Component of Statistics Canada
Catalogue no. 12-001-X Business Survey Methods Division

Article

Small area estimation under a two-level model

by Mahmoud Torabi and J.N.K. Rao

June 2008

Statistics Canada Statistique Canada

Canada

Survey Methodology, June 2008
Vol. 34, No. 1, pp. 11-17
Statistics Canada, Catalogue No. 12-001-X

Small area estimation under a two-level model

Mahmoud Torabi and J.N.K. Rao [1]

Abstract

Lehtonen and Veijanen (1999) proposed a new model-assisted generalized regression (GREG) estimator of a small area mean under a two-level model. They have shown that the proposed estimator performs better than the customary GREG estimator in terms of average absolute relative bias and average median absolute relative error. We derive the mean squared error (MSE) of the new GREG estimator under the two-level model and compare it to the MSE of the best linear unbiased prediction (BLUP) estimator. We also provide empirical results on the relative efficiency of the estimators. We show that the new GREG estimator exhibits better performance relative to the customary GREG estimator in terms of average MSE and average absolute relative error. We also show that, due to borrowing strength from related small areas, the EBLUP estimator exhibits significantly better performance relative to the customary GREG and the new GREG estimators. We provide simulation results under a model-based set-up as well as under a real finite population.

Key Words: BLUP estimator; GREG estimator; Mean squared error; Random effects; Small area means.

1. Introduction

Small area estimation has received a lot of attention in recent years due to growing demand for reliable small area statistics. Traditional area-specific direct estimators do not provide adequate precision because sample sizes in small areas are seldom large enough. This makes it necessary to employ indirect estimators that borrow strength from related areas, in particular, model-based indirect estimators. Unit level random effect models, including nested error linear regression models and two-level models, are often used in small area estimation to obtain efficient model-based estimators of small area means. Rao (2003) gives a comprehensive account of model-based small area estimation.

A two-level model is given by

$$y_{ij} = x'_{ij}\beta_i + e_{ij};$$

$$\beta_i = Z_i\beta + \nu_i,\ j = 1, ..., N_i;\ i = 1, ..., m \qquad (1)$$

where N_i is the number of units in the i^{th} area ($i = 1, ..., m$), y_{ij} is the response and x_{ij} is a $p \times 1$ vector of unit level covariates attached to the j^{th} unit in the i^{th} area. Further, Z_i is a $p \times q$ matrix of area level covariates, β is a $q \times 1$ vector of regression parameters, ν_i's are independent random vectors with mean zero and covariance Σ_ν, and e_{ij}'s are independent random variables with mean zero and variance σ_e^2 and independent of ν_i's. We can express the mean $\bar{Y}_i$ of i^{th} area as

$$\bar{Y}_i \approx \mu_i = \bar{X}'_i(Z_i\beta + \nu_i),$$

assuming N_i is large, where $\bar{X}_i$ is the known population mean of x_{ij} in the i^{th} area. The sample values $\{(y_{ij}, x_{ij}),\ j = 1, ..., n_i;\ i = 1, ..., m\}$ are assumed to obey the model (1), that is, there is no sample selection bias. The model for the sample is then given by

$$y_{ij} = x'_{ij}(Z_i\beta + \nu_i) + e_{ij},\ j = 1, ..., n_i;\ i = 1, ..., m. \qquad (2)$$

In matrix notation, (2) may be written as

$$y_i = X_i(Z_i\beta + \nu_i) + e_i,\ i = 1, ..., m$$

with $\text{Var}(y_i) = V_i = X_i\Sigma_\nu X'_i + \sigma_e^2 I_{n_i}$, where y_i is a $n_i \times 1$ vector and X_i is a $n_i \times p$ matrix. The two-level model (2) was first introduced by Moura and Holt (1999) in the context of small area estimation. This model effectively integrates the use of unit level and area level covariates into a single model, by modeling the random slopes β_i in (1) in terms of area level covariates Z_i.

Lehtonen and Veijanen (1999) proposed a model-assisted new generalized regression (GREG) estimator of a small area mean under the two-level model. Lehtonen and Veijanen (1999) showed that the new GREG estimator based on model (1) performs better than the customary GREG estimator based on a model with fixed $\beta_i = Z_i\beta$. Moura and Holt (1999) obtained the best linear unbiased prediction (BLUP) estimator of the small area mean μ_i and its MSE under the two-level model (2); see Section 2. Lehtonen, Särndal, and Veijanen (2003) studied the effect of model choice on different types of estimators (synthetic, GREG, and composite) of small area means.

