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

Survey Methodology

Small area benchmarked estimation under the basic unit level model when the sampling rates are non-negligible

by Marius Stefan and Michael A. Hidiroglou

Release date: June 24, 2021

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.

Published by authority of the Minister responsible for Statistics Canada

An HTML version is also available.

Cette publication est aussi disponible en français.

Survey Methodology, June 2021
Vol. 47, No. 1, pp. 123-149
Statistics Canada, Catalogue No. 12-001-X

Small area benchmarked estimation under the basic unit level model when the sampling rates are non-negligible

Marius Stefan and Michael A. Hidiroglou[1]

Abstract

We consider the estimation of a small area mean under the basic unit-level model. The sum of the resulting model-dependent estimators may not add up to estimates obtained with a direct survey estimator that is deemed to be accurate for the union of these small areas. Benchmarking forces the model-based estimators to agree with the direct estimator at the aggregated area level. The generalized regression estimator is the direct estimator that we benchmark to. In this paper we compare small area benchmarked estimators based on four procedures. The first procedure produces benchmarked estimators by ratio adjustment. The second procedure is based on the empirical best linear unbiased estimator obtained under the unit-level model augmented with a suitable variable that ensures benchmarking. The third procedure uses pseudo-empirical estimators constructed with suitably chosen sampling weights so that, when aggregated, they agree with the reliable direct estimator for the larger area. The fourth procedure produces benchmarked estimators that are the result of a minimization problem subject to the constraint given by the benchmark condition. These benchmark procedures are applied to the small area estimators when the sampling rates are non-negligible. The resulting benchmarked estimators are compared in terms of relative bias and mean squared error using both a design-based simulation study as well as an example with real survey data.

Key Words: Small area; Benchmarking; Empirical estimator; Pseudo-empirical estimator; Constrained estimator.

1. Introduction

Small area estimation (SAE) has grown in importance in recent years due to the demand for reliable small area statistics. Direct estimators are used to estimate parameters of interest when the sample size is reasonably large. However, they have large standard errors and coefficients of variation when it comes to applying them to small areas, as the realized sample size will be quite small. It is therefore necessary to use models that borrow strength from other related areas or from past surveys to have stable estimators for these small areas. Model-based estimates typically show a substantial improvement over direct estimates in terms of mean squared error (MSE).

The available theory for small area estimation is based on either area-level or unit-level models, depending on the level of available auxiliary information. Unit-level based methods use the data of the individual units as auxiliary information, whereas area-level based methods use aggregates or means of the data of the units within the small areas. Fay and Herriot (1979), denoted hereafter as the FH model, is the most used area-level model in small area estimation. The one-fold nested error regression model proposed in Battese, Harter and Fuller (1988), also known as the basic unit-level model, is frequently used when unit-level information is available. We denote this model as the BHF model. Both are special cases of a general linear mixed model in SAE (see Rao and Molina, 2015 for an excellent account of the small area estimation).

Small area means or totals are the most frequent linear parameters estimated in SAE. In these cases, the most popular small area method is the use of linear mixed models to derive the best linear unbiased

1. Marius Stefan, Faculty of Applied Sciences, Polytechnic University of Bucharest, Splaiul Independentei, nr. 313. E-mail: mastefan@gmail.com; Michael A. Hidiroglou, Statistics Canada Alumnus. E-mail: hidirog@yahoo.ca.

predictors (BLUP) for the small area mean or total. BLUP estimators minimize the MSE among the class of linear unbiased estimators. Alternatively, it can be shown that the BLUP estimator can be obtained by solving mixed model equations with unknowns given by the fixed and random parameters of the model. The mixed model equations result from the maximization of the joint density of the data and the vector of random small area effects. A BLUP estimator depends on the variances (and covariances) of random effects which can be estimated by the Henderson method of fitting constants (FC), the maximum likelihood (ML) or restricted maximum likelihood (REML). Using these estimated components in the BLUP estimator leads to a two-stage estimator referred to as the empirical best linear unbiased predictor (EBLUP).

A potential difficulty with EBLUP estimators is that when they are aggregated over all the small areas, they may not agree with the overall estimate for a larger area obtained via direct estimation. Statistical agencies favor an overall agreement between the sum of the model-based small area estimates and the direct estimate at a higher level that corresponds to the union of the small areas. Benchmarking is a method of modifying the model-based estimates to agree with the direct estimator for the larger area.

Existing benchmarking methods are either frequentist or Bayesian. In this paper, we focus on the frequentist approach to benchmarking (for Bayesian benchmarking procedures, see You, Rao and Dick, 2004; Datta, Ghosh, Steorts and Maples, 2011 and Nandram and Sayit, 2011). The frequentist methods can be applied to obtain benchmark small area estimates for both the area-level and unit-level models.

We briefly summarize the existing literature for both types of models. We first describe the procedures developed to benchmark area-level based estimates. Pfeffermann and Barnard (1991) obtained a constrained benchmarked estimator by maximizing the joint density of the data and the vector of random small area effects given the benchmark restriction. Their benchmark estimator was constructed with modified estimates of fixed and small area effects that are solutions to the constrained maximization problem. Wang, Fuller and Qu (2008) developed a benchmarked EBLUP for the FH area-level model, by minimizing a loss function subject to the constraint given by the benchmark condition. They obtained a second benchmarked estimator by adding a suitable auxiliary variable to the FH model without imposing a constraint. They showed that the EBLUP estimator based on the augmented FH model is self-benchmarked: the estimator satisfied the benchmark condition without further adjustments. Bell, Datta and Ghosh (2013) generalized the result in Wang et al. (2008) to the case of multiple benchmark constraints by considering a more general loss function. You, Rao and Hidiroglou (2013) obtained another self-benchmarked estimator under the FH model by replacing the regression vector used in the EBLUP estimator with an alternative estimator that depends on the benchmarking weights.

We now turn to procedures that benchmark unit-level model-based estimates. The objective is to obtain small area estimators that benchmark to a direct estimator at a given level of aggregation of the small areas. The direct estimators that are mostly used by Statistical agencies are the Generalized Regression Estimator (GREG) in Särndal, Swensson and Wretman (1989) or more generally the calibration estimator based on procedures in Deville and Särndal (1992). You and Rao (2002) developed a pseudo-EBLUP predictor (YR predictor) that incorporates survey weights. A property of this estimator is that it is

self-benchmarked, that is, the sum of the small area estimates adds up to an estimator that has the same form as the GREG. However, it is not a direct estimator because the estimated regression vector that is part of this estimator reflects the error structure of the nested error model. Assuming that the sampling rates are negligible, Stefan and Hidiroglou (2020) proposed several procedures to ensure that the EBLUP and pseudo-EBLUP estimators would benchmark to the GREG estimator, given that both the model and the GREG estimator used the same vector of auxiliary variables. Ugarte, Militino and Goicoa (2009) developed a restricted EBLUP estimator for a small area total that satisfies the benchmarking property to a synthetic estimator.

The objective of this paper is to compare several benchmarked estimators of a small area mean for the basic unit level model when the sampling rates are non-negligible. We compare six benchmarked estimators: two benchmarked estimators based on the procedures proposed by Stefan and Hidiroglou (2020), two restricted estimators based on the procedure proposed by Ugarte et al. (2009) and two ratio estimators obtained by multiplying each small area EBLUP and YR estimators by a common adjustment factor. The paper is organized as follows. Section 2 presents a summary of EBLUP and pseudo-EBLUP estimators under the basic unit-level model. Section 3 describes the six benchmarked estimators. The first two estimators are based on simple ratio adjustments. Then, we show how the two benchmarking procedures proposed by Stefan and Hidiroglou (2020) in the case of negligible sampling rates can be adapted to produce benchmarked small area mean estimators when the sampling rates are non-negligible. Finally, we describe the restricted EBLUP estimator of Ugarte et al. (2009) and propose a pseudo restricted estimator which is a variant of the restricted EBLUP that incorporates survey weights. We also propose a re-parameterized restricted maximum likelihood (reREML) method for estimating the variance components. This method of estimation is useful when computing restricted EBLUP small area mean estimators as it results in strictly positive variance components estimates. Section 4 presents the results of a Monte Carlo simulation based on generated data sets, whereas Section 5 reports the results of a simulation study based on a real data set. Finally, Section 6 gives some concluding remarks.

2. EBLUP and pseudo-EBLUP estimation

Consider the one-fold nested error regression model

$$y_{ij} = \mathbf{x}_{ij}^T\boldsymbol{\beta} + v_i + e_{ij},\ i = 1, \ldots, m;\ j = 1, \ldots, N_i, \tag{2.1}$$

where y_{ij} is the variable of interest for the j^{th} population unit in the i^{th} small area, $\mathbf{x}_{ij} = \left(x_{ij1}, \ldots, x_{ijp}\right)^T$ is a vector of auxiliary variables with $x_{ij1} = 1$, $\boldsymbol{\beta} = \left(\beta_1, \ldots, \beta_p\right)^T$ is a $p \times 1$ vector of regression parameters and N_i is the number of population units in the i^{th} small area, U_i. The random small area effects v_i are assumed to be i.i.d. $N\left(0, \sigma_v^2\right)$, and independent of the unit errors e_{ij}, which are assumed i.i.d. $N\left(0, \sigma_e^2\right)$. We draw samples s_i of size n_i independently within each small area i, according to a specified sampling design with first-order inclusion probabilities denoted by π_{ij}, for $j = 1, \ldots, N_i$. The total sample size is n, where $n = \sum_{i=1}^{m} n_i$. The resulting basic design weights are given by $d_{ij} = 1/\pi_{ij}$.

We assume that the sample design is ignorable, and that selection bias is absent. This implies that model (2.1) also holds for the sample data:

$$y_{ij} = \mathbf{x}_{ij}^T\boldsymbol{\beta} + v_i + e_{ij},\ i = 1, \ldots, m;\ j = 1, \ldots, n_i, \tag{2.2}$$

Model (2.2) is a special case of the general linear mixed model. Defining $\mathbf{y}_i = \left(y_{i1}, \ldots, y_{in_i}\right)^T$, $\mathbf{X}_i = \left(\mathbf{x}_{i1}^T, \ldots, \mathbf{x}_{in_i}^T\right)^T$, $\mathbf{v} = (v_1, \ldots, v_m)^T$ and $\mathbf{e}_i = \left(e_{i1}, \ldots, e_{in_i}\right)^T$, it follows that model (2.2) can be expressed in a matrix form by stacking the observations. The resulting equation is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{v} + \mathbf{e}, \tag{2.3}$$

where $\mathbf{y} = \text{col}_{1\le i\le m}(\mathbf{y}_i)$, $\mathbf{X} = \text{col}_{1\le i\le m}(\mathbf{X}_i)$, $\mathbf{Z} = \text{diag}_{1\le i\le m}\{\mathbf{1}_{n_i}\}$ and $\mathbf{e} = \text{col}_{1\le i\le m}(\mathbf{e}_i)$, with $\mathbf{1}_{n_i}$ a vector of dimension n_i composed of ones. We denote by $\mathbf{G}$ and $\mathbf{R}$ the variance matrices of the random vectors $\mathbf{v}$ and $\mathbf{e}$ respectively. Then $\mathbf{G} = \sigma_v^2\mathbf{I}_m$ and $\mathbf{R} = \sigma_e^2\mathbf{I}_n$. It follows that the variance matrix of vector $\mathbf{y}$, denoted as $\mathbf{V}$, is given by $\mathbf{V} = \mathbf{R} + \mathbf{Z}\mathbf{G}\mathbf{Z}^T$.

The parameters of interest are the small area means $\bar{Y}_i$, where $\bar{Y}_i = N_i^{-1}\sum_{j=1}^{N_i} y_{ij}$, $i = 1, \ldots, m$. If N_i is large, the sampling fraction $f_i = N_i^{-1}n_i$, of the i^{th} small area is negligible. This set-up corresponds to the case of an *infinite population* or negligible sampling rates. It follows that the small area means $\bar{Y}_i$ can be approximated by μ_i (see Rao and Molina, 2015, page 174), where $\mu_i = \bar{\mathbf{X}}_i^T\boldsymbol{\beta} + v_i$ and $\bar{\mathbf{X}}_i = \sum_{j=1}^{N_i}\mathbf{x}_{ij}/N_i$ is the vector of population means of the $\mathbf{x}_{ij}$'s for the i^{th} area. An estimator of μ_i is given by $\hat{\mu}_i = \bar{\mathbf{X}}_i^T\hat{\boldsymbol{\beta}} + \hat{v}_i$ (Rao and Molina, 2015, page 175), where $\hat{\boldsymbol{\beta}}$ and $\hat{v}_i$ are estimators of $\boldsymbol{\beta}$ and v_i respectively. If N_i is not large enough or if the sampling rates f_i are not negligible, parameters $\bar{Y}_i$ cannot be approximated by linear combinations of $\boldsymbol{\beta}$ and v_i. This corresponds to the case of a *finite population*. Let r_i be the set of the $N_i - n_i$ unobserved y-values in small area i. If we assume that we know the $\mathbf{x}_{ij}$'s for each individual in the population, an estimator $\hat{\bar{Y}}_i$ of $\bar{Y}_i$ is based on the observed values y_{ij}, $j \in s_i$, and predicted values $\hat{y}_{ij} = \mathbf{x}_{ij}^T\hat{\boldsymbol{\beta}} + \hat{v}_i$ for $j \in r_i$. That is, estimator $\hat{\bar{Y}}_i$ is given by

$$\hat{\bar{Y}}_i = \frac{1}{N_i}\left(\sum_{j\in s_i} y_{ij} + \sum_{j\in r_i}\hat{y}_{ij}\right). \tag{2.4}$$

Much of the SAE theory deals with the infinite population case, whereas the literature on the finite population case is more limited. In this paper we focus on finite population (or non-negligible sampling rates) case, thereby constructing estimators based on (2.4).

