_l \hat{\sigma}_l^2 - 2\sum_{k\in s_m^{(\lambda)}} w_k\left(w_k - 1\right) d_k \hat{\sigma}_k^2\right) \\ &= 2\sum_{l\in s_r} w_\lambda d_\lambda \varphi_{l\lambda}\left(w_l - 1\right) d_l \hat{\sigma}_l^2 - 2w_\lambda\left(w_\lambda - 1\right) d_\lambda \hat{\sigma}_\lambda^2. \qquad (3.5)\end{aligned}$$

In Section 2.3, the estimation of the total variance, $\hat{V}_{\text{TOT}}\left(\hat{t}_d\right)$, has been defined as $\hat{V}_{\text{TOT}}\left(\hat{t}_d\right) = \hat{V}_{\text{ORD}}\left(\hat{t}_d\right) + \hat{V}_{\text{DIF}}\left(\hat{t}_d\right) + \hat{V}_{\text{NR}}\left(\hat{t}_d\right) + \hat{V}_{\text{MIX}}\left(\hat{t}_d\right)$. Similarly, the impact of unit λ on $\hat{V}_{\text{TOT}}\left(\hat{t}_d\right)$ is defined as

$$\delta_\lambda\left(\hat{V}_{\text{TOT}}\left(\hat{t}_d\right)\right) = \delta_\lambda\left(\hat{V}_{\text{ORD}}\left(\hat{t}_d\right)\right) + \delta_\lambda\left(\hat{V}_{\text{DIFF}}\left(\hat{t}_d\right)\right) + \delta_\lambda\left(\hat{V}_{\text{NR}}\left(\hat{t}_d\right)\right) + \delta_\lambda\left(\hat{V}_{\text{MIX}}\left(\hat{t}_d\right)\right),$$

where $\delta_\lambda\left(\hat{V}_{\text{ORD}}\left(\hat{t}_d\right)\right)$, $\delta_\lambda\left(\hat{V}_{\text{DIF}}\left(\hat{t}_d\right)\right)$, $\delta_\lambda\left(\hat{V}_{\text{NR}}\left(\hat{t}_d\right)\right)$, and $\delta_\lambda\left(\hat{V}_{\text{MIX}}\left(\hat{t}_d\right)\right)$ are respectively given by equations (3.2), (3.3), (3.4) and (3.5).

It can be observed (proofs are given in the appendix) that $\hat{V}_{\text{DIF}}\left(\hat{t}_d\right) = \sum_{k\in s_m} \delta_k\left(\hat{V}_{\text{DIF}}\left(\hat{t}_d\right)\right)$ and $\hat{V}_{\text{MIX}}\left(\hat{t}_d\right) = \sum_{k\in s_m} \delta_k\left(\hat{V}_{\text{MIX}}\left(\hat{t}_d\right)\right)$. However, this linear relation doesn't hold for $\hat{V}_{\text{NR}}\left(\hat{t}_d\right)$. This property is important to consider because, for $\hat{V}_{\text{DIF}}\left(\hat{t}_d\right)$ and $\hat{V}_{\text{MIX}}\left(\hat{t}_d\right)$, the sum of the unit-level errors on all nonresponding units, $k \in s_m$, is equal to the corresponding estimated variance component. In the case of non-response variance component, the sum of the errors is different than $\hat{V}_{\text{NR}}\left(\hat{t}_d\right)$. The difference is given by

$$\sum_{k\in s_m} \delta_k\left(\hat{V}_{\text{NR}}\left(\hat{t}_d\right)\right) - \hat{V}_{\text{NR}}\left(\hat{t}_d\right) = \sum_{l\in s_r}\left(\left(\sum_{k\in s_m} w_k d_k \varphi_{lk}\right)^2 - \sum_{k\in s_m} w_k^2 d_k \varphi_{lk}^2\right)\hat{\sigma}_l^2. \qquad (3.6)$$

This difference can be relatively small, especially in business surveys characterized with asymmetric data. This is the case when $\max_{k\in s_m}\left(w_k d_k \varphi_{lk}\right) \cong \sum_{k\in s_m} w_k d_k \varphi_{lk}$. This is in line with the results shown by Mills et al. (2013).

Overall, the total variance can be considered as approximately linear in terms of the unit-level errors, especially in the case of sample surveys where $\hat{V}_{\text{ORD}}\left(\hat{t}_d\right), \hat{V}_{\text{DIF}}\left(\hat{t}_d\right),$ and $\hat{V}_{\text{MIX}}\left(\hat{t}_d\right)$ are significant contributors to the total variance.