In Section 3, we first derive the mean squared error (MSE) of the new GREG estimator and the customary GREG estimator (Section 2) under the two-level model (2), assuming known model parameters. We then compare the MSE of the GREG, new GREG and BLUP estimators, and obtain an explicit expression for the increase in MSE of the new GREG estimator relative to the MSE of the BLUP estimator. In Section 4, we provide empirical results on the relative efficiency of the estimators when the model parameters are estimated. We used a model-based set-up as

1. Mahmoud Torabi, Department of Pediatrics, University of Alberta, Edmonton, Alberta, T6G 2J3; J.N.K. Rao, School of Mathematics and Statistics, Carleton University, Ottawa, Ontario, K1S 5B6.

well as a real finite population for the simulation study. Finally, some concluding remarks are given in Section 5.

2. BLUP and GREG estimators

The two-level model (2) is a special case of a general linear mixed model with block diagonal covariance structure. Therefore, assuming known model parameters, we may calculate the BLUP estimator of μ_i as

$$\tilde{\mu}_i^B = \bar{X}_i'(Z_i\beta + \tilde{\nu}_i), \tag{3}$$

where

$$\tilde{\nu}_i = \Sigma_v X_i' V_i^{-1}(y_i - X_i Z_i \beta),$$

and the superscript B on $\tilde{\mu}_i$ stands for BLUP estimator (Rao 2003, page 107). Similarly, the BLUP estimator of a non-sampled observation j in i^{th} area can be written as

$$\tilde{y}_{ij} = x_{ij}'(Z_i\beta + \tilde{\nu}_i). \tag{4}$$

On the other hand, a model-assisted GREG estimator of μ_i (or $\bar{Y}_i$) is given by

$$\tilde{\mu}_i^G = \frac{1}{N_i}\left[\sum_{j=1}^{N_i}\tilde{y}_{ij} + \sum_{j=1}^{n_i} w_{ij}(y_{ij} - \tilde{y}_{ij})\right],\ i = 1, \ldots, m \tag{5}$$

where $\tilde{y}_{ij}$ is the predictor of y_{ij} under the assumed model, and w_{ij} is the survey weight which equals N_i/n_i in the case of simple random sampling (SRS) within areas. We focus on SRS within areas in this paper.

Using (5) with $\tilde{y}_{ij} = x_{ij}'Z_i\beta$ as the predictor of y_{ij} under the model (1) with fixed $\beta_i = Z_i\beta$, we can write the customary GREG estimator as

$$\tilde{\mu}_i^G = \bar{y}_i + (\bar{X}_i - \bar{x}_i)'Z_i\beta, \tag{6}$$

where the superscript G on $\tilde{\mu}_i$ denotes GREG (Särdnal, Swensson and Wretman 1992, page 225), $\bar{y}_i$ is sample mean of y_{ij} in the i^{th} area, and $\bar{x}_i$ is the sample mean of x_{ij} in the i^{th} area, respectively. Using the predictor (4) based on the two-level model (1) in (5), we get a new GREG estimator of μ_i (or $\bar{Y}_i$) as

$$\begin{aligned}\tilde{\mu}_i^{\text{LV}} &= [\bar{X}_i'(Z_i\beta + \tilde{\nu}_i) + (\bar{y}_i - \bar{x}_i'(Z_i\beta + \tilde{\nu}_i))] \\ &= \bar{y}_i + (\bar{X}_i - \bar{x}_i)'(Z_i\beta + \tilde{\nu}_i),\end{aligned} \tag{7}$$

where the superscript LV on $\tilde{\mu}_i$ denotes that it was first proposed by Lehtonen and Veijanen (1999). The estimators $\tilde{\mu}_i^B$, $\tilde{\mu}_i^G$ and $\tilde{\mu}_i^{\text{LV}}$ are linear in the y_{ij}'s and unbiased under the two-level model (1). In practice, we replace the parameters β, Σ_v and σ_e^2 in (3), (6) and (7) by suitable estimators. The resulting estimators are denoted by $\hat{\mu}_i^B$, $\hat{\mu}_i^G$ and $\hat{\mu}_i^{\text{LV}}$ respectively, where $\hat{\mu}_i^B$ is the empirical BLUP (EBLUP) estimator. Under normality assumption, $\hat{\mu}_i^B$ is the empirical best (EB) estimator. The EBLUP estimator of $\bar{Y}_i$ is given in Section 4.2.2. Note that $\hat{\mu}_i^G$ and $\hat{\mu}_i^{\text{LV}}$ are valid as estimators of $\bar{Y}_i$.