2.1 EBLUP estimation

We denote by $\tilde{\boldsymbol{\beta}}$ and $\tilde{\mathbf{v}} = (\tilde{v}_1, \ldots, \tilde{v}_m)^T$ the BLUP predictors of $\boldsymbol{\beta}$ and $\mathbf{v}$ respectively. These estimators are given by $\tilde{\boldsymbol{\beta}} = \left(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{y}$ and $\tilde{\mathbf{v}} = \mathbf{G}\mathbf{Z}^T\mathbf{V}^{-1}\left(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}\right)$. Under the normality assumption of $\mathbf{e}$ and $\mathbf{v}$, it can be shown that $\tilde{\boldsymbol{\beta}}$ and $\tilde{\mathbf{v}}$ can be obtained by maximizing the joint density of $\mathbf{y}$ and $\mathbf{v}$ with respect to $\boldsymbol{\beta}$ and $\mathbf{v}$. This is equivalent to minimizing the function

$$\phi = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{v})^T\mathbf{R}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{v}) + \mathbf{v}^T\mathbf{G}^{-1}\mathbf{v}. \tag{2.5}$$

This leads to the following mixed model equations

$$\mathbf{A}\begin{pmatrix}\tilde{\boldsymbol{\beta}}\\ \tilde{\mathbf{v}}\end{pmatrix} = \mathbf{b}, \tag{2.6}$$

where

$$\mathbf{A} = \begin{pmatrix}\mathbf{X}^T\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}^T\mathbf{R}^{-1}\mathbf{Z}\\ \mathbf{Z}^T\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}^T\mathbf{R}^{-1}\mathbf{Z}+\mathbf{G}^{-1}\end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix}\mathbf{X}^T\mathbf{R}^{-1}\mathbf{y}\\ \mathbf{Z}^T\mathbf{R}^{-1}\mathbf{y}\end{pmatrix}. \tag{2.7}$$

(see Rao and Molina, 2015, page 99 for details). The variance components (σ_v^2, σ_e^2) in equations (2.6) and (2.7) are generally unknown. Three methods of estimation, FC, ML and REML, are commonly used in SAE to estimate the variance components (σ_v^2, σ_e^2). A well-known difficulty with these methods is that the estimate of σ_v^2 can take on negative values. This estimate is truncated to zero when this occurs, that is $\hat{\sigma}_v^2$ is set to 0. Empirical versions of $\mathbf{A}$ and $\mathbf{b}$, denoted as $\hat{\mathbf{A}}$ and $\hat{\mathbf{b}}$, are obtained if the unknown variance components (σ_v^2, σ_e^2) are replaced by estimators $(\hat{\sigma}_v^2, \hat{\sigma}_e^2)$. It follows from equation (2.6) that EBLUP estimators of model parameters $(\boldsymbol{\beta}, \mathbf{v})$, denoted as $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{v}} = (\hat{v}_1, \ldots, \hat{v}_m)^T$, are given by

$$\begin{pmatrix}\hat{\boldsymbol{\beta}}\\ \hat{\mathbf{v}}\end{pmatrix} = \hat{\mathbf{A}}^{-1}\hat{\mathbf{b}}. \tag{2.8}$$

Using (2.8), it can be proved that $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{v}}$ are

$$\begin{pmatrix}\hat{\boldsymbol{\beta}}\\ \hat{\mathbf{v}}\end{pmatrix} = \begin{pmatrix}\left(\mathbf{X}^T\hat{\mathbf{V}}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\hat{\mathbf{V}}^{-1}\mathbf{y}\\ \hat{\mathbf{G}}\mathbf{Z}^T\hat{\mathbf{V}}^{-1}\left(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}\right)\end{pmatrix}, \tag{2.9}$$

where $\hat{\mathbf{G}} = \hat{\sigma}_v^2\mathbf{I}_m$ and $\hat{\mathbf{V}} = \hat{\sigma}_e^2\mathbf{I}_n + \hat{\sigma}_v^2\mathbf{Z}\mathbf{Z}^T$.

Remark 1. It is easier to invert matrices $\hat{\mathbf{G}} = \hat{\sigma}_v^2\mathbf{I}_m$ and $\hat{\mathbf{R}} = \hat{\sigma}_e^2\mathbf{I}_n$ than $\hat{\mathbf{V}}$. Consequently, it is simpler to use the mixed model equations (2.8) than equations (2.9) for computing $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{v}}$. However, when $\hat{\sigma}_v^2$ is equal to zero, equations (2.8) cannot be used because the $\hat{\mathbf{G}}^{-1}$ term in matrix $\hat{\mathbf{A}}$ does not exist. In such cases, $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{v}}$ can only be computed using (2.9).

Under model (2.2), it can be shown that $\hat{\boldsymbol{\beta}}$ and $\hat{v}_i$ in $\hat{\mathbf{v}} = (\hat{v}_1, \ldots, \hat{v}_m)^T$ satisfy

$$\sum_{i=1}^{m}\sum_{j\in s_i}\mathbf{x}_{ij}\left(y_{ij} - \mathbf{x}_{ij}^T\hat{\boldsymbol{\beta}} - \hat{v}_i\right) = 0. \tag{2.10}$$

Estimators $\hat{\boldsymbol{\beta}}$ and $\hat{v}_i$ are used to compute EBLUP predictions $\hat{y}_{ij}^{\text{EBLUP}}$ for the $N_i - n_i$ unobserved units in small area i: $\hat{y}_{ij}^{\text{EBLUP}} = \mathbf{x}_{ij}^T\hat{\boldsymbol{\beta}} + \hat{v}_i$ for $j \in r_i$. An EBLUP estimator of $\bar{Y}_i$, denoted as $\hat{\bar{Y}}_i^{\text{EBLUP}}$, is obtained by replacing in (2.4) $\hat{y}_{ij}$ by $\hat{y}_{ij}^{\text{EBLUP}}$. It follows that $\hat{\bar{Y}}_i^{\text{EBLUP}}$ is

$$\hat{\bar{Y}}_i^{\text{EBLUP}} = \frac{1}{N_i}\left[\sum_{j\in s_i} y_{ij} + \mathbf{x}_{ir}^T\hat{\boldsymbol{\beta}} + (N_i - n_i)\,\hat{v}_i\right], \tag{2.11}$$

where $\mathbf{x}_{ir} = \sum_{j \in r_i} \mathbf{x}_{ij}$ represents the sum of non sampled values $\mathbf{x}_{ij}$.

2.2 You-Rao estimation

You and Rao (2002) proposed a pseudo-EBLUP small area mean estimator (YR estimator) that incorporates the design weights d_{ij} into the formula of the EBLUP estimator. A property of the pseudo-EBLUP estimator is that the design consistency is preserved as the area sample size increases. Furthermore, the YR predictor offers protection against model failure or an informative sampling design (see among others Hidiroglou and Estevao, 2016 and Verret, Rao and Hidiroglou, 2015 for details). Pseudo EBLUP estimators can be constructed using the procedure in You and Rao (2002) with survey weights w_{ij} that may be calibrated on some vector of auxiliary variables. Let $\hat{\boldsymbol{\beta}}^{\mathrm{YR}}$ and $\hat{\mathbf{v}}^{\mathrm{YR}} = \left(\hat{v}_1^{\mathrm{YR}}, \ldots, \hat{v}_m^{\mathrm{YR}}\right)^T$ be the YR estimators of $\boldsymbol{\beta}$ and $\mathbf{v}$ respectively based on weights w_{ij} (see You and Rao, 2002 for details). The estimators $\hat{\boldsymbol{\beta}}^{\mathrm{YR}}$ and $\hat{v}_i^{\mathrm{YR}}$ satisfy the estimating unit-level based equations

$$\sum_{i=1}^{m} \sum_{j \in s_i} w_{ij} \mathbf{x}_{ij} \left(y_{ij} - \mathbf{x}_{ij}^T \hat{\boldsymbol{\beta}}^{\mathrm{YR}} - \hat{v}_i^{\mathrm{YR}}\right) = 0. \tag{2.12}$$

Equations (2.12) represent the survey-weighted version of equations (2.10). You-Rao predictions $\hat{y}_{ij}^{\mathrm{YR}}$ of y_{ij} are computed as $\hat{y}_{ij}^{\mathrm{YR}} = \mathbf{x}_{ij}^T \hat{\boldsymbol{\beta}}^{\mathrm{YR}} + \hat{v}_i^{\mathrm{YR}}$ for $j \in r_i$. Replacing $\hat{y}_{ij}$ by $\hat{y}_{ij}^{\mathrm{YR}}$ in (2.4) leads to the YR estimator of $\bar{Y}_i$ in the case of non negligible sampling rates:

$$\hat{\bar{Y}}_i^{\mathrm{YR}} = \frac{1}{N_i}\left[\sum_{j \in s_i} y_{ij} + \mathbf{x}_{ir}^T \hat{\boldsymbol{\beta}}^{\mathrm{YR}} + \left(N_i - n_i\right) \hat{v}_i^{\mathrm{YR}}\right]. \tag{2.13}$$

Estimators $\hat{\boldsymbol{\beta}}^{\mathrm{YR}}$ and $\hat{\mathbf{v}}^{\mathrm{YR}}$ can alternatively be obtained as solutions to weighted mixed model equations similar to (2.6) (see Huang and Hidiroglou, 2003 for details). To this end, we define matrices $\mathbf{W}_i = \mathrm{diag}_{1 \le j \le n_i}\{w_{ij}\}$, $\mathbf{W} = \mathrm{diag}_{1 \le i \le m}\{\mathbf{W}_i\}$ and $\boldsymbol{\Omega} = \mathrm{diag}_{1 \le i \le m}\{\omega_i\}$, where $\omega_i = \sum_{j \in s_i} w_{ij}^2 \Big/ \sum_{j \in s_i} w_{ij}$ for $i = 1, \ldots, m$. Let ϕ_w be the sample weighted version of ϕ, where

$$\phi_w = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{v})^T \mathbf{W}^{1/2}\mathbf{R}^{-1}\mathbf{W}^{1/2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{v}) + \mathbf{v}^T \boldsymbol{\Omega}^{1/2}\mathbf{G}^{-1}\boldsymbol{\Omega}^{1/2}\mathbf{v}, \tag{2.14}$$

with $\mathbf{W}^{1/2}$ and $\boldsymbol{\Omega}^{1/2}$ representing the square root of matrices $\mathbf{W}$ and $\boldsymbol{\Omega}$ respectively. In the first term of ϕ_w, the model error associated with the observation y_{ij} is weighted by the corresponding survey weight w_{ij}, whereas in the second term of ϕ_w, the factor ω_i in the diagonal matrix $\boldsymbol{\Omega}$ represents the weight attached to the small area effect v_i. It can be shown that the minimization of ϕ_w with respect to $\boldsymbol{\beta}$ and $\mathbf{v}$ leads to $\left(\hat{\boldsymbol{\beta}}^{\mathrm{YR}}, \hat{\mathbf{v}}^{\mathrm{YR}}\right)$. It follows that $\left(\hat{\boldsymbol{\beta}}^{\mathrm{YR}}, \hat{\mathbf{v}}^{\mathrm{YR}}\right)$ are given by

$$\begin{pmatrix} \hat{\boldsymbol{\beta}}^{\mathrm{YR}} \\ \hat{\mathbf{v}}^{\mathrm{YR}} \end{pmatrix} = \hat{\mathbf{A}}_w^{-1} \hat{\mathbf{b}}_w, \tag{2.15}$$

where the known values of $\mathbf{A}_w$ and $\mathbf{b}_w$ are given by

$$\mathbf{A}_w = \begin{pmatrix} \mathbf{X}^T\mathbf{W}^{1/2}\mathbf{R}^{-1}\mathbf{W}^{1/2}\mathbf{X} & \mathbf{X}^T\mathbf{W}^{1/2}\mathbf{R}^{-1}\mathbf{W}^{1/2}\mathbf{Z} \\ \mathbf{Z}^T\mathbf{W}^{1/2}\mathbf{R}^{-1}\mathbf{W}^{1/2}\mathbf{X} & \mathbf{Z}^T\mathbf{W}^{1/2}\mathbf{R}^{-1}\mathbf{W}^{1/2}\mathbf{Z} + \boldsymbol{\Omega}^{1/2}\mathbf{G}^{-1}\boldsymbol{\Omega}^{1/2} \end{pmatrix} \text{ and } \mathbf{b}_w = \begin{pmatrix} \mathbf{X}^T\mathbf{W}^{1/2}\mathbf{R}^{-1}\mathbf{W}^{1/2}\mathbf{y} \\ \mathbf{Z}^T\mathbf{W}^{1/2}\mathbf{R}^{-1}\mathbf{W}^{1/2}\mathbf{y} \end{pmatrix}, \tag{2.16}$$

and $\hat{\mathbf{A}}_w$ and $\hat{\mathbf{b}}_w$ are empirical versions of $\mathbf{A}_w$ and $\mathbf{b}_w$ obtained by estimating $\mathbf{G}$ and $\mathbf{R}$ by $\hat{\mathbf{G}} = \hat{\sigma}_v^2 \mathbf{I}_m$ and $\hat{\mathbf{R}} = \hat{\sigma}_e^2 \mathbf{I}_n$ respectively. Equation (2.15) can alternatively be written as

$$\begin{pmatrix} \hat{\boldsymbol{\beta}}^{\mathrm{YR}} \\ \hat{\mathbf{v}}^{\mathrm{YR}} \end{pmatrix} = \begin{pmatrix} \left(\mathbf{X}^T \hat{\mathbf{V}}_w^{-1} \mathbf{X}\right)^{-1} \mathbf{X}^T \hat{\mathbf{V}}_w^{-1} \mathbf{y} \\ \hat{\mathbf{G}}_\omega \mathbf{Z}^T \hat{\mathbf{V}}_w^{-1} \left(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}^{\mathrm{YR}}\right) \end{pmatrix}, \tag{2.17}$$

where $\hat{\mathbf{G}}_\omega = \boldsymbol{\Omega}^{-1/2} \hat{\mathbf{G}} \boldsymbol{\Omega}^{-1/2}$ and $\hat{\mathbf{V}}_w = \mathbf{W}^{-1/2} \hat{\mathbf{R}} \mathbf{W}^{-1/2} + \mathbf{Z}\hat{\mathbf{G}}_\omega \mathbf{Z}^T$.

3. Benchmarked estimators

We now proceed to develop benchmarked estimators of the small area means $\bar{Y}_i$ using unit level model (2.2) or augmented versions of it. We assume that a reliable direct estimator $\hat{Y}_w = \sum_{i=1}^m \sum_{j \in s_i} w_{ij} y_{ij}$ of the population total Y is available, where $Y = \sum_{i=1}^m Y_i$, and $Y_i = N_i \bar{Y}_i$ is the total of small area i. Let $\hat{\bar{Y}}_i$ be the model-based small area estimator of $\bar{Y}_i$. It is desirable to ensure that the aggregated values of $\hat{\bar{Y}}_i$, agree with the reliable estimator $\hat{Y}_w$. The small area means estimators $\hat{\bar{Y}}_i$, $i = 1, \ldots, m$, are said to be benchmarked to $\hat{Y}_w$ if

$$\sum_{i=1}^m N_i \hat{\bar{Y}}_i = \hat{Y}_w. \tag{3.1}$$

Let $\hat{Y}_w$ be a GREG estimator with weights calibrated at the population level on a vector of auxiliary variables $\mathbf{x}_{ij}^*$. This estimator is analogous to the combined regression estimator if one views the small areas as strata. The vector of auxiliary variables $\mathbf{x}_{ij}^*$ may or may not be the same as $\mathbf{x}_{ij}$. We distinguish two cases in this context: $\mathbf{x}_{ij} \subseteq \mathbf{x}_{ij}^*$ and $\mathbf{x}_{ij} \not\subset \mathbf{x}_{ij}^*$. The first case, $\mathbf{x}_{ij} \subseteq \mathbf{x}_{ij}^*$, implies that all the components of $\mathbf{x}_{ij}$ also belong to $\mathbf{x}_{ij}^*$, and that $\mathbf{x}_{ij}^*$ may or may not have additional components that are different from those contained in $\mathbf{x}_{ij}$. The second case, $\mathbf{x}_{ij} \not\subset \mathbf{x}_{ij}^*$, implies that some of the components of $\mathbf{x}_{ij}$ do not appear in $\mathbf{x}_{ij}^*$. We assume that the first component of both vectors $\mathbf{x}_{ij}$ and $\mathbf{x}_{ij}^*$ are equal to one, as they represent an intercept term.