4 Simulation study

The sum of the item contributions is expected to be close enough to the estimated variance due to non-response. Simulations were conducted to assess the validity of the proposed score. The goal was then to evaluate if the proposed item contribution is a good approximation of the real contribution to the total variance of a given unit. In order to do so, the total contributions of a random subset of s_m were compared to the difference of the estimated variances where this subset is respectively considered as nonresponding units and responding units.

The following steps explain how simulations were performed.

1. A population was created, starting from an auxiliary variable x generated according to a gamma distribution with a mean of 48 and a variance of 768. The variable of interest y was created conditionally on x from a gamma distribution with a mean of $1.5x$ and a variance of $16x.$ These parameters are the same as the ones set by Beaumont and Bissonnette (2011).
2. A simple random sample s was selected from this population and an independent non-response subset s_m was generated using Bernoulli sampling.
   a. The nonresponding units from s_m were imputed using ratio imputation, where $y_k^* = x_k\left(\sum_{l \in s_r} y_l\right)\left(\sum_{l \in s_r} x_l\right)^{-1}$ and
   $$\hat{\sigma}_k^2 = x_k \frac{\sum_{l \in s_r}\left(y_l - y_l^*\right)^2}{\sum_{l \in s_r} x_l}.$$
   b. The population total $\hat{t},$ the variance components $\hat{V}_{\bullet}\left(\hat{t}\right),$ and the unit-level decompositions $\delta_k\left(\hat{V}_{\bullet}\left(\hat{t}\right)\right)$ were estimated, where the subscript $\bullet$ represents any of the variance components.
3. A subset, $\Lambda,$ of units, $\lambda,$ from $s_m,$ independently selected from a Bernoulli experiment, was moved from s_m to s_r to simulate non-response conversion. Therefore, we have a new partition, $P_s^{(\Lambda)},$ with $s_m^{(\Lambda)} = s_m \setminus \Lambda$ and $s_r^{(\Lambda)} = s_r \cup \Lambda.$
   a. The nonresponding units k from $s_m^{(\Lambda)}$ were re-imputed using a ratio model which is given by $y_k^{**} = x_k\left(\sum_{l \in s_r^{(\Lambda)}} y_l\right)\left(\sum_{l \in s_r^{(\Lambda)}} x_l\right)^{-1}$ and
   $$\hat{\sigma}_k^{*2} = x_k \frac{\sum_{l \in s_r^{(\Lambda)}}\left(y_l - y_l^{**}\right)^2}{\sum_{l \in s_r^{(\Lambda)}} x_l}.$$
   b. The population total, $\hat{t}^{(\Lambda)},$ and the variance components, $\hat{V}_{\bullet}\left(\hat{t}^{(\Lambda)}\right),$ were estimated.

4. The total of the unit-level decompositions, $\sum_{\lambda \in \Lambda} \delta_\lambda \left(\hat{V}_\bullet\left(\hat{t}\right)\right)$, for units λ from Λ was compared to the difference in the variance component estimates, $\hat{V}_\bullet\left(\hat{t}\right) - \hat{V}_\bullet\left(\hat{t}^{(\Lambda)}\right)$. The relative difference in the decomposition error, DRel, was calculated as

$$\text{DRel} = \frac{\left(\hat{V}_\bullet\left(\hat{t}^{(\Lambda)}\right) + \sum_{\lambda \in \Lambda} \delta_\lambda \left(\hat{V}_\bullet\left(\hat{t}\right)\right)\right) - \hat{V}_\bullet\left(\hat{t}\right)}{\hat{V}_\bullet\left(\hat{t}\right)}. \tag{4.1}$$

Steps 1 to 4 were independently repeated with different combinations of population size, sample size, response rate, and conversion rate as described in 4.1, 4.2 and 4.3.

4.1 Simulation scenario 1: Fixed parameters

In scenario 1, population size, sample size, response rate, and conversion rate were respectively set to 400, 100, 70%, and 33.3%, with 200 independent iterations. The results are shown in Figures 4.1 and 4.2.

Both Figures 4.1 and 4.2 show that the sum of the unit-level decomposition is a good predictor of the change in the non-response component estimates. The average relative difference in the variance estimates is low at 2.1%, but the standard error is large at 5.8%. Out of 200 relative differences, only 19 are not within the +/– 10% range but they are all above 10%. If a nonrespondent is converted to a respondent, we conclude that the non-response component of the variance will approximately be reduced by the measured contribution of this unit.