3. Mean squared error

The mean squared error (MSE) of the customary GREG estimator $\tilde{\mu}_i^G$ under the two-level model can be written as

$$\begin{aligned}\text{MSE}(\tilde{\mu}_i^G) &= E(\tilde{\mu}_i^G - \mu_i)^2 \\ &= E[\bar{y}_i + (\bar{X}_i - \bar{x}_i)'Z_i\beta - \bar{X}_i'(Z_i\beta + \nu_i)]^2 \\ &= E[(\bar{x}_i - \bar{X}_i)'\nu_i + \bar{e}_i]^2 \\ &= (\bar{x}_i - \bar{X}_i)'\Sigma_v(\bar{x}_i - \bar{X}_i) + \frac{\sigma_e^2}{n_i},\end{aligned}$$

as stated in Theorem 1.

Theorem 1. *The MSE of the GREG estimator* (*6*) *is given by*

$$\text{MSE}(\tilde{\mu}_i^G) = (\bar{x}_i - \bar{X}_i)'\Sigma_v(\bar{x}_i - \bar{X}_i) + \frac{\sigma_e^2}{n_i}. \tag{8}$$

Further, we may write the MSE of the BLUP estimator $\tilde{\mu}_i^B$ as follows:

$$\begin{aligned}\text{MSE}(\tilde{\mu}_i^B) &= E(\tilde{\mu}_i^B - \mu_i)^2 \\ &= E[\bar{X}_i'(\tilde{\nu}_i - \nu_i)]^2 \\ &= \bar{X}_i'\left(\Sigma_v - \Sigma_v X_i' V_i^{-1} X_i \Sigma_v\right)\bar{X}_i,\end{aligned}$$

as stated in Theorem 2.

Theorem 2. *The MSE of the BLUP estimator* (*3*) *is given by*

$$\text{MSE}(\tilde{\mu}_i^B) = \bar{X}_i'\left(\Sigma_v - \Sigma_v X_i' V_i^{-1} X_i \Sigma_v\right)\bar{X}_i. \tag{9}$$

Theorem 3 gives the MSE of the new GREG estimator $\tilde{\mu}_i^{\text{LV}}$.

Theorem 3. *The MSE of the new GREG estimator* (*7*) *is given by*

$$\begin{aligned}\text{MSE}(\tilde{\mu}_i^{\text{LV}}) &= \text{MSE}(\tilde{\mu}_i^B) \\ &\quad + \left\{\bar{x}_i'\Sigma_v X_i' V_i^{-1} X_i \Sigma_v \bar{x}_i - \bar{x}_i'\Sigma_v \bar{x}_i + \frac{\sigma_e^2}{n_i}\right\}.\end{aligned} \tag{10}$$

Proof of Theorem 3 is given in the Appendix.

By definition, we have $\text{MSE}(\tilde{\mu}_i^B) \le \text{MSE}(\tilde{\mu}_i^{\text{LV}})$ and (10) gives an explicit expression for the increase in MSE of $\tilde{\mu}_i^{\text{LV}}$ relative to the MSE of the BLUP estimator $\tilde{\mu}_i^B$.