For a given sample s, auxiliary data $\mathbf{x}_{ij}^*$ and basic design weights $d_{ij} = 1/\pi_{ij}$, the GREG estimator of the population total Y is given by

$$\hat{Y}^{\mathrm{GREG}} = \sum_{i=1}^m \sum_{j \in s_i} w_{ij}^{\mathrm{GREG}} y_{ij},$$

where the GREG weights w_{ij}^{GREG} are given by

$$w_{ij}^{\mathrm{GREG}} = d_{ij} \left(1 + \left(\mathbf{X}^* - \hat{\mathbf{X}}^{*\mathrm{HT}}\right)^T \left(\sum_{i=1}^m \sum_{j \in s_i} d_{ij} \mathbf{x}_{ij}^* \mathbf{x}_{ij}^{*T}\right)^{-1} \mathbf{x}_{ij}^*\right). \tag{3.2}$$

In equation (3.2), $\mathbf{X}^* = \sum_{i=1}^m \mathbf{X}_i^*$, where $\mathbf{X}_i^* = \sum_{j=1}^{N_i} \mathbf{x}_{ij}^*$ represents the known small area total, whereas $\hat{\mathbf{X}}^{*\mathrm{HT}} = \sum_{i=1}^m \hat{\mathbf{X}}_i^{*\mathrm{HT}}$ and $\hat{\mathbf{X}}_i^{*\mathrm{HT}} = \sum_{j \in s_i} d_{ij} \mathbf{x}_{ij}^*$ represent respectively the direct design-based Horvitz-Thompson estimators of $\mathbf{X}^*$ and $\mathbf{X}_i^*$. Note that

$$\sum_{i=1}^{m}\sum_{j\in s_i} w_{ij}^{\mathrm{GREG}}\mathbf{x}_{ij}^{*} = \mathbf{X}^{*}. \tag{3.3}$$

Using the GREG weights w_{ij}^{GREG}, estimators of N_i and $\mathbf{X}_i$ are given by

$$\hat{N}_i^{\mathrm{GREG}} = \sum_{j\in s_i} w_{ij}^{\mathrm{GREG}} \quad \text{and} \quad \hat{\mathbf{X}}_i^{\mathrm{GREG}} = \sum_{j\in s_i} w_{ij}^{\mathrm{GREG}}\mathbf{x}_{ij}. \tag{3.4}$$

The small area estimates $\hat{\bar{Y}}_i^{\mathrm{EBLUP}}$ and $\hat{\bar{Y}}_i^{\mathrm{YR}}$ given respectively by (2.11) and (2.13), do not satisfy the benchmarking equation (3.1) for $\hat{Y}_w = \hat{Y}^{\mathrm{GREG}}$: that is the total estimates $\hat{Y}^{\mathrm{EBLUP}} = \sum_{i=1}^{m} N_i\hat{\bar{Y}}_i^{\mathrm{EBLUP}}$, and $\hat{Y}^{\mathrm{YR}} = \sum_{i=1}^{m} N_i\hat{\bar{Y}}_i^{\mathrm{YR}}$ do not match the GREG estimator $\hat{Y}^{\mathrm{GREG}}$. We need to adjust $\hat{\bar{Y}}_i^{\mathrm{EBLUP}}$ and $\hat{\bar{Y}}_i^{\mathrm{YR}}$ so that the sum of these modified small area estimators add up to $\hat{Y}^{\mathrm{GREG}}$ when they are summed over all the m small areas.

A very simple modification to the $\hat{\bar{Y}}_i^{\mathrm{EBLUP}}$'s and $\hat{\bar{Y}}_i^{\mathrm{YR}}$'s is called ratio benchmarking. It consists of multiplying each $\hat{\bar{Y}}_i^{\mathrm{EBLUP}}$ and $\hat{\bar{Y}}_i^{\mathrm{YR}}$ by the common adjustment factors $\hat{Y}^{\mathrm{GREG}}\Big/\sum_{i=1}^{m} N_i\hat{\bar{Y}}_i^{\mathrm{EBLUP}}$ and $\hat{Y}^{\mathrm{GREG}}\Big/\sum_{i=1}^{m} N_i\hat{\bar{Y}}_i^{\mathrm{YR}}$ respectively, leading to the ratio benchmarked estimators

$$\hat{\bar{Y}}_{ib}^{\mathrm{EBRat}} = \hat{\bar{Y}}_i^{\mathrm{EBLUP}}\frac{\hat{Y}^{\mathrm{GREG}}}{\sum_{i=1}^{m} N_i\hat{\bar{Y}}_i^{\mathrm{EBLUP}}} \quad \text{and} \quad \hat{\bar{Y}}_{ib}^{\mathrm{YRat}} = \hat{\bar{Y}}_i^{\mathrm{YR}}\frac{\hat{Y}^{\mathrm{GREG}}}{\sum_{i=1}^{m} N_i\hat{\bar{Y}}_i^{\mathrm{YR}}}. \tag{3.5}$$

It readily follows that both $\hat{\bar{Y}}_{ib}^{\mathrm{EBRat}}$ and $\hat{\bar{Y}}_{ib}^{\mathrm{YRat}}$ satisfy equation (3.1) with $\hat{Y}_w = \hat{Y}^{\mathrm{GREG}}$. In equation (3.5) and hereafter the subscript b denotes that the estimators are benchmarked to $\hat{Y}^{\mathrm{GREG}}$.

Note that the $\hat{\bar{Y}}_i^{\mathrm{EBLUP}}$'s and $\hat{\bar{Y}}_i^{\mathrm{YR}}$'s in equation (3.5) are multiplied by the same factor regardless of their precision and ignoring the particular small area characteristics, such as the variability of the units within a small area, or the small area sample size. Consequently, the resulting benchmarked estimators, $\hat{\bar{Y}}_{ib}^{\mathrm{EBRat}}$ and $\hat{\bar{Y}}_{ib}^{\mathrm{YRat}}$, based on this simple procedure, are just proportional modifications of estimators $\hat{\bar{Y}}_i^{\mathrm{EBLUP}}$ and $\hat{\bar{Y}}_i^{\mathrm{YR}}$ respectively, to obtain the desired concordance. This limitation can be avoided by using the small area model (2.2) to construct the benchmarked estimators.

We now proceed to show how model (2.2) can be used to obtain estimators benchmarked to $\hat{Y}^{\mathrm{GREG}}$. In Sections 3.1 and 3.2 we adapt the procedures in Stefan and Hidiroglou (2020) for obtaining benchmarked estimators to the case of non-negligible sampling rates. In Sections 3.3 and 3.4 we introduce two restricted benchmarked estimators based on the procedure proposed by Ugarte et al. (2009). The benchmarked estimators of Sections 3.1 and 3.2 rely on the assumption that $\mathbf{x}_{ij} \subseteq \mathbf{x}_{ij}^{*}$, whereas the estimators of Sections 3.3 and 3.4 can be computed for any vector $\mathbf{x}_{ij}$ or $\mathbf{x}_{ij}^{*}$.

3.1 Augmented EBLUP benchmarked estimators

The GREG weights w_{ij}^{GREG} should be used in the estimation to achieve benchmarking to $\hat{Y}^{\mathrm{GREG}}$. A possible way that w_{ij}^{GREG} can be incorporated in the estimation is by augmenting the small area model (2.2) with a suitable auxiliary variable that is a function of w_{ij}^{GREG}. This procedure is based on the augmented model approach used by Wang et al. (2008), whereby estimates obtained using the FH

area-level model could be forced to add up to specified totals. Stefan and Hidiroglou (2020) adapted the Wang et al. (2008) approach under the basic unit-level model and for negligible sampling rates. They showed that benchmarking to $\hat{Y}^{\text{GREG}}$ could be obtained by augmenting model (2.2) with the GREG weights w_{ij}^{GREG}. We extend Stefan and Hidiroglou (2020) to the case when the sampling rates are non-negligible. For this case, benchmarking to $\hat{Y}^{\text{GREG}}$ is achieved by augmenting model (2.2) with $q_{ij} = w_{ij}^{\text{GREG}} - 1$. This leads to the augmented model given by

$$y_{ij} = \mathbf{x}_{ij}^T \boldsymbol{\beta}_{1a} + q_{ij}\beta_{2a} + v_{ia} + e_{ija},\ i = 1, \ldots, m;\ j \in s_i. \tag{3.6}$$

The random effects v_{ia} are assumed to be i.i.d. $N\left(0, \sigma_{va}^2\right)$ and independent of the unit errors e_{ija}, and the e_{ija}'s are assumed to be i.i.d. $N\left(0, \sigma_{ea}^2\right)$. The EBLUP estimators of $\boldsymbol{\beta}_a = \left(\boldsymbol{\beta}_{1a}^T, \beta_{2a}\right)^T$ and v_{ia} in (3.6) are respectively denoted by $\hat{\boldsymbol{\beta}}_a = \left(\hat{\boldsymbol{\beta}}_{1a}^T, \hat{\beta}_{2a}\right)^T$ and $\hat{v}_{ia}$. We can now spell Result 1 for $\hat{\boldsymbol{\beta}}_a$ and $\hat{v}_{ia}$.

Result 1. The EBLUP estimators $\hat{\boldsymbol{\beta}}_a$ and $\hat{v}_{ia}$ based on model (3.6) obey the following equation

$$\sum_{i=1}^{m}\sum_{j\in s_i} y_{ij} + \left(\sum_{i=1}^{m} \mathbf{x}_{ir}\right)^T \hat{\boldsymbol{\beta}}_{1a} + \sum_{i=1}^{m} q_{iw}\hat{\beta}_{2a} + \sum_{i=1}^{m}\left(\hat{N}_i^{\text{GREG}} - n_i\right)\hat{v}_{ia} = \hat{Y}^{\text{GREG}}, \tag{3.7}$$

where $q_{iw} = \sum_{j\in s_i} q_{ij}^2 = \sum_{j\in s_i} (w_{ij}^{\text{GREG}} - 1)^2$.

Proof: See Appendix A.

It follows from equation (3.7) that small area estimators benchmarked to $\hat{Y}^{\text{GREG}}$ are given by

$$\hat{\bar{Y}}_{iab}^{\text{EBLUP}} = \frac{1}{N_i}\left[\sum_{j\in s_i} y_{ij} + \mathbf{x}_{ir}^T\hat{\boldsymbol{\beta}}_{1a} + q_{iw}\hat{\beta}_{2a} + \left(\hat{N}_i^{\text{GREG}} - n_i\right)\hat{v}_{ia}\right]. \tag{3.8}$$

The subscript a indicates that $\hat{\bar{Y}}_{iab}^{\text{EBLUP}}$ is based on an augmented small area model.

3.2 You-Rao benchmarked estimators

The procedure proposed by You and Rao (2002) can be used with any survey weights w_{ij}. However, there is no guarantee that the resulting YR estimator will be benchmarked to $\hat{Y}^{\text{GREG}}$. When the sampling rates are negligible, Stefan and Hidiroglou (2020) obtained benchmarked estimators with the You and Rao's (2002) procedure based on the weights $w_{ij} = w_{ij}^{\text{GREG}}$ of the GREG estimator. When the sampling rates are non-negligible, we now show that the weights $w_{ij} = w_{ij}^{\text{GREG}} - 1$ lead to YR benchmarked estimators.

Let $\hat{\boldsymbol{\beta}}^{\text{YR}}$ and $\hat{\mathbf{v}}^{\text{YR}} = \left(\hat{v}_1^{\text{YR}}, \ldots, \hat{v}_m^{\text{YR}}\right)^T$ be YR estimators of $\boldsymbol{\beta}$ and $\mathbf{v}$ respectively with w_{ij} replaced by $w_{ij}^{\text{GREG}} - 1$. Using $\hat{\boldsymbol{\beta}}^{\text{YR}}$, $\hat{v}_i^{\text{YR}}$ and the $N_i - n_i$ estimates $\hat{y}_{ij}^{\text{YR}} = \mathbf{x}_{ij}^T\hat{\boldsymbol{\beta}}^{\text{YR}} + \hat{v}_i^{\text{YR}}$ for $j \in r_i$, a YR estimator, denoted as $\hat{\bar{Y}}_i^{\text{YR}}$, can be computed with equation (2.13). However, $\hat{\bar{Y}}_i^{\text{YR}}$ is not benchmarked to $\hat{Y}^{\text{GREG}}$ even if it uses the weights $w_{ij}^{\text{GREG}} - 1$. The original YR procedure leads to a self-benchmarked estimator in a limited number of cases.

To achieve the benchmark to $\hat{Y}^{\text{GREG}}$, a YR modified estimator, denoted as $\hat{\bar{Y}}_{ib}^{\text{YR}}$, is defined as follows:

$$\hat{\bar{Y}}_{ib}^{\mathrm{YR}} = \frac{1}{N_i}\left[\sum_{j\in s_i} y_{ij} + \mathbf{x}_{ir}^T\hat{\boldsymbol{\beta}}^{\mathrm{YR}} + \left(\hat{N}_i^{\mathrm{GREG}} - n_i\right)\hat{v}_i^{\mathrm{YR}}\right]. \tag{3.9}$$

The following proves that $\hat{\bar{Y}}_{ib}^{\mathrm{YR}}$ defined by (3.9) benchmarks to $\hat{Y}^{\mathrm{GREG}}$.

Result 2. Let $\hat{\boldsymbol{\beta}}^{\mathrm{YR}}$ and $\hat{\mathbf{v}}^{\mathrm{YR}} = \left(\hat{v}_1^{\mathrm{YR}}, \ldots, \hat{v}_m^{\mathrm{YR}}\right)^T$ be respectively the YR estimators of $\boldsymbol{\beta}$ and $\mathbf{v}$, constructed with weights $w_{ij}^{\mathrm{GREG}} - 1$. Then, $\left(\hat{\boldsymbol{\beta}}^{\mathrm{YR}}, \hat{\mathbf{v}}^{\mathrm{YR}}\right)$ satisfy the following equation:

$$\sum_{i=1}^{m}\sum_{j\in s_i} y_{ij} + \sum_{i=1}^{m}\mathbf{x}_{ir}^T\hat{\boldsymbol{\beta}}^{\mathrm{YR}} + \sum_{i=1}^{m}\left(\hat{N}_i^{\mathrm{GREG}} - n_i\right)\hat{v}_i^{\mathrm{YR}} = \hat{Y}^{\mathrm{GREG}}.$$

Proof: See Appendix A.

Given $\mathbf{x}_{ij}^*$, the weights w_{ij}^{GREG} are calibrated on $\mathbf{x}_{ij}^*$ at the small area level if they satisfy the following equations

$$\sum_{j\in s_i} w_{ij}^{\mathrm{GREG}}\mathbf{x}_{ij}^* = \mathbf{X}_i^*, \quad \text{for} \quad i = 1, \ldots, m. \tag{3.10}$$

Equations (3.10) implies equation (3.3), however, the reverse is not true. If the weights w_{ij}^{GREG} satisfy (3.10), and since $\mathbf{x}_{ij} \subseteq \mathbf{x}_{ij}^*$, it follows that the weights w_{ij}^{GREG} are also calibrated on $\mathbf{x}_{ij}$ at the small area level. In turn, this implies that $\hat{N}_i^{\mathrm{GREG}} = N_i$, as we assume that vector $\mathbf{x}_{ij}$ contains the constant regressor equal to 1. It follows that $\hat{\bar{Y}}_i^{\mathrm{YR}} = \hat{\bar{Y}}_{ib}^{\mathrm{YR}}$. Thus, the YR estimator $\hat{\bar{Y}}_i^{\mathrm{YR}}$ constructed with $w_{ij}^{\mathrm{GREG}} - 1$ is self-benchmarked to $\hat{Y}^{\mathrm{GREG}}$ in the special case when the GREG weights are calibrated at the small area level (see You and Rao, 2002).