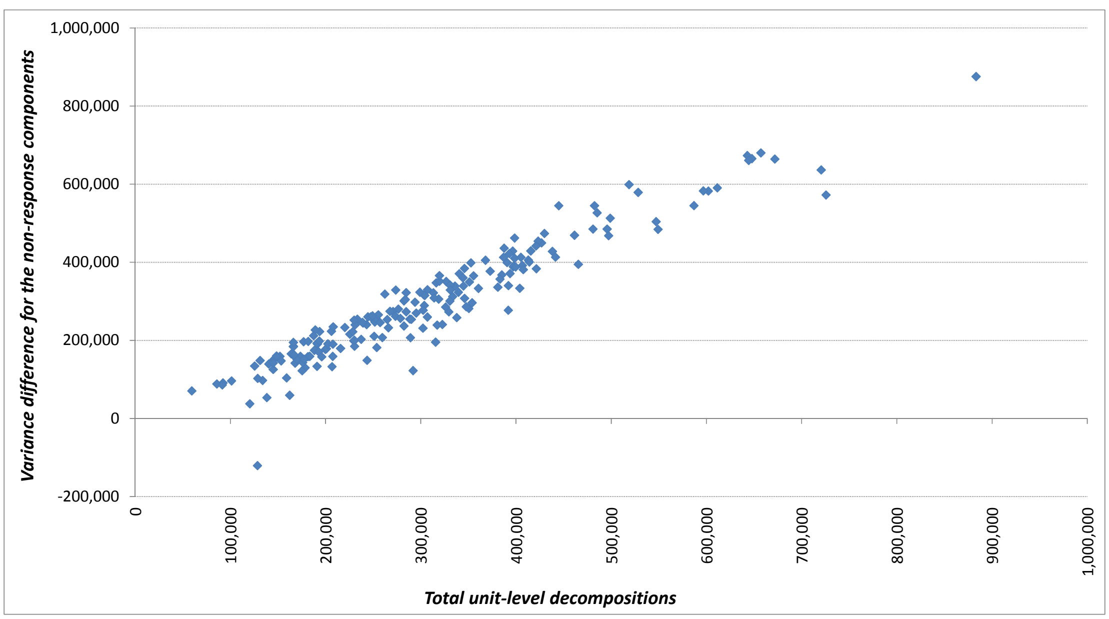

Figure 4.1 Variance difference for the non-response components versus total unit-level decompositions with fixed parameters.

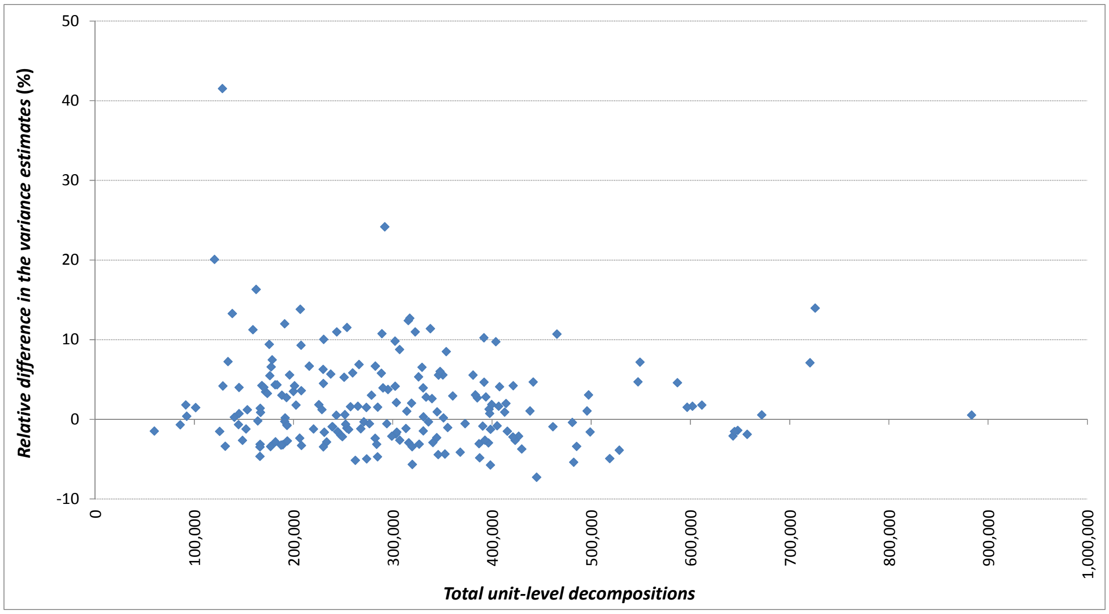

Figure 4.2 Relative difference in the variance estimates versus total unit-level decompositions with fixed parameters.

4.2 Simulation scenario 2: Varying population and sample sizes

In scenario 2, the population size ranged from 100 to 50,000, with sample rate, response rate, and conversion set to 20%, 70%, and 33.3% respectively. More iterations (40) were created for the smallest population $(N = 100)$, and less (10) for the largest $(N = 50{,}000)$, for operational considerations. The results are shown in Figures 4.3 and 4.4.