4. Empirical results

4.1 Empirical comparison of MSE values

In order to study the efficiency of the new GREG estimator, we used data from Moura and Holt (1999) based on 38,740 households in the enumeration districts (small areas) in one county in Brazil. The income of household's head was treated as the response variable y. Two unit level independent variables were identified as the educational attainment of household's head (ordinal scale of 0-5) and the number of rooms in the household (1-11+). The following two-level model was assumed for this data:

$$y_{ij} = \beta_{i0} + \beta_{i1}x_{1ij} + \beta_{i2}x_{2ij} + e_{ij},$$
$$j = 1, \ldots, N_i;\ i = 1, \ldots, m \qquad (11)$$

with

$$\beta_{i0} = \beta_0 + \nu_{i0};\ \beta_{i1} = \beta_1 + \nu_{i1};\ \beta_{i2} = \beta_2 + \nu_{i2}, \quad (12)$$

where

$$\nu = (\nu_{i0}, \nu_{i1}, \nu_{i2})' \overset{\text{iid}}{\sim} N_3(0, \Sigma_\nu),\ e_{ij} \overset{\text{iid}}{\sim} N(0, \sigma_e^2)$$

and x_1 and x_2 respectively represent the number of rooms and the educational attainment of household's head (centered about their respective population means). Note that the model (12) for the random β_i-coefficients does not contain area level covariates Z.

Moura and Holt (1999) also studied another model with an area level covariate Z for modeling β_i's in (12). For this data, average number of cars per household in each area was used as a covariate z for modeling the random coefficients β_{i1} and β_{i2} corresponding to the variables x_1 and x_2, but not for the random intercept term, β_{i0}. The two-level model (11) with the area level covariate z is given by

$$y_{ij} = \beta_{i0} + \beta_{i1}x_{1ij} + \beta_{i2}x_{2ij} + \varepsilon_{ij},$$
$$j = 1, \ldots, N_i;\ i = 1, \ldots, m \qquad (13)$$

with

$$\beta_{i0} = \beta_0 + \nu_{i0};$$
$$\beta_{i1} = \beta_1 + \alpha_1 z_i + \nu_{i1};\ \beta_{i2} = \beta_2 + \alpha_2 z_i + \nu_{i2}. \quad (14)$$

Moura and Holt (1999) fitted models (11)-(12) and (13)-(14) to the full data set mentioned above. We summarize their results in Table 1.

Table 1
Parameter estimates based on Moura and Holt's (1999) data set, where σ_0^2, σ_1^2 and σ_2^2 are the diagonal elements and σ_{01}, σ_{02} and σ_{12} are the off-diagonal elements of the covariance matrix Σ_ν

Parameter	Diagonal Covariance: Model (14) with z	Diagonal Covariance: Model (12) without z	General Covariance: Model (12) without z
β_0	8.442	8.688	8.456
β_1	0.451	1.321	1.223
β_2	0.744	2.636	2.596
α_1	3.779	-	-
α_2	1.659	-	-
σ_0^2	0.745	0.637	1.385
σ_1^2	0.237	0.471	0.234
σ_2^2	0.700	1.472	0.926
σ_{01}	-	-	0.354
σ_{02}	-	-	0.492
σ_{12}	-	-	0.333
σ_e^2	44.00	44.01	47.74

The area means of x_1 and x_2 were calculated from the whole data and treated as the population means $\bar{X}_{1i}$ and $\bar{X}_{2i}$. A random subsample of 10% of the records was selected from each small area. The overall sample size is $n = 3{,}876$ and the number of small areas is $m = 140$. Using the sample values of x_1, x_2 and z and the population means $\bar{X}_{1i}$ and $\bar{X}_{2i}$, we computed $\text{MSE}(\tilde{\mu}_i^G)$, $\text{MSE}(\tilde{\mu}_i^B)$ and $\text{MSE}(\tilde{\mu}_i^{\text{LV}})$ using (8), (9) and (10) respectively, treating the estimates of regression parameters, Σ_ν and σ_e^2 in Table 1 for the full data as true values. We then calculated the average MSE values over the areas:

$$\overline{\text{MSE}}_G = \frac{1}{m}\sum_{i=1}^m \text{MSE}(\tilde{\mu}_i^G),$$
$$\overline{\text{MSE}}_B = \frac{1}{m}\sum_{i=1}^m \text{MSE}(\tilde{\mu}_i^B)$$

and

$$\overline{\text{MSE}}_{\text{LV}} = \frac{1}{m}\sum_{i=1}^m \text{MSE}(\tilde{\mu}_i^{\text{LV}}).$$