3.3 Restricted EBLUP benchmarked estimator

In Section 2 we showed that the EBLUP estimators of $(\boldsymbol{\beta}, \mathbf{v})$ can be obtained if the function ϕ defined in (2.5) is minimized with respect to $(\boldsymbol{\beta}, \mathbf{v})$. It therefore follows that an EBLUP estimator can be viewed as the solution to an unrestricted minimization problem. The idea of restricted EBLUP estimators is to obtain new estimators of $(\boldsymbol{\beta}, \mathbf{v})$ by minimizing ϕ subject to the restriction given by the benchmark condition. The procedure was used by Pfeffermann and Barnard (1991) under the FH area-level model. More recently, Ugarte et al. (2009) applied the procedure under the BHF unit-level model to obtain benchmarking to a synthetic estimator. Ugarte et al. (2009) described the restricted estimator as a generalized least squares estimator subject to a restriction by noticing that the minimization can be conducted as in the econometrics theory of regression estimation under linear constraints. We now describe the procedure in Ugarte et al. (2009).

We denote by $\hat{\boldsymbol{\beta}}^R$ and $\hat{\mathbf{v}}^R = \left(\hat{v}_1^R, \ldots, \hat{v}_m^R\right)^T$ the new restricted EBLUP estimators of $(\boldsymbol{\beta}, \mathbf{v})$. Then, the restricted EBLUP estimator of $\bar{Y}_i$, denoted as $\hat{\bar{Y}}_{ib}^{\mathrm{REBLUP}}$, is given by equation (2.4), where $\hat{y}_{ij}$ are replaced by $\hat{y}_{ij}^R = \mathbf{x}_{ij}^T\hat{\boldsymbol{\beta}}^R + \hat{v}_i^R$, for $j \in r_i$. We impose that the estimators $\hat{\bar{Y}}_{ib}^{\mathrm{REBLUP}}$, $i = 1, \ldots, m$ be benchmarked to $\hat{Y}^{\mathrm{GREG}}$, that is they satisfy equation (3.1) with $\hat{Y}_w = \hat{Y}^{\mathrm{GREG}}$. After carrying out some algebra, it can be shown that the benchmark to $\hat{Y}^{\mathrm{GREG}}$ of estimators $\hat{\bar{Y}}_{ib}^{\mathrm{REBLUP}}$, $i = 1, \ldots, m$ is equivalent to the following linear constraint equation

$$\mathbf{a}_1^T \hat{\boldsymbol{\beta}}^R + \mathbf{a}_2^T \hat{\mathbf{v}}^R = \hat{Y}_r^{\text{GREG}}, \tag{3.11}$$

where $\mathbf{a}_1 = \sum_{i=1}^m \mathbf{x}_{ir}$, $\mathbf{a}_2 = (N_1 - n_1, \ldots, N_m - n_m)^T$, $Y_r = Y - \sum_{i=1}^m \sum_{j \in s_i} y_{ij}$ is the total of non-observed y_{ij} values with $i = 1, \ldots, m$; $j \in r_i$, and $\hat{Y}_r^{\text{GREG}} = \hat{Y}^{\text{GREG}} - \sum_{i=1}^m \sum_{j \in s_i} y_{ij}$ is an estimator of Y_r based on $\hat{Y}^{\text{GREG}}$. The restricted EBLUP estimators $(\hat{\boldsymbol{\beta}}^R, \hat{\mathbf{v}}^R)$ are therefore obtained as the solution to the minimization of function ϕ given by (2.5) subject to the linear constraint (3.11).

The Lagrange multiplier method can be used to solve the constrained minimization of ϕ. After straightforward algebra, it can be shown that estimators $(\hat{\boldsymbol{\beta}}^R, \hat{\mathbf{v}}^R)$ are given by

$$\begin{pmatrix} \hat{\boldsymbol{\beta}}^R \\ \hat{\mathbf{v}}^R \end{pmatrix} = \begin{pmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{v}} \end{pmatrix} + \frac{1}{\mathbf{a}^T \hat{\mathbf{A}} \mathbf{a}} \hat{\mathbf{A}}^{-1} \mathbf{a} \left[\hat{Y}_r^{\text{GREG}} - \mathbf{a}^T \begin{pmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{v}} \end{pmatrix} \right], \tag{3.12}$$

where $(\hat{\boldsymbol{\beta}}, \hat{\mathbf{v}})$ are the (unconstrained) EBLUP estimators of $(\boldsymbol{\beta}, \mathbf{v})$, $\hat{\mathbf{A}}$ is the empirical version of matrix $\mathbf{A}$ defined in (2.7), and $\mathbf{a} = (\mathbf{a}_1^T \ \mathbf{a}_2^T)^T$. Then, using $\hat{y}_{ij}^R$ in (2.4), the estimator $\hat{\bar{Y}}_{ib}^{\text{REBLUP}}$ can be rewritten as

$$\hat{\bar{Y}}_{ib}^{\text{REBLUP}} = \frac{1}{N_i} \left[\sum_{j \in s_i} y_{ij} + \mathbf{x}_{ir}^T \hat{\boldsymbol{\beta}}^R + (N_i - n_i) \hat{v}_i^R \right]. \tag{3.13}$$

Remark 2. The matrix $\hat{\mathbf{A}}$ does not exist for samples when $\hat{\sigma}_v^2 = 0$. In such cases, we noticed that equation (2.8) cannot be used to compute the unconstrained estimators $(\hat{\boldsymbol{\beta}}, \hat{\mathbf{v}})$. However $(\hat{\boldsymbol{\beta}}, \hat{\mathbf{v}})$ can still be computed when $\hat{\sigma}_v^2 = 0$ because the alternative equation (2.9) can be used for $(\hat{\boldsymbol{\beta}}, \hat{\mathbf{v}})$. Equation (3.12) clearly shows that the constrained $(\hat{\boldsymbol{\beta}}^R, \hat{\mathbf{v}}^R)$ cannot be computed for samples when estimator $\hat{\sigma}_v^2$ is truncated to zero, and no alternative equation exists in these cases.

It, therefore, follows that the methods of estimation for the variance components commonly used in SAE cannot be used to compute the restricted EBLUP estimator. In Section 3.4 and Appendix B we describe an alternative method that produces a strictly positive estimation of σ_v^2 that can be applied in conjunction with $(\hat{\boldsymbol{\beta}}^R, \hat{\mathbf{v}}^R)$ such that a restricted benchmarked estimator of $\bar{Y}_i$ always exists.

3.4 Restricted You-Rao benchmarked estimator

We showed in Section 2.2 that YR estimators of $\boldsymbol{\beta}$ and $\mathbf{v}$ can be obtained as a solution to mixed model equations obtained by minimizing the sample weighted function ϕ_w given by (2.14). That is, we showed that, by defining a function ϕ_w with weights $\{w_{ij}\}$, $i = 1, \ldots, m$; $j \in s_i$ and $\{\omega_i\}$, $i = 1, \ldots, m$, and then minimizing ϕ_w, we obtain the same estimators as those given by the You and Rao's (2002) procedure. We now minimize function ϕ_w under the benchmark constraint given by (3.11). The result is a restricted YR estimator that is benchmarked to $\hat{Y}^{\text{GREG}}$.

Minimization of ϕ_w given the benchmark restriction (3.11) results in estimators of $\bar{Y}_i$, $i = 1, \ldots, m$ that are guaranteed to be benchmarked for any weights that define the function ϕ_w. Thus, one may choose any set of weights w_{ij} in ϕ_w. In a limited design-based simulation study, we compared three restricted YR estimators based on three options with respect to w_{ij}: i. $w_{ij} = w_{ij}^{\text{GREG}} - 1$; ii. $w_{ij} = w_{ij}^{\text{GREG}}$ and iii. $w_{ij} = d_{ij}$. We found no significant difference between these three estimators in terms of design mean

squared error. Given this last point and that the unrestricted benchmarked YR estimators described in Section 3.2 were based on $w_{ij} = w_{ij}^{\mathrm{GREG}} - 1$, we chose to define the restricted YR estimator based on these weights.

Let ϕ_w be defined in terms of $w_{ij} = w_{ij}^{\mathrm{GREG}} - 1$ and $\omega_i = \sum_{j \in s_i} w_{ij}^2 \Big/ \sum_{j \in s_i} w_{ij}$. Minimization of ϕ_w with respect to $(\boldsymbol{\beta}, \mathbf{v})$ subject to the benchmark constraint (3.11) results in the restricted YR estimators of $(\boldsymbol{\beta}, \mathbf{v})$, denoted as $\left(\hat{\boldsymbol{\beta}}^{\mathrm{RYR}}, \hat{\mathbf{v}}^{\mathrm{RYR}}\right)$. They are given by:

$$\begin{pmatrix} \hat{\boldsymbol{\beta}}^{\mathrm{RYR}} \\ \hat{\mathbf{v}}^{\mathrm{RYR}} \end{pmatrix} = \begin{pmatrix} \hat{\boldsymbol{\beta}}^{\mathrm{YR}} \\ \hat{\mathbf{v}}^{\mathrm{YR}} \end{pmatrix} + \frac{1}{\mathbf{a}^T \hat{\mathbf{A}}_w \mathbf{a}} \hat{\mathbf{A}}_w^{-1} \mathbf{a} \left[\hat{Y}_r^{\mathrm{GREG}} - \mathbf{a}^T \begin{pmatrix} \hat{\boldsymbol{\beta}}^{\mathrm{YR}} \\ \hat{\mathbf{v}}^{\mathrm{YR}} \end{pmatrix} \right], \tag{3.14}$$

where estimators $\left(\hat{\boldsymbol{\beta}}^{\mathrm{YR}}, \hat{\mathbf{v}}^{\mathrm{YR}}\right)$ are given by (2.15), and $\hat{\mathbf{A}}_w$ is the empirical version of $\mathbf{A}_w$ given by (2.16). Using $\hat{\boldsymbol{\beta}}^{\mathrm{RYR}}$ and $\hat{v}_i^{\mathrm{RYR}}$ of $\hat{\mathbf{v}}^{\mathrm{RYR}} = \left(\hat{v}_1^{\mathrm{RYR}}, \ldots, \hat{v}_m^{\mathrm{RYR}}\right)^T$, restricted YR estimates $\hat{y}_{ij}^{\mathrm{RYR}} = \mathbf{x}_{ij}^T \hat{\boldsymbol{\beta}}^{\mathrm{RYR}} + \hat{v}_i^{\mathrm{RYR}}$ of unobserved y_{ij} for $j \in r_i$ are then used to compute a benchmarked restricted YR estimator:

$$\hat{\bar{Y}}_{ib}^{\mathrm{RYR}} = \frac{1}{N_i} \left[\sum_{j \in s_i} y_{ij} + \mathbf{x}_{ir}^T \hat{\boldsymbol{\beta}}^{\mathrm{RYR}} + (N_i - n_i)\, \hat{v}_i^{\mathrm{RYR}} \right]. \tag{3.15}$$

As in the case of the restricted EBLUP estimator, the estimators $\left(\hat{\boldsymbol{\beta}}^{\mathrm{RYR}}, \hat{\mathbf{v}}^{\mathrm{RYR}}\right)$ given by (3.14) do not exist if FC, ML or REML results in a truncated estimate for σ_v^2. Consequently, $\bar{Y}_i$ can only be estimated by $\hat{\bar{Y}}_{ib}^{\mathrm{RYR}}$ with a method of estimation for the variance components that always leads to strictly positive estimates for σ_v^2.

A null estimate of σ_v^2 poses no problem in computing EBLUP and YR estimators. However, we noticed that the restricted EBLUP and the restricted YR estimators cannot be computed if $\hat{\sigma}_v^2 = 0$. In order to get around this problem, we use a method proposed by Moghtased-Azar, Tehranchi and Amiri-Simkooei (2014) that guarantees that the estimator of σ_v^2 will be strictly positive. This method is based on the concept of a re-parameterized restricted maximum likelihood estimation (reREML). Their idea is to use functions whose range is the set of all positive real numbers, namely positive-valued functions (PVFs), for unknown variance components in the stochastic model instead of using variance components themselves. Their numerical results showed the successful estimation of non-negativity estimation of variance components (as positive values) as well as covariance components (as negative or positive values).

We used a Fisher-scoring algorithm to obtain iteratively the reREML estimates of the variance components of the basic unit-level model given by (2.2) (see Appendix B for details). We also carried out a small simulation and found out that for area sample sizes equal to or larger than 3, the Fisher-scoring algorithm converged in less than 15 iterations. When we only considered the samples that produced a null estimate $\hat{\sigma}_v^2 = 0$, we observed that the algorithm converged even faster (see Figure 4.1 in Section 4).

4. Simulation study

We report in this section the results of a design-based simulation study as it is in line with measures that are computed by the National Statistical Offices. A design-based study is one where a fixed finite

population is first generated using an assumed model, and then for each simulation run, a sample is drawn employing the fixed finite population. The aim of the simulation study is to evaluate the properties of the benchmarked estimators described in Section 3 in terms of design bias and design mean squared error. We considered two scenarios: *Scenario 1* corresponds to the case of correct modeling, whereas *Scenario 2* corresponds to the case of incorrect modelling. Model diagnostics such as those given in Rao and Molina (2015, pages 114-118), can be used to test whether the models are correct or not. Such model diagnostics include residual analysis to detect departures from the assumed model, selection of auxiliary variables for the model, and case-deletion diagnostics to detect influential observations.

4.1 Simulation set-up for generating the finite populations

For each scenario, we considered five populations. Each population had $m = 30$ small areas, with $N_i = 100$ population units within each small area. The populations corresponding to *Scenario 1* were created using the following model

$$y_{ij} = \beta_0 + x_{ij}\beta_1 + v_i + e_{ij},\ i = 1, \ldots, m;\ j = 1, \ldots, N_i, \tag{4.1}$$

where $\beta_0 = 10$ and $\beta_1 = 5$. For generating the populations in *Scenario 2*, we split the 30 small areas into three equal groups of small areas, denoted as G_ℓ, for $\ell = 1, 2, 3$. The first group G_1 contains areas $i = 1, \ldots, 10$, the second group G_2 contains areas $i = 11, \ldots, 20$, and the third group G_3 contains areas $i = 21, \ldots, 30$. The model within a given group is given by

$$y_{ij} = \beta_{0,\ell} + x_{ij}\beta_{1,\ell} + v_i + e_{ij},\ i \in G_\ell;\ j = 1, \ldots, N_i, \tag{4.2}$$

where $(\beta_{0,1} = 10,\ \beta_{1,1} = 1)$ for areas $i \in G_1$, $(\beta_{0,2} = 20,\ \beta_{1,2} = 5)$ for areas $i \in G_2$, and $(\beta_{0,3} = 30,\ \beta_{1,3} = 10)$ for areas $i \in G_3$. Both (4.1) and (4.2) use the auxiliary variable $\mathbf{x}_{ij} = (1,\ x_{ij})^T$ whose values $x_{ij},\ j = 1, \ldots, N_i$ were generated from an exponential distribution with mean equal to 5 and variance equal to 25.