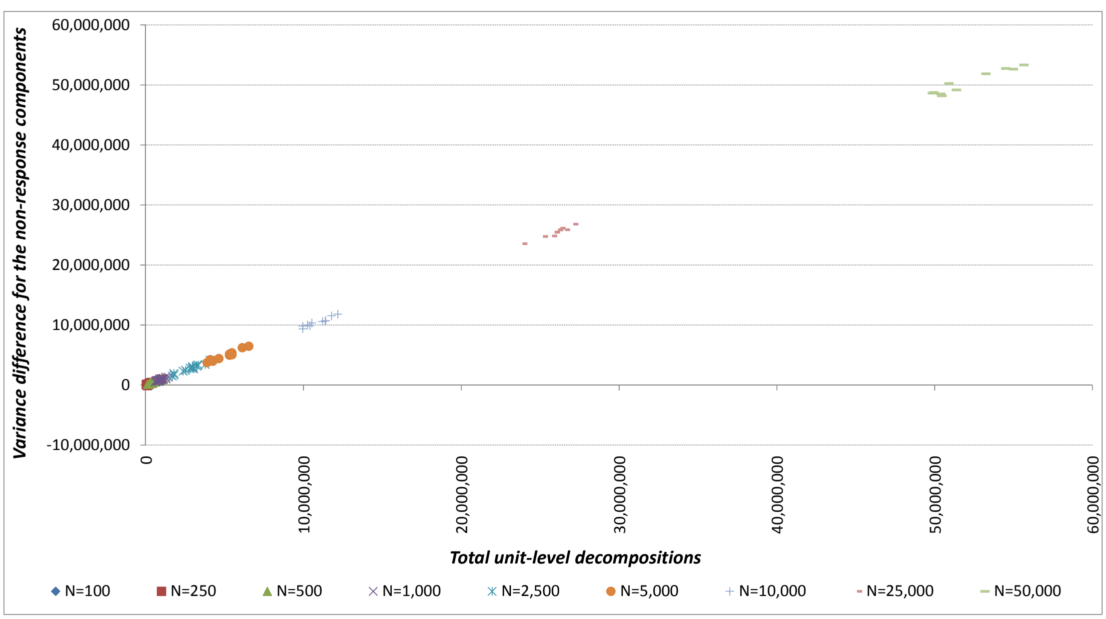

Figure 4.3 Variance difference for the non-response components versus total unit-level decompositions, varying population sizes.

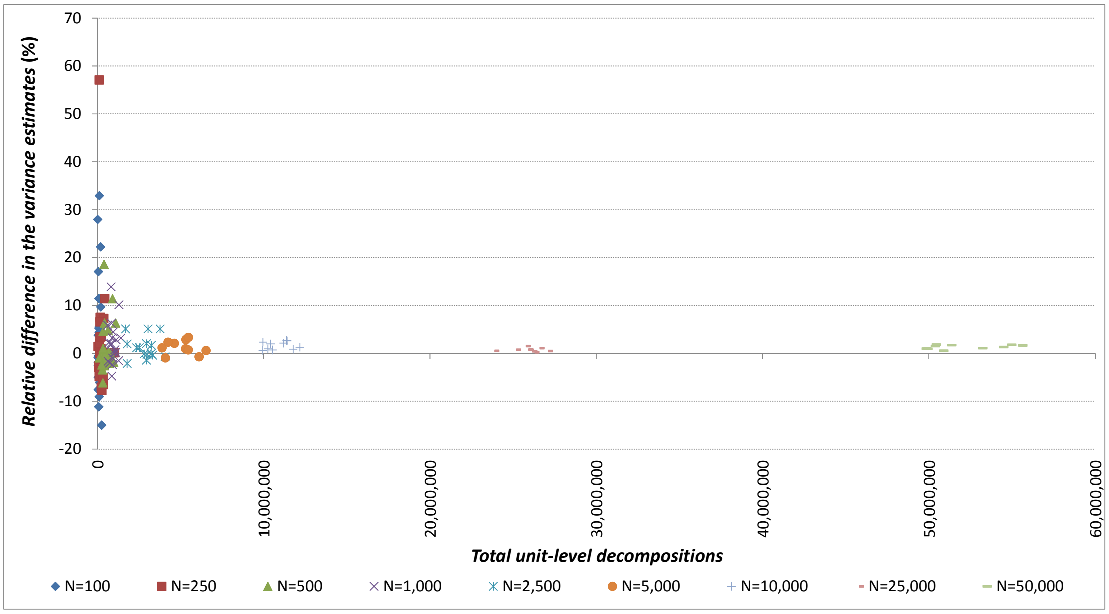

Figure 4.4 Relative difference in the variance estimates versus total unit-level decompositions, varying population sizes.