We define the relative efficiency of $\tilde{\mu}^B$ over $\tilde{\mu}^G$ as EFF_B and the relative efficiency of $\tilde{\mu}^{\text{LV}}$ over $\tilde{\mu}^G$ as EFF_{LV}, where

$$\text{EFF}_B = \frac{\overline{\text{MSE}}_G}{\overline{\text{MSE}}_B}; \qquad \text{EFF}_{\text{LV}} = \frac{\overline{\text{MSE}}_G}{\overline{\text{MSE}}_{\text{LV}}}.$$

We summarize the results in Tables 2 and 3. Tables 2 and 3 reveal that the new GREG estimator is slightly more efficient than the usual GREG estimator in terms of average MSE: $\text{EFF}_{\text{LV}} \le 112\%$. However, the new GREG estimator is substantially less efficient than the BLUP estimator, under the assumed two-level model. For example, for the model with z and diagonal covariance matrix (Table 2), $\text{EFF}_B = 292\%$ compared to $\text{EFF}_{\text{LV}} = 106\%$, and $\overline{\text{MSE}}_B = 0.62$ compared to $\overline{\text{MSE}}_{\text{LV}} = 1.72$.

Table 2
Comparison of small area estimators: relative efficiency (EFF) and average MSE (MSE) for the case of diagonal covariance matrix based on Moura and Holt's (1999) data set

Quality Measure	Model without z			Model with z		
	GREG	New GREG	BLUP	GREG	New GREG	BLUP
EFF	100%	112%	306%	100%	106%	292%
MSE	1.92	1.71	0.62	1.83	1.72	0.62

Table 3
Comparison of small area estimators: relative efficiency (EFF) and average MSE ($\overline{\text{MSE}}$) for the case of a general covariance matrix based on Moura and Holt's (1999) data set

Quality Measure	Model without z		
	GREG	New GREG	BLUP
EFF	100%	108%	253%
MSE	2.02	1.87	0.80

4.2 Simulation study

4.2.1 Simulation study under a model-based framework

In order to investigate the efficiency of the new GREG estimator with estimated model parameters, a small

simulation study based on the two-level models (11)-(12) and (13)-(14) was undertaken. We only considered a diagonal covariance structure Σ_v with diagonal elements σ_0^2, σ_1^2 and σ_2^2. We again used the data from Moura and Holt (1999). The estimates of $\beta_0, \beta_1, \beta_2, \alpha_1, \alpha_2, \sigma_0^2, \sigma_1^2, \sigma_2^2$ and σ_e^2 reported in Table 1 are treated as true values.

In our simulation study, we took (x_{1ij}, x_{2ij}, z_i) from Moura and Holt (1999) and then generated y_{ij} based on the models (11)-(12) and (13)-(14). By using the generated samples $(y_{ij}^{(b)}, x_{1ij}, x_{2ij}, z_i)$, $b = 1, ..., B = 1{,}000$, we calculated $\hat{\beta}^{(b)}$ by generalized least squares for the new GREG method as well as for the BLUP method. For the GREG method we used ordinary least squares to estimate β as $\hat{\beta}_{\text{ols}}^{(b)}$. In addition, $\hat{\Sigma}_v^{(b)}$ and $\hat{\sigma}_e^{2(b)}$ were computed based on the restricted maximum likelihood (REML) method. For each generated sample, we calculated

$$\mu_i^{(b)} = \bar{X}'_i(Z_i\beta + \nu_i^{(b)}),\ i = 1, ..., m;\ b = 1, ..., B.$$

We computed the new GREG estimator of $\mu_i^{(b)}$ as $\hat{\mu}_i^{\text{LV}(b)} = \bar{y}_i^{(b)} + (\bar{X}_i - \bar{x}_i)'(Z_i\hat{\beta}^{(b)} + \hat{\nu}_i^{(b)})$, the GREG estimator of $\mu_i^{(b)}$ as $\hat{\mu}_i^{G(b)} = \bar{y}_i^{(b)} + (\bar{X}_i - \bar{x}_i)' Z_i \hat{\beta}_{\text{ols}}^{(b)}$ and the empirical BLUP (EBLUP) estimator of $\mu_i^{(b)}$ as $\hat{\mu}_i^{B(b)} = \bar{X}_i'(Z_i\hat{\beta}^{(b)} + \hat{\nu}_i^{(b)})$, where $\hat{\nu}_i^{(b)} = \hat{\Sigma}_v^{(b)} X_i' \hat{V}_i^{-1(b)}(y_i^{(b)} - X_i Z_i \hat{\beta}^{(b)})$.