The random components in (4.1) and (4.2) were generated from the normal distributions $v_i \sim N(0, \sigma_v^2)$ and $e_{ij} \sim N(0, \sigma_e^2)$. The five populations corresponding to *Scenario 1*, denoted as A1, B1, C1, D1 and E1, were generated based on (4.1) and the following variance parameters doublets: i. $(\sigma_{Av}^2 = 0.2,\ \sigma_e^2 = 20)$ for population A1; ii. $(\sigma_{Bv}^2 = 1,\ \sigma_e^2 = 20)$ for population B1; iii. $(\sigma_{Cv}^2 = 2,\ \sigma_e^2 = 20)$ for population C1; iv. $(\sigma_{Dv}^2 = 4,\ \sigma_e^2 = 20)$ for population D1; and $(\sigma_{Ev}^2 = 20,\ \sigma_e^2 = 20)$ for population E1. Note that, for populations A1 through E1, the value of σ_e^2 is kept fixed, whereas the values for σ_v^2 vary. The σ_v^2's are chosen to obtain the following variance ratios $\delta = \sigma_v^2 / \sigma_e^2$ as 0.01, 0.05, 0.1, 0.2 and 1. The five populations in *Scenario 2*, denoted as A2, B2, C2, D2 and E2, were generated based on (4.2) with the same variance parameters doublets as for *Scenario 1*.

A stratified sampling design was used by drawing independent probability proportional to size samples (pps) of size n_i within the i^{th} small area. The small area sample sizes were taken $n_i = 3$ for $i = 1, \ldots, m$. The selection probabilities were computed as $p_{ij} = b_{ij} \Big/ \sum_{j=1}^{N_i} b_{ij}$, where the size measures

are $b_{ij} = x_{ij}$. We used Conditional Poisson Sampling (CPS) to select the pps samples within each small area (see Tillé (2006), Chapter 5). The basic design weights are given by $d_{ij} = 1/(n_i p_{ij})$.

In *Scenario 1*, we fitted the nested regression model (4.1) and its augmented version to pps sampling data selected from one of the five populations generated with model (4.1). This scenario represents correct modeling as the model fitted and the model used to generate the finite population coincide. In *Scenario 2*, we fitted the nested regression model (4.1) and its augmented version to pps sampling data selected from one of the five populations generated with model (4.2). This scenario represents incorrect modeling as the model fitted and the model used to generate the finite population do not coincide.

We selected $G =$ 30,000 stratified pps samples from each of the ten finite populations: populations A1 to E1 corresponding to *Scenario 1*, and populations A2 to E2 corresponding to *Scenario 2*. For $g = 1, \ldots, G$ let $(\hat{\sigma}_v^{2\mathrm{RE}(g)}, \hat{\sigma}_e^{2\mathrm{RE}(g)})$ and $(\hat{\sigma}_v^{2\mathrm{reRE}(g)}, \hat{\sigma}_e^{2\mathrm{reRE}(g)})$ denote respectively the estimates of (σ_v^2, σ_e^2) given by the truncated REML method and its re-parameterized version, that correspond to the g^{th} sample. The starting values in equation (B.2) were $\alpha_1^{(0)} = \log(0.1 + \hat{\sigma}_v^{2\mathrm{RE}(g)})$ and $\alpha_2^{(0)} = \log(\hat{\sigma}_e^{2\mathrm{RE}(g)})$. Equation (B.2) reached convergence in less than 15 iterations for all the populations and both scenarios. Based on the G simulated samples selected in each of the five populations corresponding to *Scenario 1*, we computed the Monte Carlo value of the probability to obtain a zero truncated REML estimate for σ_v^2 as

$$P_{\mathrm{MC}}(\hat{\sigma}_v^{2\mathrm{RE}} = 0) = \frac{1}{G}\sum_{g=1}^{G} I(\hat{\sigma}_v^{2\mathrm{RE}(g)} = 0),$$

where $I(A)$ is an indicator function with value 1 if condition A holds, and 0 otherwise.

Table 4.1 displays the Monte Carlo values of the probability to get a zero estimate for $\hat{\sigma}_v^{2\mathrm{RE}}$. It can be seen that the simulated probability $P_{\mathrm{MC}}(\hat{\sigma}_v^{2\mathrm{RE}} = 0)$ can be as high as 0.47 for $\delta = 0.01$. As δ increases, this empirical probability decreases. Table 4.1 clearly shows that estimates $(\hat{\sigma}_v^{2\mathrm{RE}}, \hat{\sigma}_e^{2\mathrm{RE}})$ cannot be used to compute the restricted EBLUP and YR estimators for samples selected in populations A1, B1, C1 and D1.

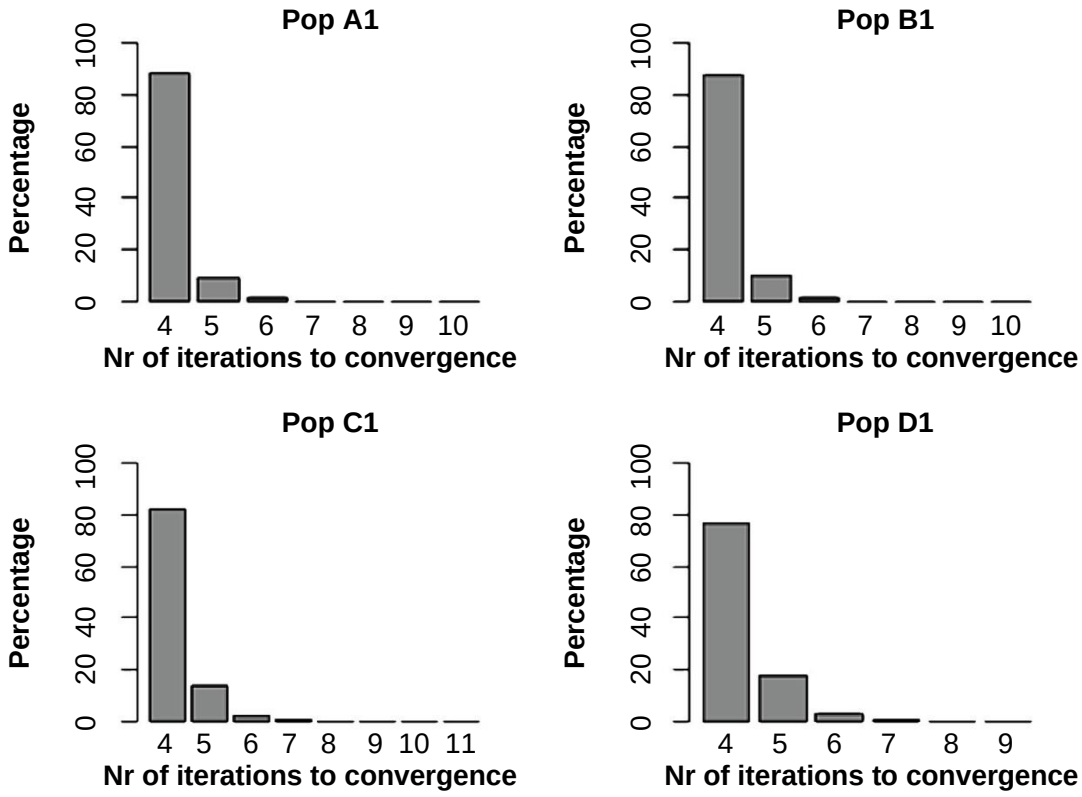
Table 4.1
Values of $P(\hat{\sigma}_v^{2\mathrm{RE}} = 0)$: *Scenario 1*

	Pop A1 $\delta = 0.01$	**Pop B1** $\delta = 0.05$	**Pop C1** $\delta = 0.1$	**Pop D1** $\delta = 0.2$	**Pop E1** $\delta = 1$
$P_{\mathrm{MC}}(\hat{\sigma}_v^{2\mathrm{RE}} = 0)$	0.47	0.40	0.21	0.06	0.00

Figure 4.1 displays the number of iterations to convergence of the Fisher-scoring algorithm for the estimate $\hat{\sigma}_v^{2\mathrm{reRE}}$ of σ_v^2. The algorithm stops when the value of $|\hat{\sigma}_v^{2\mathrm{reRE}(r+1)} - \hat{\sigma}_v^{2\mathrm{reRE}(r)}|$ is less than 10^{-5}, where $\hat{\sigma}_v^{2\mathrm{reRE}(r)}$ represents the r^{th} iteration computed with equation (B.2) in Appendix B. The percentages of Figure 4.1 are based only on samples with a truncated REML estimate of σ_v^2, that is $\hat{\sigma}_v^{2\mathrm{RE}} = 0$. We only considered populations A1, B1, C1 and D1, as these four populations have non-negligible

probabilities for $\hat{\sigma}_v^{2\text{RE}}$ to be null. Figure 4.1 clearly shows that the convergence is attained in a maximum of 11 iterations.

Figure 4.1 Percentage of iterations to convergence in samples with $\hat{\sigma}_v^{2\text{RE}} = 0$.



4.2 Comparison between the benchmarked estimators

The aim of the simulation study is to compare the benchmarked estimators described in Section 3 in terms of design bias and design mean squared error. We used both scenarios as we wanted to check how benchmarking protects against incorrect modeling. Furthermore, we considered the benchmark to two GREG estimators: $\hat{Y}_1^{\text{GREG}}$ and $\hat{Y}_2^{\text{GREG}}$. Estimator $\hat{Y}_1^{\text{GREG}}$ has weights given by (3.2) calibrated on the auxiliary vector $\mathbf{x}_{ij} = (1, x_{ij})$ associated with the small area model. It follows that estimator $\hat{Y}_1^{\text{GREG}}$ corresponds to the case $\mathbf{x}_{ij} \subseteq \mathbf{x}_{ij}^*$. The second GREG estimator $\hat{Y}_2^{\text{GREG}}$ has weights given by (3.2) based on auxiliary vector $\mathbf{x}_{ij}^* = (1, x_{ij}^*)$, where the values x_{ij}^*, $j = 1, \ldots, N_i$ were generated from an exponential distribution with mean equal to 5 and variance equal to 25, and independently of the values x_{ij}, $j = 1, \ldots, N_i$. It follows that estimator $\hat{Y}_2^{\text{GREG}}$ corresponds to the case $\mathbf{x}_{ij} \not\subset \mathbf{x}_{ij}^*$, since the auxiliary variable x_{ij} associated with the unit-level model (4.1) do not belong to vector $\mathbf{x}_{ij}^*$ used to obtain the weights associated with $\hat{Y}_2^{\text{GREG}}$.

For a fixed finite population, let $\bar{Y}_i$ be the mean of the small area i and $\hat{\bar{Y}}_i$ a generic estimator of $\bar{Y}_i$. We denote by $\hat{\bar{Y}}_i^{(g)}$ the value of $\hat{\bar{Y}}_i$ based on the g^{th} simulated sample, for $g = 1, \ldots, G$. The estimators described in Section 3 respect the benchmark property regardless of the method used to estimate the variance components. Since the restricted benchmarked estimators are based on estimates $\left(\hat{\sigma}_v^{2\text{reRE}(g)}, \hat{\sigma}_e^{2\text{reRE}(g)}\right)$, we decided to use reREML for computing $\hat{\bar{Y}}_i^{(g)}$ for each estimator $\hat{\bar{Y}}_i$ evaluated in this simulation study.

We considered the following performance measures:

Average Absolute Relative Bias

$$\overline{\text{ARB}} = \frac{1}{m}\sum_{i=1}^{m}\text{ARB}_i \quad \text{with} \quad \text{ARB}_i = \left|\frac{1}{G}\sum_{g=1}^{G}\frac{\hat{\bar{Y}}_i^{(g)}}{\bar{Y}_i} - 1\right|$$

Average Relative Root Mean Squared Error

$$\overline{\text{RRMSE}} = \frac{1}{m}\sum_{i=1}^{m}\text{RRMSE}_i \quad \text{with} \quad \text{RRMSE}_i = \sqrt{\frac{1}{G}\sum_{g=1}^{G}\left(\frac{\hat{\bar{Y}}_i^{(g)}}{\bar{Y}_i} - 1\right)^2}.$$

This portion of the simulation is summarized in four tables. We provide the results separately for *Scenarios 1* and *2*. The results for the case when the benchmarking is to $\hat{Y}_1^{\text{GREG}}$ (the case $\mathbf{x}_{ij} \subseteq \mathbf{x}_{ij}^*$) are summarized in Tables 4.2 (*Scenario 1*) and 4.3 (*Scenario 2*). Those for the case when the benchmarking is to $\hat{Y}_2^{\text{GREG}}$ (the case $\mathbf{x}_{ij} \not\subset \mathbf{x}_{ij}^*$) are summarized in Tables 4.4 (*Scenario 1*) and 4.5 (*Scenario 2*).

Benchmarking to $\hat{Y}_1^{\text{GREG}}$ *(the case* $\mathbf{x}_{ij} \subseteq \mathbf{x}_{ij}^*$*)*

We computed the $\overline{\text{ARB}}$ and $\overline{\text{RRMSE}}$ for two non benchmarked estimators, $\hat{\bar{Y}}_i^{\text{EBLUP}}$ and $\hat{\bar{Y}}_i^{\text{YR}}$, as well as their corresponding estimators benchmarked to $\hat{Y}_1^{\text{GREG}}$. For $\hat{\bar{Y}}_i^{\text{EBLUP}}$, we have three benchmarked estimators $\hat{\bar{Y}}_{ib}^{\text{EBRat}}$, $\hat{\bar{Y}}_{iab}^{\text{EBLUP}}$ and $\hat{\bar{Y}}_{ib}^{\text{REBLUP}}$, given respectively by equations (3.5), (3.8) and (3.13). For $\hat{\bar{Y}}_i^{\text{YR}}$, the corresponding benchmarked estimators are $\hat{\bar{Y}}_{ib}^{\text{YRat}}$, $\hat{\bar{Y}}_{ib}^{\text{YR}}$ and $\hat{\bar{Y}}_{ib}^{\text{RYR}}$, given respectively by equations (3.5), (3.9) and (3.15).

We first discuss their properties when the model is correct (*Scenario 1*). Comparing the $\overline{\text{ARB}}$'s across all the estimators in Table 4.2, we observe that there is not much difference between the estimators. The EBLUP estimators have somewhat smaller $\overline{\text{ARB}}$'s than the estimators based on the YR procedure. The benchmarked estimator $\hat{\bar{Y}}_{iab}^{\text{EBLUP}}$ has the smallest $\overline{\text{ARB}}$'s, whereas the $\overline{\text{ARB}}$'s of the benchmarked estimators $\hat{\bar{Y}}_{ib}^{\text{EBRat}}$ and $\hat{\bar{Y}}_{ib}^{\text{REBLUP}}$ are identical to those of $\hat{\bar{Y}}_i^{\text{EBLUP}}$. The $\overline{\text{ARB}}$ values associated with estimators $\hat{\bar{Y}}_i^{\text{YR}}$, $\hat{\bar{Y}}_{ib}^{\text{YRat}}$ and $\hat{\bar{Y}}_{ib}^{\text{RYR}}$ are close, whereas estimator $\hat{\bar{Y}}_{ib}^{\text{YR}}$ has a somewhat larger relative bias, especially for larger values of $\delta = \sigma_v^2/\sigma_e^2$. For all the estimators, the $\overline{\text{ARB}}$'s increase as δ increases: slight exceptions occur when $\delta = 1$.