Both Figures 4.3 and 4.4 show that the relative differences in the decomposition errors are more volatile for smaller populations but rapidly converge close to 0 as population and sample sizes increase. This is further confirmed by Table 4.1.

Table 4.1
Count, average and standard deviation of relative differences in the variance estimates by population sizes

Population Size (N)	Relative Differences in the Variance Estimates in percentage		
	Count	Average	Standard Deviation
100	33(*)	2.2	10.6
250	30	1.6	11.4
500	25	1.0	5.3
1,000	20	2.2	4.4
2,500	10	1.2	2.3
5,000	10	1.2	1.4
10,000	10	1.6	0.8
25,000	10	0.7	0.4
50,000	10	1.3	0.4
Grand Total	163	1.6	7.3

(*): Out of 40 replicates created, only 33 had converted units.

To identify the sources of this instability, the relative differences between the estimated imputation variance, $\text{DRel}\left(\hat{\sigma}_k^2\right) = \left(\hat{\sigma}_k^2 - \hat{\sigma}_k^{2(\Lambda)}\right) \big/ \hat{\sigma}_k^{2(\Lambda)}$, and the relative difference between the estimated imputation relationship element, $\text{DRel}\left(\varphi_{lk}\right) = \left(\varphi_{lk} - \varphi_{lk}^{(\lambda)}\right) \big/ \varphi_{lk}^{(\lambda)}$, were measured for all units, $k \notin \Lambda$ and $l \notin \Lambda$. Note that under the ratio imputation model, both are constant for a given replicate, i.e., $\mathrm{DRel}\left(\hat{\sigma}_k^2\right) = \mathrm{DRel}\left(\hat{\sigma}^2\right)$ and $\mathrm{DRel}\left(\varphi_{lk}\right) = \mathrm{DRel}\left(\varphi\right)$. After the deletion of 2 extreme replicates, the correlation between the relative difference in the variance estimates DRel and $\mathrm{DRel}\left(\hat{\sigma}^2\right)$ is 0.78 while the correlation between DRel and $\mathrm{DRel}\left(\varphi\right)$ is 0.01. This illustrates that the instability is primarily caused by the variability of the $\hat{\sigma}_l^{(\lambda)} = \hat{\sigma}_l$ estimates. From this scenario, the conclusions are:

- Assumption 2 becomes valid for large enough sample sizes and leads to more accurate unit-level decomposition for consistent imputation model variance estimators.
- The unit-level decomposition is robust to assumption 3 validity.

4.3 Simulation scenario 3: Varying conversion rates

In scenario 3, the population and sample sizes were fixed to 2,500 and 500 respectively, and response rate is set to 50%. The conversion rates (CR) varied from 10% to 100% by increments of 10%, in order to generate different sizes of subset Λ, with 15 iterations each. The results are shown in Figures 4.5 and 4.6.

Both Figures 4.5 and 4.6 show that the relative difference in the decomposition errors becomes biased as the size of Λ increases, as confirmed in Table 4.2. This is primarily due to non-linearity of $\hat{V}_{\mathrm{NR}}\left(\hat{t}_d\right)$, as demonstrated in equation (3.6). The monotone nature of the relationship in Figure 4.5 suggests that the ordering of the error contributors is not affected, i.e., the large estimated contributors will have larger effect on the variance than the ones with a small estimated contribution.