Next, we report on the $\overline{\text{RRMSE}}$'s. As expected, the smallest $\overline{\text{RRMSE}}$'s are associated with $\hat{\bar{Y}}_i^{\text{EBLUP}}$, whereas estimator $\hat{\bar{Y}}_i^{\text{YR}}$ has somewhat larger $\overline{\text{RRMSE}}$ values due to the use of survey weights under

correct modeling. Benchmarking results in an increase of the $\overline{\text{RRMSE}}$. Note that the $\overline{\text{RRMSE}}$'s of the benchmarked estimators $\hat{\bar{Y}}_{iab}^{\text{EBLUP}}$ and $\hat{\bar{Y}}_{ib}^{\text{YR}}$ given in Sections 3.1 and 3.2 respectively, are higher than those associated with the restricted methods $\hat{\bar{Y}}_{ib}^{\text{REBLUP}}$ and $\hat{\bar{Y}}_{ib}^{\text{RYR}}$ given in Sections 3.3 and 3.4 respectively. The naïve ratio procedures $\hat{\bar{Y}}_{ib}^{\text{EBRat}}$ and $\hat{\bar{Y}}_{ib}^{\text{YRat}}$ have $\overline{\text{RRMSE}}$'s that are quite comparable to those of the benchmarked that use the restricted methods. The $\overline{\text{RRMSE}}$'s increase as δ increases.

We conclude the following in the case $\mathbf{x}_{ij} \subseteq \mathbf{x}_{ij}^{*}$ and when the small area model is correctly specified. The restricted benchmarked or ratio type estimators perform better than those that use an augmented model for EBLUP or a modified YR method. When the restricted or the ratio benchmarking techniques are used, the resulting estimators have bias values that are similar to those associated with their non benchmarked versions, whereas their mean squared error values are slightly larger than those of the non benchmarked versions. The small area estimators and the GREG estimator $\hat{Y}_1^{\text{GREG}}$ are based on the same auxiliary variables, whereas the model is correct. Consequently, $\hat{\bar{Y}}_i^{\text{EBLUP}}$ and $\hat{\bar{Y}}_i^{\text{YR}}$ do not have to be severely modified to achieve benchmarking to $\hat{Y}_1^{\text{GREG}}$.

Table 4.2
$\overline{\text{ARB}}$ (%) and $\overline{\text{RRMSE}}$ (%) for *Scenario 1*: the benchmark to $\hat{Y}_1^{\text{GREG}}$ $\left(\mathbf{x}_{ij} \subseteq \mathbf{x}_{ij}^{*}\right)$

Estimator	Measure	Pop A1 $\delta = 0.01$	Pop B1 $\delta = 0.05$	Pop C1 $\delta = 0.1$	Pop D1 $\delta = 0.2$	Pop E1 $\delta = 1$
$\hat{\bar{Y}}_i^{\text{EBLUP}}$	$\overline{\text{ARB}}$	1.1	1.9	2.3	2.7	2.6
	$\overline{\text{RRMSE}}$	2.7	3.4	3.9	4.9	6.5
$\hat{\bar{Y}}_i^{\text{YR}}$	$\overline{\text{ARB}}$	1.2	2.0	2.4	2.9	3.1
	$\overline{\text{RRMSE}}$	3.1	3.7	4.2	5.3	7.2
$\hat{\bar{Y}}_{ib}^{\text{EBRat}}$	$\overline{\text{ARB}}$	1.1	1.9	2.3	2.7	2.6
	$\overline{\text{RRMSE}}$	3.2	3.8	4.3	5.2	6.9
$\hat{\bar{Y}}_{ib}^{\text{YRat}}$	$\overline{\text{ARB}}$	1.2	2.0	2.4	2.9	3.1
	$\overline{\text{RRMSE}}$	3.1	3.7	4.3	5.3	7.4
$\hat{\bar{Y}}_{iab}^{\text{EBLUP}}$	$\overline{\text{ARB}}$	1.0	1.6	2.1	2.4	2.3
	$\overline{\text{RRMSE}}$	9.6	9.8	10.1	11.1	13.9
$\hat{\bar{Y}}_{ib}^{\text{YR}}$	$\overline{\text{ARB}}$	1.2	2.0	2.5	3.0	3.7
	$\overline{\text{RRMSE}}$	3.5	4.8	5.4	11.7	14.5
$\hat{\bar{Y}}_{ib}^{\text{REBLUP}}$	$\overline{\text{ARB}}$	1.1	1.9	2.3	2.7	2.6
	$\overline{\text{RRMSE}}$	3.2	3.8	4.3	5.3	7.0
$\hat{\bar{Y}}_{ib}^{\text{RYR}}$	$\overline{\text{ARB}}$	1.2	2.0	2.4	2.9	3.2
	$\overline{\text{RRMSE}}$	3.1	3.7	4.3	5.3	7.5

The results for not using the correct model are given in Table 4.3. The value of δ does not have much impact on the $\overline{\text{ARB}}$'s and $\overline{\text{RRMSE}}$'s across all estimators. The $\overline{\text{ARB}}$'s and $\overline{\text{RRMSE}}$'s of the EBLUP estimators, whether they are benchmarked or not, are higher than those associated with the YR estimators. It follows that if we have incorrect modeling, the use of the YR estimators is recommended. Since $\hat{Y}_1^{\text{GREG}}$

and the estimators based on the YR procedure use the same vector of auxiliary information, it follows that there is not much difference in terms of $\overline{\text{ARB}}$ and $\overline{\text{RRMSE}}$ between the non benchmarked estimator $\hat{\bar{Y}}_i^{\text{YR}}$ and its benchmarked versions, $\hat{\bar{Y}}_{ib}^{\text{YRat}}$, $\hat{\bar{Y}}_{ib}^{\text{YR}}$ and $\hat{\bar{Y}}_{ib}^{\text{RYR}}$. However, it can be noticed that the benchmarked estimator $\hat{\bar{Y}}_{ib}^{\text{YR}}$ has the smallest $\overline{\text{ARB}}$ values, whereas the restricted benchmarked estimator $\hat{\bar{Y}}_{ib}^{\text{RYR}}$ has the smallest $\overline{\text{RRMSE}}$'s.

Table 4.3
$\overline{\text{ARB}}$ (%) and $\overline{\text{RRMSE}}$ (%) for *Scenario 2*: the benchmark to $\hat{Y}_1^{\text{GREG}}$ $\left(\mathbf{x}_{ij} \subseteq \mathbf{x}_{ij}^*\right)$

Estimator	Measure	Pop A2 $\delta = 0.01$	Pop B2 $\delta = 0.05$	Pop C2 $\delta = 0.1$	Pop D2 $\delta = 0.2$	Pop E2 $\delta = 1$
$\hat{\bar{Y}}_i^{\text{EBLUP}}$	$\overline{\text{ARB}}$	42.3	42.7	43.2	43.0	41.5
	$\overline{\text{RRMSE}}$	59.8	60.5	61.1	60.6	59.0
$\hat{\bar{Y}}_i^{\text{YR}}$	$\overline{\text{ARB}}$	13.5	13.8	13.8	13.6	13.5
	$\overline{\text{RRMSE}}$	42.8	43.2	43.5	43.2	42.4
$\hat{\bar{Y}}_{ib}^{\text{EBRat}}$	$\overline{\text{ARB}}$	42.9	43.4	43.9	43.6	42.1
	$\overline{\text{RRMSE}}$	61.2	61.9	62.7	62.1	60.3
$\hat{\bar{Y}}_{ib}^{\text{YRat}}$	$\overline{\text{ARB}}$	13.8	14.1	14.1	13.9	13.8
	$\overline{\text{RRMSE}}$	43.9	44.4	44.7	44.4	43.5
$\hat{\bar{Y}}_{iab}^{\text{EBLUP}}$	$\overline{\text{ARB}}$	19.8	20.2	20.2	20.2	19.6
	$\overline{\text{RRMSE}}$	66.2	66.7	67.6	67.3	66.6
$\hat{\bar{Y}}_{ib}^{\text{YR}}$	$\overline{\text{ARB}}$	10.9	10.6	11.5	12.5	10.7
	$\overline{\text{RRMSE}}$	47.3	47.6	48.1	47.9	47.8
$\hat{\bar{Y}}_{ib}^{\text{REBLUP}}$	$\overline{\text{ARB}}$	41.2	41.8	41.8	41.7	40.6
	$\overline{\text{RRMSE}}$	58.2	59.0	59.1	58.9	57.4
$\hat{\bar{Y}}_{ib}^{\text{RYR}}$	$\overline{\text{ARB}}$	12.5	12.7	12.6	12.5	12.5
	$\overline{\text{RRMSE}}$	42.4	42.9	43.1	42.9	42.1

Benchmarking to $\hat{Y}_2^{\text{GREG}}$ (the case $\mathbf{x}_{ij} \not\subset \mathbf{x}_{ij}^*$)

The results of this case are given in Tables 4.4 and 4.5 for *Scenarios 1* and *2*, respectively. The weighting is with respect to w_{ij}^{GREG} given by (3.2) .We investigated the following four estimators ($\hat{\bar{Y}}_{ib}^{\text{EBRat}}$, $\hat{\bar{Y}}_{ib}^{\text{REBLUP}}$, $\hat{\bar{Y}}_{ib}^{\text{YRat}}$, and $\hat{\bar{Y}}_{ib}^{\text{RYR}}$) that are benchmarked to $\hat{Y}_2^{\text{GREG}}$. The first two estimators, $\hat{\bar{Y}}_{ib}^{\text{EBRat}}$ and $\hat{\bar{Y}}_{ib}^{\text{REBLUP}}$, are given by equations (3.5) and (3.13) respectively, while the last two, $\hat{\bar{Y}}_{ib}^{\text{YRat}}$ and $\hat{\bar{Y}}_{ib}^{\text{RYR}}$, are given by equations (3.5) and (3.15).

In Table 4.4, we summarize the average ARB and RRMSE values when the model is correct. That is, both the sample and the population data respect model (4.1). We first discuss their properties in terms of the $\overline{\text{ARB}}$'s. Comparing the $\overline{\text{ARB}}$'s across all the estimators in Table 4.4, we observe once more that, under correct modeling, the original EBLUP estimator, $\hat{\bar{Y}}_i^{\text{EBLUP}}$, has the smallest $\overline{\text{ARB}}$'s. The $\overline{\text{ARB}}$'s increase when benchmarking is required, and this is different from what we noticed from Table 4.2. There is not much difference in terms of $\overline{\text{ARB}}$ between the benchmarked estimators obtained using ratio

adjustment methods, $\hat{\bar{Y}}_{ib}^{\text{EBRat}}$ and $\hat{\bar{Y}}_{ib}^{\text{YRat}}$, and those obtained by restricted methods, $\hat{\bar{Y}}_{ib}^{\text{REBLUP}}$ and $\hat{\bar{Y}}_{ib}^{\text{RYR}}$. The $\overline{\text{ARB}}$'s increase as δ increases: slight exceptions occur when $\delta = 1$.

Next, we report on the $\overline{\text{RRMSE}}$'s. As expected, the smallest $\overline{\text{RRMSE}}$'s are associated with $\hat{\bar{Y}}_i^{\text{EBLUP}}$ which is optimal under correct modeling. Benchmarking results in an increase of $\overline{\text{RRMSE}}$. Note that the $\overline{\text{RRMSE}}$'s associated with all four benchmarking procedures in Table 4.4 are quite high compared to the $\overline{\text{RRMSE}}$'s associated with the non benchmarked estimators $\hat{\bar{Y}}_i^{\text{EBLUP}}$ and $\hat{\bar{Y}}_i^{\text{YR}}$. The estimators $\hat{\bar{Y}}_{ib}^{\text{EBRat}}$ and $\hat{\bar{Y}}_{ib}^{\text{YRat}}$ have similar efficiency, whereas $\hat{\bar{Y}}_{ib}^{\text{REBLUP}}$ and $\hat{\bar{Y}}_{ib}^{\text{RYR}}$ have $\overline{\text{RRMSE}}$ values that are somewhat larger than those of $\hat{\bar{Y}}_{ib}^{\text{EBRat}}$ and $\hat{\bar{Y}}_{ib}^{\text{YRat}}$. The $\overline{\text{RRMSE}}$'s increase as δ increases.

When $\mathbf{x}_{ij} \not\subset \mathbf{x}_{ij}^*$, there are larger differences between the small area estimators based on model (2.2) that uses the vector $\mathbf{x}_{ij}$, and the GREG estimator that uses $\mathbf{x}_{ij}^*$. Notice that we considered a somewhat extreme situation when $\mathbf{x}_{ij}$ and $\mathbf{x}_{ij}^*$ have no variable in common. It follows that the modifications needed to obtain benchmarked estimators are more accentuated in this case as compared to the case $\mathbf{x}_{ij} \subseteq \mathbf{x}_{ij}^*$. This explains why in Table 4.4 the benchmarked estimators have significantly larger $\overline{\text{ARB}}$ and $\overline{\text{RRMSE}}$ values than the estimators that are not benchmarked to $\hat{Y}_2^{\text{GREG}}$.

Table 4.4
$\overline{\text{ARB}}$ (%) and $\overline{\text{RRMSE}}$ (%) for *Scenario 1*: the benchmark to $\hat{Y}_2^{\text{GREG}}$ $\left(\mathbf{x}_{ij} \not\subset \mathbf{x}_{ij}^*\right)$

Estimator	Measure	Pop A1 $\delta = 0.01$	Pop B1 $\delta = 0.05$	Pop C1 $\delta = 0.1$	Pop D1 $\delta = 0.2$	Pop E1 $\delta = 1$
$\hat{\bar{Y}}_i^{\text{EBLUP}}$	$\overline{\text{ARB}}$	1.1	1.9	2.3	2.7	2.6
	$\overline{\text{RRMSE}}$	2.7	3.4	3.9	4.9	6.5
$\hat{\bar{Y}}_i^{\text{YR}}$	$\overline{\text{ARB}}$	1.2	2.0	2.4	2.9	3.1
	$\overline{\text{RRMSE}}$	3.1	3.7	4.2	5.3	7.2
$\hat{\bar{Y}}_{ib}^{\text{EBRat}}$	$\overline{\text{ARB}}$	4.2	4.3	4.5	4.9	4.6
	$\overline{\text{RRMSE}}$	13.0	13.2	13.5	14.0	14.6
$\hat{\bar{Y}}_{ib}^{\text{YRat}}$	$\overline{\text{ARB}}$	4.2	4.3	4.5	5.0	4.8
	$\overline{\text{RRMSE}}$	13.0	13.2	13.5	14.0	14.0
$\hat{\bar{Y}}_{ib}^{\text{REBLUP}}$	$\overline{\text{ARB}}$	4.2	4.3	4.5	5.0	4.8
	$\overline{\text{RRMSE}}$	13.1	13.3	13.5	14.1	15.0
$\hat{\bar{Y}}_{ib}^{\text{RYR}}$	$\overline{\text{ARB}}$	4.2	4.3	4.6	5.1	5.0
	$\overline{\text{RRMSE}}$	13.5	13.7	13.8	14.5	16.2

The impact of using an incorrect model is given in Table 4.5. We see that $\hat{\bar{Y}}_i^{\text{EBLUP}}$ suffers the most in terms of both $\overline{\text{ARB}}$ and $\overline{\text{RRMSE}}$ because the EBLUP procedure assumes that the model is correct. The benchmarked versions of EBLUP, $\hat{\bar{Y}}_{ib}^{\text{EBRat}}$ and $\hat{\bar{Y}}_{ib}^{\text{REBLUP}}$, also have high $\overline{\text{ARB}}$'s and $\overline{\text{RRMSE}}$'s. Although the original You and Rao (2002) estimator, $\hat{\bar{Y}}_i^{\text{YR}}$, has much smaller $\overline{\text{ARB}}$ than the EBLUP estimator, its $\overline{\text{RRMSE}}$ is fairly high. The $\overline{\text{ARB}}$ and $\overline{\text{RRMSE}}$ associated with the ratio benchmarked version of $\hat{\bar{Y}}_i^{\text{YR}}$, $\hat{\bar{Y}}_{ib}^{\text{YRat}}$, are a bit higher than those associated with $\hat{\bar{Y}}_i^{\text{YR}}$. The benchmarked YR estimator, $\hat{\bar{Y}}_{ib}^{\text{RYR}}$, which is based on the restricted procedure given in Section 3.4, has an $\overline{\text{ARB}}$ that is the smallest amongst the

estimators in Table 4.5. Due to benchmarking, its $\overline{\text{RRMSE}}$ is slightly larger than the one associated with $\hat{\bar{Y}}_i^{\text{YR}}$.

Table 4.5
$\overline{\text{ARB}}$ (%) and $\overline{\text{RRMSE}}$ (%) for *Scenario 2*: the benchmark to $\hat{Y}_2^{\text{GREG}}\left(\mathbf{x}_{ij} \not\subset \mathbf{x}_{ij}^*\right)$

Estimator	Measure	Pop A2 $\delta = 0.01$	Pop B2 $\delta = 0.05$	Pop C2 $\delta = 0.1$	Pop D2 $\delta = 0.2$	Pop E2 $\delta = 1$
$\hat{\bar{Y}}_i^{\text{EBLUP}}$	$\overline{\text{ARB}}$	42.3	42.6	43.2	43.0	41.6
	$\overline{\text{RRMSE}}$	59.8	60.4	61.1	60.7	59.1
$\hat{\bar{Y}}_i^{\text{YR}}$	$\overline{\text{ARB}}$	13.6	13.6	13.9	13.7	13.5
	$\overline{\text{RRMSE}}$	42.8	43.1	43.5	43.3	42.4
$\hat{\bar{Y}}_{ib}^{\text{EBRat}}$	$\overline{\text{ARB}}$	43.8	44.4	44.9	44.6	43.3
	$\overline{\text{RRMSE}}$	65.4	66.1	67.0	66.4	64.5
$\hat{\bar{Y}}_{ib}^{\text{YRat}}$	$\overline{\text{ARB}}$	15.0	15.2	15.6	15.2	14.9
	$\overline{\text{RRMSE}}$	47.9	48.2	48.7	48.3	47.3
$\hat{\bar{Y}}_{ib}^{\text{REBLUP}}$	$\overline{\text{ARB}}$	37.3	38.0	38.1	37.8	37.1
	$\overline{\text{RRMSE}}$	57.4	58.2	58.5	58.2	56.7
$\hat{\bar{Y}}_{ib}^{\text{RYR}}$	$\overline{\text{ARB}}$	9.9	10.1	10.4	10.0	10.1
	$\overline{\text{RRMSE}}$	43.4	43.8	44.2	43.9	43.1

5. Real data example

In this section, we compare the benchmarked estimators through a real data analysis. The data set we studied is the corn and soybean data provided by Battese et al. (1988). They considered the estimation of mean hectares of corn and soybeans per segment for twelve counties in north-central Iowa. The response variable y_{ij} is the number of hectares of corn in the j^{th} segment of the i^{th} county. The auxiliary variables, x_{1ij} and x_{2ij}, are the number of pixels classified as corn and soybeans respectively, in the j^{th} segment of the i^{th} county. We report only results for $\bar{Y}_i$, the mean number of hectares of corn per segment for county i.

Following Battese et al. (1988), we deleted the sample data from the second sample segment in Hardin county because the corn area for that segment looked erroneous. Among the twelve counties, there were three counties with a single sample segment. Following Prasad and Rao (1990), we combined these three counties into a single one, resulting in 10 counties in our data set with sample size n_i ranging from 2 to 5 in each county. The total number of segments N_i (population size) within each county ranged from 402 to 1,505. Following You and Rao (2002), we assumed simple random sampling within each county, and the basic design weight was computed as $d_{ij} = N_i / n_i$ for unit j in the i^{th} county.

We base our calculations on the unit level sampling model given by

$$y_{ij} = \beta_0 + x_{1ij}\beta_1 + x_{2ij}\beta_2 + v_i + e_{ij},\ j = 1, \ldots, n_i;\ i = 1, \ldots, 10, \tag{5.1}$$

where v_i and e_{ij} are normally distributed errors with common variances σ_v^2 and σ_e^2. We fitted model (5.1) to the sample data to obtain EBP estimates of $\boldsymbol{\beta}$ and v_i, denoted as $\hat{\boldsymbol{\beta}} = \left(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2\right)^T$ and $\hat{v}_i$, and re-parameterized REML estimates of the variance components, denoted as $\left(\hat{\sigma}_v^{2\text{reRE}}, \hat{\sigma}_e^{2\text{reRE}}\right)$. The EBLUP estimates of the model fixed effects are $\hat{\beta}_0 = 58.5$, $\hat{\beta}_1 = 0.316$ and $\hat{\beta}_2 = -0.150$, whereas the reREML estimates of the variance components are $\hat{\sigma}_v^{2\text{reRE}} = 135.6$ and $\hat{\sigma}_e^{2\text{reRE}} = 155.9$. The estimated δ is 0.869 which is close to 1. For each unit in the sample, we replicated the vector $\mathbf{x}_{ij} = \left(1, x_{1ij}, x_{2ij}\right)^T$ several times equal to $\left[d_{ij}\right]$, the closest integer to the sampling weight $d_{ij} = N_i / n_i$. Thus, we obtained a pseudo-population of x-values, denoted as $\mathbf{x}_{ij}^{\text{ps}} = \left(1, x_{1ij}^{\text{ps}}, x_{2ij}^{\text{ps}}\right)^T$, with county population size equal to $N_i^{\text{ps}} = n_i\left[N_i / n_i\right]$. The y-values of our pseudo-population, denoted as y_{ij}^{ps}, are defined as: $y_{ij}^{\text{ps}} = y_{ij}$ for $j \in s_i$, and $y_{ij}^{\text{ps}} = \hat{\beta}_0 + x_{1ij}^{\text{ps}}\hat{\beta}_1 + x_{2ij}^{\text{ps}}\hat{\beta}_2 + \hat{v}_i + e_{ij}^{\text{ps}}$ for $j \in r_i^{\text{ps}}$, where $e_{ij}^{\text{ps}} \sim N\left(0, \hat{\sigma}_e^{2\text{reRE}}\right)$ and r_i^{ps} is composed of the $N_i^{\text{ps}} - n_i$ non-observed units in the i^{th} small area. Prasad and Rao (1990) used a similar procedure to generate a pseudo-population with a larger number of counties than the data set provided by Battese et al. (1988). Their pseudo population composed of twenty counties was obtained in two steps: first, the values of the auxiliary variables associated with the original data set were duplicated; then, the values of the response variable were computed from the model, by using the duplicated x-values and the estimates of the model parameters.

Let $\bar{Y}_i = \left(N_i^{\text{ps}}\right)^{-1}\sum_{j=1}^{N_i^{\text{ps}}} y_{ij}^{\text{ps}}$ and $Y = \sum_{i=1}^{10} N_i^{\text{ps}}\bar{Y}_i$ be respectively the mean of the i^{th} small area and the total of the pseudo-population. At the population level we estimate Y by the GREG estimator $\hat{Y}^{\text{GREG}}$ based on weights given by (3.2) where the vector $\mathbf{x}_{ij}^{\text{ps}*}$ is the two-dimensional vector $\mathbf{x}_{ij}^{\text{ps}*} = \left(1, x_{1ij}^{\text{ps}}\right)^T$. It follows that $\mathbf{x}_{ij}^{\text{ps}} \not\subset \mathbf{x}_{ij}^{\text{ps}*}$ given that $\mathbf{x}_{ij}^{\text{ps}} = \left(1, x_{1ij}^{\text{ps}}, x_{2ij}^{\text{ps}}\right)^T$ and $\mathbf{x}_{ij}^{\text{ps}*} = \left(1, x_{1ij}^{\text{ps}}\right)^T$.

From the pseudo-population $\left(y_{ij}^{\text{ps}}, \mathbf{x}_{ij}^{\text{ps}}\right)$, $j = 1, \ldots, N_i^{\text{ps}}$; $i = 1, \ldots, 10$, we drew $G = 30{,}000$ stratified simple random samples without replacement of size n_i, and treating each county as a stratum. These sample sizes were equal to those of the original data set. We used the design relative bias (RB) and mean squared error (RRMSE) to evaluate the performance of six estimators: two non benchmarked estimators, $\hat{\bar{Y}}_i^{\text{EBLUP}}$ and $\hat{\bar{Y}}_i^{\text{YR}}$, and four benchmarked estimators, $\hat{\bar{Y}}_{ib}^{\text{EBRat}}$, $\hat{\bar{Y}}_{ib}^{\text{YRat}}$, $\hat{\bar{Y}}_{ib}^{\text{REBLUP}}$ and $\hat{\bar{Y}}_{ib}^{\text{RYR}}$, that can be computed in the case $\mathbf{x}_{ij}^{\text{ps}} \not\subset \mathbf{x}_{ij}^{\text{ps}*}$. Let $\hat{\bar{Y}}_i$ be a generic estimator of the i^{th} small area mean $\bar{Y}_i$, and $\hat{\bar{Y}}_i^{(g)}$ its value associated with the g^{th} sample, for $g = 1, \ldots, G$. Its RB and RRMSE values are given by

$$\text{RB}_i = \frac{1}{G}\sum_{g=1}^{G}\frac{\hat{\bar{Y}}_i^{(g)}}{\bar{Y}_i} - 1 \quad \text{and} \quad \text{RRMSE}_i = \sqrt{\frac{1}{G}\sum_{g=1}^{G}\left(\frac{\hat{\bar{Y}}_i^{(g)}}{\bar{Y}_i} - 1\right)^2}.$$

Table 5.1 reports on the design RB and RRMSE of the six estimators of $\bar{Y}_i$ for the ten counties of the pseudo population. From this example, we see that the RBs and RRMSEs are quite similar across all estimators and sample sizes. This follows because the model that generated the population data is correct, whereas both the small area model and the GREG estimator have in common the auxiliary variable equal to the number of pixels classified as corn.

Table 5.1
RB (%) and RRMSE (%): the benchmark to $\hat{Y}^{\text{GREG}}\left(\mathbf{x}_{ij}^{\text{ps}} \not\subset \mathbf{x}_{ij}^{\text{ps*}}\right)$

County	n_i	Measure	$\hat{\bar{Y}}_i^{\text{EBLUP}}$	$\hat{\bar{Y}}_i^{\text{YR}}$	$\hat{\bar{Y}}_{ib}^{\text{EBRat}}$	$\hat{\bar{Y}}_{ib}^{\text{YRat}}$	$\hat{\bar{Y}}_{ib}^{\text{REBLUP}}$	$\hat{\bar{Y}}_{ib}^{\text{RYR}}$
Cerro Hamilton Worth	3	RB	1.6	1.4	1.3	1.3	1.0	1.2
		RRMSE	5.2	5.4	5.3	5.4	5.6	5.4
Humboldt	2	RB	2.0	1.9	1.7	1.8	1.8	1.8
		RRMSE	4.5	4.5	4.5	4.5	4.4	4.5
Franklin	3	RB	-3.3	-3.4	-3.5	-3.5	-3.5	-3.5
		RRMSE	5.2	5.4	5.5	5.5	5.4	5.4
Pocahontas	3	RB	-3.1	-3.4	-3.4	-3.5	-3.3	-3.5
		RRMSE	6.2	6.5	6.4	6.6	6.4	6.6
Winnebago	3	RB	2.6	2.3	2.3	2.2	2.3	2.2
		RRMSE	5.4	5.3	5.3	5.3	5.3	5.2
Wright	3	RB	-0.4	-0.6	-0.7	-0.7	-0.6	-0.6
		RRMSE	3.7	3.8	3.9	3.9	3.8	3.9
Webster	4	RB	-2.6	-2.9	-2.9	-3.0	-2.8	-2.9
		RRMSE	5.2	5.4	5.5	5.5	5.4	5.5
Hancock	5	RB	0.9	0.7	0.6	0.6	0.8	0.7
		RRMSE	4.2	4.1	4.2	4.2	4.2	4.2
Kossuth	5	RB	3.5	3.3	3.2	3.2	3.2	3.2
		RRMSE	5.9	5.8	5.8	5.8	5.8	5.8
Hardin	5	RB	-1.5	-1.7	-1.8	-1.8	-1.7	-1.8
		RRMSE	4.2	4.3	4.4	4.5	4.3	4.4

6. Conclusion

In general, the sum of model-based small area estimates is not equal to a direct estimate obtained across the union of these small areas. The weight that is associated with the direct estimator can be the sampling weight or one obtained as a result of using the GREG estimator. The auxiliary data that are used to obtain the GREG and the unit-level small area estimates may not necessarily coincide. In this paper, we have suggested several benchmarking procedures for two well-known small area estimators (EBLUP and YR) that are based on the unit level model. We considered the case when the sampling rates are not negligible, and that the sample design is ignorable. In the event that it is deemed that the sample design is not ignorable for some of the survey items, the auxiliary data vector $\mathbf{x}_{ij}$ in model (2.2) could be augmented by including an additional variable g_{ij} specified function of the survey weights to offset the potential bias of the EBLUP or YR estimators. Verret et al. (2015) proposed a number of choices for g_{ij} that included the survey weight w_{ij}. In the case of the EBLUP estimator, benchmarking is achieved by adding the variable $q_{ij} = w_{ij}^{\text{GREG}} - 1$. Since q_{ij} should be highly correlated to w_{ij}, the suggested procedure for benchmarking EBLUP should provide good protection against possible non ignorable sampling. The simulations in Verret et al. (2015) illustrated that the YR procedure, on its own, provides good protection as well against possible non ignorable sampling. Their simulation also showed that further protection can be obtained by their setting g_{ij} equal to $n_i w_{ij}$.

We extended the benchmarking procedures in Stefan and Hidiroglou (2020) to the case of non-negligible sampling rates within each small area. These procedures are based on estimators that were initially developed by Battese et al. (1988) (EBLUP estimator), and You and Rao (2002) (YR estimator)

when the sampling rates within each small area are negligible. Ugarte et al. (2009) proposed a different benchmarked estimator which is a restricted EBLUP estimator. We extended the procedure in Ugarte et al. (2009) to obtain a benchmarked estimator that incorporates the survey weights, and that is essentially a restricted YR estimator. We also considered two benchmarked estimators based on simple ratio adjustments applied on the EBLUP and YR estimators respectively. We carried out a simulation study to compare the properties of these six benchmarked estimators.

If the auxiliary data used to estimate the small area means are the same as those used in the GREG, and if the model is correct, the restricted procedure in Ugarte et al. (2009) and the ratio adjusted EBLUP estimator will have the smaller $\overline{\text{ARB}}$'s and $\overline{\text{RRMSE}}$'s. On the other hand, if the model is incorrect and the auxiliary data are the same ones, the YR estimator based on Stefan and Hidiroglou (2020) procedure, adapted to non-negligible sampling rates, has the smallest $\overline{\text{ARB}}$'s, whereas the restricted YR estimator has the smallest $\overline{\text{RRMSE}}$'s. On the other hand, if the auxiliary data used to estimate the small area means are not the same as those used in the GREG, we come to the following conclusions. The restricted EBLUP and the ratio adjusted EBLUP estimators are the benchmarked estimators that have the smallest $\overline{\text{ARB}}$'s and $\overline{\text{RRMSE}}$'s if the model is correct. If the model is not correct, the restricted YR estimator is the preferred choice both in terms of $\overline{\text{ARB}}$ and $\overline{\text{RRMSE}}$.