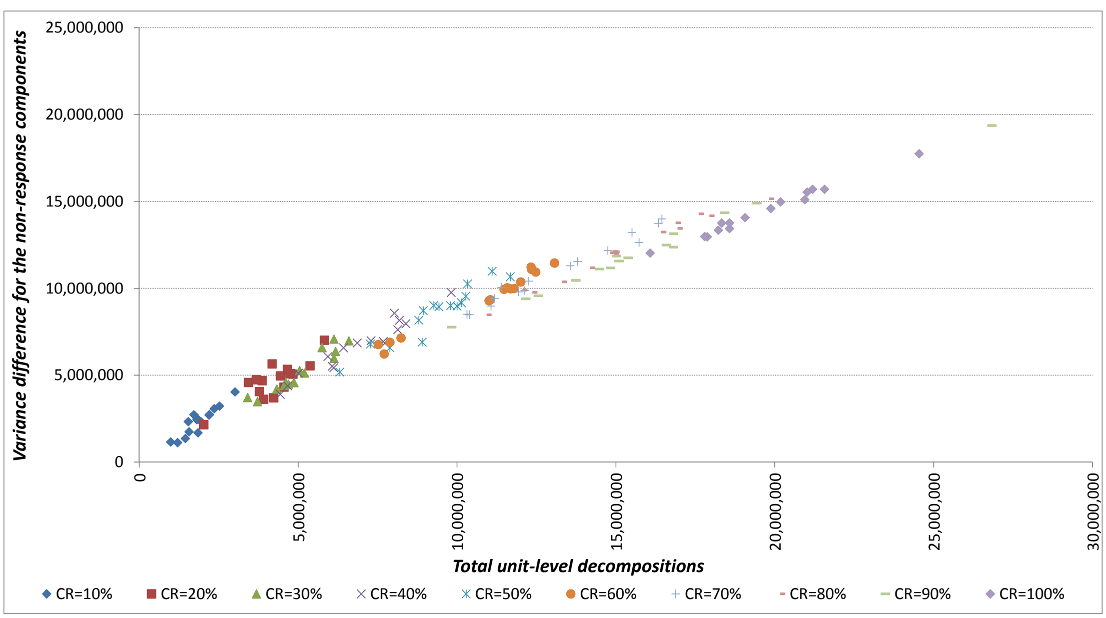

Figure 4.5 Variance difference for the non-response components versus total unit-level decompositions, varying conversion rates (CR).

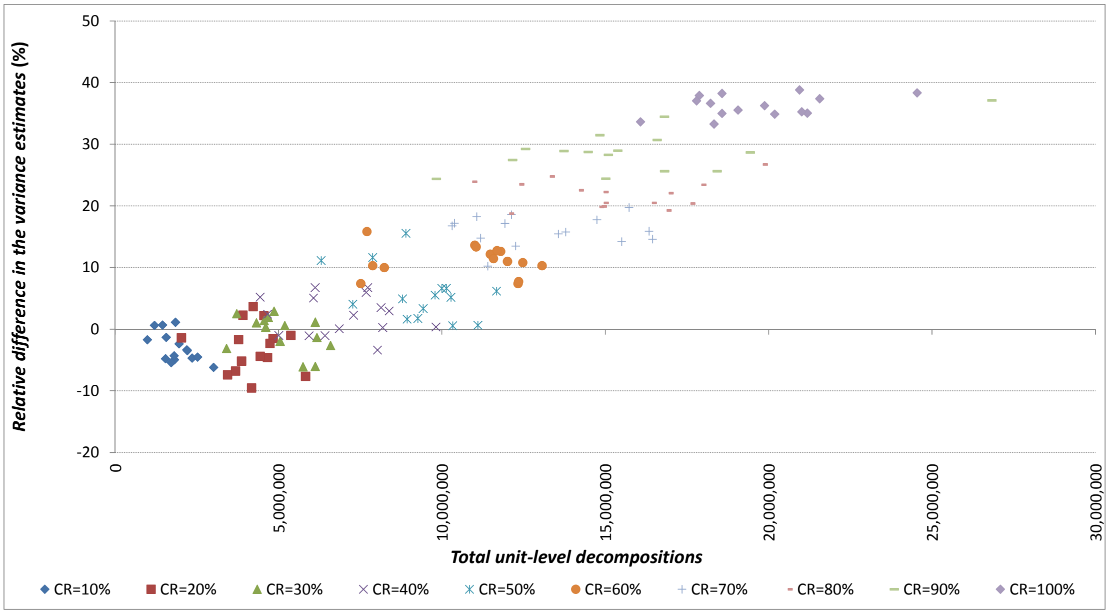

Figure 4.6 Relative difference in the variance estimates versus total unit-level decompositions, varying conversion rates (CR).

Table 4.2
Count, average, and standard deviation of relative differences in the variance estimates by conversion rates (CR)

Conversion Rate (CR)	Relative Differences in the Variance Estimates in percentage		
	Count	Average	Standard Deviation
10%	15	-3.0	2.4
20%	15	-3.0	3.9
30%	15	-0.5	3.0
40%	15	2.2	3.3
50%	15	5.7	4.3
60%	15	11.1	2.4
70%	15	16.0	2.4
80%	15	21.9	2.3
90%	15	28.9	3.5
100%	15	36.2	1.7
Grand Total	150	11.5	13.5

Despite the fact that the relative differences in the variance estimates are not null on average, it doesn't prevent the use of the proposed decomposition of errors to identify the largest sources of variance, especially in asymmetric populations. Mills et al. (2013) showed through a simulation how this could be successfully adapted into an efficient active collection strategy.

5 Conclusion

The proposed unit-level score is a good approximation of the unit impact on the variance due to non-response. It is applicable for different survey designs, compliant with calibration estimators for domain totals and works with many common imputation methods. The assumptions on which the decomposition relies are generally valid in common surveys using unbiased imputation methods and consistent estimators of imputation model parameters. The simulation results show that this approach becomes more accurate with larger sample sizes. The decomposition of the non-response variance is biased due to its non-linearity. However, the bias is smaller in asymmetric populations and when focusing on a small number of nonresponding units. The fact that the ordering of units using the estimated contribution to variance due to non-response is similar to the real order is an important aspect when the priority is to identify the largest contributors, not necessarily their actual contributions, to the total error.