Benchmarking should be based on the EBLUP procedure if the linear mixed effects model is appropriate. If the linear model and the benchmark (the GREG estimator) have in common a large amount of auxiliary information, the benchmarked estimators are similar to their non benchmarked versions, otherwise the loss of efficiency due to benchmarking may be important. If the model is not correct, the YR procedure should be used to achieve benchmarking. In this case, benchmarking may bring about important gains in terms of $\overline{\text{ARB}}$ and $\overline{\text{RRMSE}}$, especially if the small area model and the GREG estimator share a small number of auxiliary variables. The finite populations associated with incorrect modeling were generated based on model (4.2), with mean function incorrectly specified. However, there are many ways in which a model may be wrong, and the conclusions associated with these cases may be different.

Acknowledgements

We would like to thank the two anonymous referees and the associate editor for their constructive suggestions.

Appendix A

Proof of Result 1. The EBLUP estimators $\hat{\boldsymbol{\beta}}_a = \left(\left(\hat{\boldsymbol{\beta}}_{1a} \right)^T, \hat{\beta}_{2a} \right)^T$ and $\hat{v}_{ia}$, that are based on model (3.6), satisfy the equation

$$\sum_{i=1}^{m} \sum_{j \in s_i} \begin{pmatrix} \mathbf{x}_{ij} \\ q_{ij} \end{pmatrix} \left(y_{ij} - \mathbf{x}_{ij}^T \hat{\boldsymbol{\beta}}_{1a} - q_{ij} \hat{\beta}_{2a} - \hat{v}_{ia} \right) = 0. \tag{A.1}$$

Equation (A.1) has the form of equation (2.10) and corresponds to augmented model (3.6). Expanding the second equation in (A.1), we obtain that

$$\sum_{i=1}^{m}\sum_{j\in s_i} q_{ij}\mathbf{x}_{ij}^T\hat{\boldsymbol{\beta}}_{1a} + \sum_{i=1}^{m}\sum_{j\in s_i} q_{ij}^2\hat{\beta}_{2a} + \sum_{i=1}^{m}\sum_{j\in s_i} q_{ij}\hat{v}_{ia} = \sum_{i=1}^{m}\sum_{j\in s_i} q_{ij}y_{ij}. \tag{A.2}$$

The variable q_{ij} is defined as $q_{ij} = w_{ij}^{\mathrm{GREG}} - 1$. The right-hand side of (A.2) is

$$\sum_{i=1}^{m}\sum_{j\in s_i} q_{ij}y_{ij} = \sum_{i=1}^{m}\sum_{j\in s_i}\left(w_{ij}^{\mathrm{GREG}} - 1\right) y_{ij} = \hat{Y}^{\mathrm{GREG}} - \sum_{i=1}^{m}\sum_{j\in s_i} y_{ij}. \tag{A.3}$$

The sums that appear on the left-hand side of (A.2) are given by

$$\sum_{i=1}^{m}\sum_{j\in s_i} q_{ij}\mathbf{x}_{ij}^T = \sum_{i=1}^{m}\sum_{j\in s_i}\left(w_{ij}^{\mathrm{GREG}} - 1\right)\mathbf{x}_{ij}^T = \left(\sum_{i=1}^{m}\mathbf{X}_i - \sum_{j\in s_i}\mathbf{x}_{ij}\right)^T = \left(\sum_{i=1}^{m}\mathbf{x}_{ir}\right)^T, \tag{A.4}$$

$$\sum_{i=1}^{m}\sum_{j\in s_i} q_{ij}^2 = \sum_{i=1}^{m} q_{iw}, \tag{A.5}$$

$$\sum_{i=1}^{m}\sum_{j\in s_i} q_{ij} = \sum_{i=1}^{m}\sum_{j\in s_i}\left(w_{ij}^{\mathrm{GREG}} - 1\right) = \sum_{i=1}^{m}\left(\hat{N}_i^{\mathrm{GREG}} - n_i\right). \tag{A.6}$$

In establishing that last equality of (A.4), we used that $\mathbf{x}_{ij} \subseteq \mathbf{x}_{ij}^*$ and that weights w_{ij}^{GREG} satisfy equation (3.3). Result 1 follows by replacing (A.3), (A.4), (A.5) and (A.6) into (A.2).

Proof of Result 2. The survey-weighted estimating equations that defines $\hat{\boldsymbol{\beta}}^{\mathrm{YR}}$ and $\hat{\mathbf{v}}^{\mathrm{YR}}$ are given by (2.12) constructed with the weights $w_{ij}^{\mathrm{GREG}} - 1$:

$$\sum_{i=1}^{m}\sum_{j\in s_i}\left(w_{ij}^{\mathrm{GREG}} - 1\right)\mathbf{x}_{ij}\left(y_{ij} - \mathbf{x}_{ij}^T\hat{\boldsymbol{\beta}}^{\mathrm{YR}} - \hat{v}_i^{\mathrm{YR}}\right) = 0.$$

Since the first term of $\mathbf{x}_{ij}$ is one (representing an intercept), it follows that

$$\sum_{i=1}^{m}\sum_{j\in s_i}\left(w_{ij}^{\mathrm{GREG}} - 1\right)\left(y_{ij} - \mathbf{x}_{ij}^T\hat{\boldsymbol{\beta}}^{\mathrm{YR}} - \hat{v}_i^{\mathrm{YR}}\right) = 0. \tag{A.7}$$

The terms in (A.7) are given by:

$$\sum_{i=1}^{m}\sum_{j\in s_i}\left(w_{ij}^{\mathrm{GREG}} - 1\right) y_{ij} = \hat{Y}^{\mathrm{GREG}} - \sum_{i=1}^{m}\sum_{j\in s_i} y_{ij}, \tag{A.8}$$

$$\sum_{i=1}^{m}\sum_{j\in s_i}\left(w_{ij}^{\mathrm{GREG}} - 1\right)\mathbf{x}_{ij}^T\hat{\boldsymbol{\beta}}^{\mathrm{YR}} = \left(\sum_{i=1}^{m}\mathbf{X}_i - \sum_{j\in s_i}\mathbf{x}_{ij}\right)^T\hat{\boldsymbol{\beta}}^{\mathrm{YR}} = \left(\sum_{i=1}^{m}\mathbf{x}_{ir}\right)^T\hat{\boldsymbol{\beta}}^{\mathrm{YR}}, \tag{A.9}$$

and

$$\sum_{i=1}^{m}\sum_{j\in s_i}\left(w_{ij}^{\mathrm{GREG}}-1\right)\hat{v}_i^{\mathrm{YR}}=\sum_{i=1}^{m}\left(\hat{N}_i^{\mathrm{GREG}}-n_i\right)\hat{v}_i^{\mathrm{YR}}. \tag{A.10}$$

Plugging (A.8), (A.9) and (A.10) into (A.7) leads to

$$\sum_{i=1}^{m}\sum_{j\in s_i}y_{ij}+\left(\sum_{i=1}^{m}\mathbf{x}_{ri}\right)^{T}\hat{\boldsymbol{\beta}}^{\mathrm{YR}}+\sum_{i=1}^{m}\left(\hat{N}_i^{\mathrm{GREG}}-n_i\right)\hat{v}_i^{\mathrm{YR}}=\hat{Y}^{\mathrm{GREG}}. \tag{A.11}$$

Equation (A.11) proves Result 2.

Appendix B

Re-parameterized REML estimation of variance components

Let $\boldsymbol{\delta}=(\delta_1,\,\delta_2)$ be the vector of variance components, where $\delta_1=\sigma_v^2$ and $\delta_2=\sigma_e^2$. We define the vector $\boldsymbol{\alpha}=(\alpha_1,\,\alpha_2)$ such that $\sigma_v^2=e^{\alpha_1}$ and $\sigma_e^2=e^{\alpha_2}$. The restricted maximum log-likelihood function, denoted as $l(\boldsymbol{\alpha})$ is

$$l(\boldsymbol{\alpha})=l(\alpha_1,\,\alpha_2)=c-\frac{1}{2}\log|\mathbf{V}|-\frac{1}{2}\log|\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}|-\frac{1}{2}\mathbf{y}^T\mathbf{P}\mathbf{y}, \tag{B.1}$$

where c is a generic constant, $\mathbf{V}=e^{\alpha_1}\mathbf{Z}\mathbf{Z}^T+e^{\alpha_2}\mathbf{I}_n$ and $\mathbf{P}=\mathbf{V}^{-1}-\mathbf{V}^{-1}\mathbf{X}\left(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{V}^{-1}$. Notice that $\mathbf{P}\mathbf{X}=\mathbf{0}$. The solution to the maximization of $l(\boldsymbol{\alpha})$ is obtained iteratively using the Fisher-scoring algorithm by updating the following equation

$$\boldsymbol{\alpha}^{(r+1)}=\boldsymbol{\alpha}^{(r)}+\mathbf{I}\left(\boldsymbol{\alpha}^{(r)}\right)^{-1}\mathbf{s}\left(\boldsymbol{\alpha}^{(r)}\right). \tag{B.2}$$

Here, $\mathbf{s}\left(\boldsymbol{\alpha}^{(r)}\right)=\left(\partial l\left(\boldsymbol{\alpha}^{(r)}\right)/\partial\alpha_1,\,\partial l\left(\boldsymbol{\alpha}^{(r)}\right)/\partial\alpha_2\right)^T$ is the vector of first-order partial derivatives of $l(\boldsymbol{\alpha})$ with respect to $\boldsymbol{\alpha}$, and $\mathbf{I}\left(\boldsymbol{\alpha}^{(r)}\right)=\left(I_{jk}\left(\boldsymbol{\alpha}^{(r)}\right)\right)_{j,k=1,2}$ is the matrix of expected second-order derivatives of $-l(\boldsymbol{\alpha})$ with respect to $\boldsymbol{\alpha}$, where $I_{jk}\left(\boldsymbol{\alpha}^{(r)}\right)=E\left(-\partial^2 l\left(\boldsymbol{\alpha}^{(r)}\right)/\partial\alpha_j\partial\alpha_k\right)$.

Under the BHF model, the first-order partial derivatives of $l(\boldsymbol{\alpha})$ are given by

$$\frac{\partial l}{\partial\alpha_j}(\boldsymbol{\alpha})=\left[-\frac{1}{2}\mathrm{tr}\left(\mathbf{P}\mathbf{V}_{(j)}\right)+\frac{1}{2}\mathbf{y}^T\mathbf{P}\mathbf{V}_{(j)}\mathbf{P}\mathbf{y}\right]e^{\alpha_j},\ j=1,\,2, \tag{B.3}$$

where $\mathbf{V}_{(1)}=\mathbf{Z}\mathbf{Z}^T$ and $\mathbf{V}_{(2)}=\mathbf{I}_n$. The expected values of the second-order partial derivatives of $l(\boldsymbol{\alpha})$ are

$$E\left(\frac{\partial^2 l}{\partial\alpha_j\partial\alpha_k}(\boldsymbol{\alpha})\right)=-\frac{1}{2}\mathrm{tr}\left(\mathbf{P}\mathbf{V}_{(j)}\mathbf{P}\mathbf{V}_{(k)}\right)e^{\alpha_j+\alpha_k},\ j,\,k=1,\,2. \tag{B.4}$$

The re-parameterized REML estimator of $\boldsymbol{\delta}$ is obtained as

$$\hat{\boldsymbol{\delta}}^{\mathrm{reRE}}=\left(\hat{\sigma}_v^{2\mathrm{reRE}},\,\hat{\sigma}_e^{2\mathrm{reRE}}\right)=\left(e^{\hat{\alpha}_1},\,e^{\hat{\alpha}_2}\right). \tag{B.5}$$

References

Battese, G.E., Harter, R.M. and Fuller, W.A. (1988). An error component model for prediction of county crop areas using survey and satellite data. *Journal of the American Statistical Association*, 83, 28-36.

Bell, W.R., Datta, G.S. and Ghosh, M. (2013). Benchmarking small area estimators. *Biometrika*, 100, 189-202.

Datta, G.S., Ghosh, M., Steorts, R. and Maples, J. (2011). Bayesian benchmarking with applications to small area estimation. *Test*, 20, 574-588.

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

Fay, R.E., and Herriot, R.A. (1979). Estimation of income for small places: An application of James-Stein procedures to census data. *Journal of the American Statistical Association*, 74, 269-277.

Hidiroglou, M.A., and Estevao, V.M. (2016). A comparison of small area and traditional estimators via simulation. *Statistics in Transition*, 17, 133-154.

Huang, R., and Hidiroglou, M.A. (2003). Design consistent estimators for a mixed linear model on survey data. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 1897-1904.

Moghtased-Azar, K., Tehranchi, R. and Amiri-Simkooei, A.R. (2014). An alternative method for non-negative estimation of variance components. *Journal of Geodesy*, 88, 427-439.

Nandram, B., and Sayit, H. (2011). A Bayesian analysis of small area probabilities under a constraint. *Survey Methodology*, 37, 2, 137-152. Paper available at https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2011002/article/11603-eng.pdf.

Pfeffermann, D., and Barnard, C. (1991). Some new estimators for small area means with applications to the assessment of farmland values. *Journal of Business and Economic Statistics*, 9, 73-83.

Prasad, N.G.N., and Rao, J.N.K. (1990). The estimation of the mean squared error of small-area estimators. *Journal of the American Statistical Association*, 85, 163-171.

Rao, J.N.K., and Molina, I. (2015). *Small Area Estimation*. New York: John Wiley & Sons, Inc.

Särndal, C.-E., Swensson, B. and Wretman, J.H. (1989). The weighted residual technique for estimating the variance of the general regression estimator of the finite population total. *Biometrika*, 76, 527-537.

Stefan, M., and Hidiroglou, M.A. (2020). Benchmarked estimators for a small area mean under a one-fold nested regression model. *International Statistical Review*, (To appear).

Tillé, Y. (2006). *Sampling Algorithms*. New York: Springer.

Ugarte, M.D., Militino, A.F. and Goicoa, T. (2009). Benchmarked estimates in small areas using linear mixed models with restrictions. *Test*, 18, 342-364.

Verret, F., Rao, J.N.K. and Hidiroglou, M.A. (2015). Model-based small area estimation under informative sampling. *Survey Methodology*, 41, 2, 333-347. Paper available at https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2015002/article/14248-eng.pdf.

Wang, J., Fuller, W.A. and Qu, Y. (2008). Small area estimation under a restriction. *Survey Methodology*, 34, 1, 29-36. Paper available at https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2008001/article/10619-eng.pdf.

You, Y., and Rao, J.N.K. (2002). A pseudo-empirical best linear unbiased prediction approach to small area estimation using survey weights. *The Canadian Journal of Statistics*, 30, 431-439.

You, Y., Rao, J.N.K. and Dick, P. (2004). Benchmarking hierarchical Bayes small area estimators in the Canadian census undercoverage estimation. *Statistics in Transition*, 6, 631-640.

You, Y., Rao, J.N.K. and Hidiroglou, M.A. (2013). On the performance of self benchmarked small area estimators under the Fay-Herriot area level model. *Survey Methodology*, 39, 1, 217-229. Paper available at https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2013001/article/11830-eng.pdf.