This paper presented the method in a univariate context but it can be easily extended to a multivariate framework, using a distance function to combine the item contributions into a unit contribution. The idea remains to focus our attention in terms of collection treatments or manual verification on cases where the unit scores are the highest. In this case the non-response follow-up treatment might be different for unit non-response compared to partial non-response. For example, a telephone follow-up could be used to collect all the items for the total nonresponding units with the larger score; and the partial nonrespondents with a large score could be sent to an analyst for review, depending on the budget for follow-up. Moreover, if this score can be computed several times during the collection period, then non-response follow-ups will be more efficient because the unit score will be more accurate and the quality might become satisfactory for some estimates. Simulation results show that the proposed score is a good approximation to the contribution of a unit to the variance due to non-response. Subsequently, this score could be used to determine how many and which nonresponding units should be followed in order to reach a given estimated coefficient of variation.

This work was initially done for non-response prioritization under the Rolling Estimate iterative adaptive design process for IBSP. Following the original plan, key item estimates would be computed with their associated quality indicators at several specific times during the collection period. After each specific time, the units with the largest contributions according to our method would be prioritized for follow-up.

Acknowledgements

The authors want to thank the reviewers (Cynthia Bocci and Jessica Andrews), the associate editor, the referees and the assistant editor for their valuable feedback.

Appendix

Proof 1

$$\sum_{k \in s_m} \delta_k \left(\hat{V}_{\text{DIF}} \left(\hat{t}_d \right) \right) = \sum_{k \in s_m} (1 - \pi_k) \, w_k^2 d_k \hat{\sigma}_k^2 = \hat{V}_{\text{DIF}} \left(\hat{t}_d \right).$$

Proof 2

$$\begin{aligned}
\sum_{k\in s_m}\delta_k\left(\hat{V}_{\text{MIX}}\left(\hat{t}_d\right)\right) &= \sum_{k\in s_m}\left(2\sum_{l\in s_r} w_k d_k \varphi_{lk}\left(w_l-1\right)d_l\hat{\sigma}_l^2 - 2w_k\left(w_k-1\right)d_k\hat{\sigma}_k^2\right)\\
&= 2\sum_{k\in s_m}\sum_{l\in s_r} w_k d_k \varphi_{lk}\left(w_l-1\right)d_l\hat{\sigma}_l^2 - 2\sum_{k\in s_m} w_k\left(w_k-1\right)d_k\hat{\sigma}_k^2\\
&= 2\sum_{l\in s_r}\sum_{k\in s_m} w_k d_k \varphi_{lk}\left(w_l-1\right)d_l\hat{\sigma}_l^2 - 2\sum_{k\in s_m} w_k\left(w_k-1\right)d_k\hat{\sigma}_k^2\\
&= 2\sum_{l\in s_r}\left(\sum_{k\in s_m} w_k d_k \varphi_{lk}\right)\left(w_l-1\right)d_l\hat{\sigma}_l^2 - 2\sum_{k\in s_m} w_k\left(w_k-1\right)d_k\hat{\sigma}_k^2\\
&= 2\sum_{l\in s_r} W_{dl}\left(w_l-1\right)d_l\hat{\sigma}_l^2 - 2\sum_{k\in s_m} w_k\left(w_k-1\right)d_k\hat{\sigma}_k^2\\
&= \hat{V}_{\text{MIX}}\left(\hat{t}_d\right).
\end{aligned}$$

Proof 3

$$\begin{aligned}
\sum_{k\in s_m}\delta_k\left(\hat{V}_{\text{NR}}\left(\hat{t}_d\right)\right) &= \sum_{k\in s_m}\left(\sum_{l\in s_r}\left(2W_{dl}w_k d_k\varphi_{lk} - w_k^2 d_k\varphi_{lk}^2\right)\hat{\sigma}_l^2 + w_k^2 d_k\hat{\sigma}_k^2\right)\\
&= \sum_{k\in s_m}\left(\sum_{l\in s_r}\left(2W_{dl}w_k d_k\varphi_{lk} - w_k^2 d_k\varphi_{lk}^2\right)\hat{\sigma}_l^2\right) + \sum_{k\in s_m} w_k^2 d_k\hat{\sigma}_k^2\\
&= \sum_{l\in s_r}\left(\sum_{k\in s_m}\left(2W_{dl}w_k d_k\varphi_{lk} - w_k^2 d_k\varphi_{lk}^2\right)\hat{\sigma}_l^2\right) + \sum_{k\in s_m} w_k^2 d_k\hat{\sigma}_k^2\\
&= \sum_{l\in s_r}\left(2W_{dl}\sum_{k\in s_m} w_k d_k\varphi_{lk}\hat{\sigma}_l^2 - \sum_{k\in s_m} w_k^2 d_k\varphi_{lk}^2\hat{\sigma}_l^2\right) + \sum_{k\in s_m} w_k^2 d_k\hat{\sigma}_k^2\\
&= \sum_{l\in s_r}\left(2W_{dl}^2\hat{\sigma}_l^2 - \sum_{k\in s_m} w_k^2 d_k\varphi_{lk}^2\hat{\sigma}_l^2\right) + \sum_{k\in s_m} w_k^2 d_k\hat{\sigma}_k^2\\
&= \sum_{l\in s_r} 2W_{dl}^2\hat{\sigma}_l^2 - \sum_{l\in s_r}\sum_{k\in s_m} w_k^2 d_k\varphi_{lk}^2\hat{\sigma}_l^2 + \sum_{k\in s_m} w_k^2 d_k\hat{\sigma}_k^2\\
&= \hat{V}_{\text{NR}}\left(\hat{t}_d\right) + \sum_{l\in s_r} W_{dl}^2\hat{\sigma}_l^2 - \sum_{l\in s_r}\sum_{k\in s_m} w_k^2 d_k\varphi_{lk}^2\hat{\sigma}_l^2\\
&= \hat{V}_{\text{NR}}\left(\hat{t}_d\right) + \sum_{l\in s_r}\left(W_{dl}^2 - \sum_{k\in s_m} w_k^2 d_k\varphi_{lk}^2\right)\hat{\sigma}_l^2\\
&= \hat{V}_{\text{NR}}\left(\hat{t}_d\right) + \sum_{l\in s_r}\left(\left(\sum_{k\in s_m} w_k d_k\varphi_{lk}\right)^2 - \sum_{k\in s_m} w_k^2 d_k\varphi_{lk}^2\right)\hat{\sigma}_l^2.
\end{aligned}$$

References

Beaumont, J.-F., and Bissonnette, J. (2011). Variance estimation under composite imputation: The methodology behind SEVANI. *Survey Methodology*, 37, 2, 171-179. Paper available at https://www150.statcan.gc.ca/n1/pub/12-001-x/2011002/article/11605-eng.pdf.

Beaumont, J.-F., and Bocci, C. (2009). Variance estimation when donor imputation is used to fill in missing values. *Canadian Journal of Statistics*, 37, 400-416.

Beaumont, J.-F., Bocci, C. and Haziza, D. (2014). An adaptive data collection procedure for call prioritization. *Journal of Official Statistics*, 30, 607-621.

Beaumont, J.-F., Haziza, D. and Bocci, C. (2011). On variance estimation under auxiliary value imputation in sample surveys. *Statistica Sinica*, 21, 515-537.

Biemer, P.P. (2010). Total survey error: Design, implementation, and evaluation. *Public Opinion Quarterly*, 74, 5, 817-848.

Bosa, K., and Godbout, S. (2014). *IBSP Quality Measures – Methodology Guide*. Business Survey Methods Division. Internal document.

Godbout, S., Beaucage, Y. and Turmelle, C. (2011). Achieving quality and efficiency using a top-down approach in the Canadian integrated business statistics Program. *Proceedings of the Conference of European Statisticians*. United Nations Statistical Commission and Economic Commission for Europe. Work Session on Statistical Data Editing. Ljubljana, Slovenia, 9-11 May 2011.

Groves, R.M., and Heeringa, S.G. (2006). Responsive design for household surveys: Tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society, Series A*, 169, No. 3, 439-457.

Mills, F., Godbout, S., Bosa, K. and Turmelle, C. (2013). Multivariate selective editing in the integrated business statistics program. *Proceedings of the Joint Statistical Meeting 2013 - Survey Research Methods Section*. August 2013. Montréal, Canada.

Särndal, C.-E. (1992). Methods for estimating the precision of survey estimates when imputation has been used. *Survey Methodology*, 18, 2, 241-252. Paper available at https://www150.statcan.gc.ca/n1/pub/12-001-x/1992002/article/14483-eng.pdf.

Schouten, B., Calinescu, M. and Luiten, A. (2013). Optimizing quality of response through adaptive survey designs. *Survey Methodology*, 39, 1, 29-58. Paper available at https://www150.statcan.gc.ca/n1/pub/12-001-x/2013001/article/11824-eng.pdf.

Statistics Canada (2015). *Integrated Business Statistics Program Overview*. Statistics Canada Catalogue no. 68-515-X. Ottawa.

Turmelle, C., Godbout, S. and Bosa, K. (2012). Methodological challenges in the development of Statistics Canada's new integrated business statistics program. *Proceedings of the International Conference on Establishment Surveys IV*. Montréal, Canada.