We then computed the average mean squared error ($\overline{\text{MSE}}_1$) and average absolute relative error ($\overline{\text{ARE}}_1$)

$$\overline{\text{MSE}}_1 = \frac{1}{m}\sum_i^m \text{MSE}_{1i} \text{ where } \text{MSE}_{1i} = B^{-1}\sum_{b=1}^B (\hat{\mu}_i^{(b)} - \mu_i^{(b)})^2,$$

$$\overline{\text{ARE}}_1 = \frac{1}{m}\sum_i^m \text{ARE}_{1i} \text{ where } \text{ARE}_{1i} = B^{-1}\sum_{b=1}^B |\hat{\mu}_i^{(b)} - \mu_i^{(b)}| / \mu_i^{(b)},$$

where $\hat{\mu}_i^{(b)}$ denotes $\hat{\mu}_i^{\text{LV}(b)}$, $\hat{\mu}_i^{G(b)}$ or $\hat{\mu}_i^{B(b)}$. We report the results in Table 4. Both models with area level covariate z and without z have slightly smaller values of $\overline{\text{MSE}}_1$ and $\overline{\text{ARE}}_1$ for the new GREG estimator relative to the GREG estimator. However, $\overline{\text{MSE}}_1$ and $\overline{\text{ARE}}_1$ are significantly smaller for the EBLUP estimator due to borrowing strength from related areas. Moreover, comparing Tables 2 and 4, we can see that the values of $\overline{\text{MSE}}_1$ in Table 4 are slightly larger than the corresponding values in Table 2 due to estimating model parameters.

Table 4
Comparison of small area estimators: average MSE ($\overline{\text{MSE}}_1$) and average absolute relative error ($\overline{\text{ARE}}_1$) under a model-based framework

Quality Measure	Model without z			Model with z		
	GREG	New GREG	EBLUP	GREG	New GREG	EBLUP
$\overline{\text{MSE}}_1$	1.93	1.73	0.67	1.84	1.73	0.73
$\overline{\text{ARE}}_1$	0.14	0.13	0.08	0.13	0.12	0.08

4.2.2 Simulation study under a finite population framework

To study the performance of the estimators under a finite population framework, we created a synthetic finite population from the Brazilian data consisting of $n = 3{,}876$ sample values $(y_{ij}, x_{1ij}, x_{2ij}, z_i)$. By duplicating the sample values $(y_{ij}, x_{1ij}, x_{2ij}, z_i)$ five times, we treated the new (y, x_1, x_2, z) -data of size 19,380 as our real population.

We generated 500 independent samples ($B = 500$), each of size $n = 700$ and $n = 1{,}400$, by taking simple random samples of size $n_i = 5$ and $n_i = 10$ in each area $i = 1, ..., 140$. As before, for each sample we calculated $\hat{\beta}^{(b)}$ for the new GREG and the BLUP methods and $\hat{\beta}_{\text{ols}}^{(b)}$ for the GREG method. In addition, $\hat{\Sigma}_v^{(b)}$ and $\hat{\sigma}_e^{2(b)}$ were calculated based on the REML method. We also computed the population mean of y_{ij} for each area i as

$$\bar{Y}_i = \sum_{j=1}^{N_i} y_{ij}/N_i,\ i = 1, ..., 140,$$

where N_i is the population size in i^{th} area. Further, for each sample $b = 1, ..., B$, we calculated the new GREG estimate of the i^{th} area mean as $\hat{\bar{Y}}_i^{\text{LV}(b)} = \bar{y}_i^{(b)} + (\bar{X}_i - \bar{x}_i^{(b)})'(Z_i\hat{\beta}^{(b)} + \hat{\nu}_i^{(b)})$, the GREG estimate as $\hat{\bar{Y}}_i^{G(b)} = \bar{y}_i^{(b)} + (\bar{X}_i - \bar{x}_i^{(b)})' Z_i \hat{\beta}_{\text{ols}}^{(b)}$ and the EBLUP estimate as $\hat{\bar{Y}}_i^{B(b)} = f_i\bar{y}_i + (1 - f_i)[\bar{X}_i^{*'}(Z_i\hat{\beta}^{(b)} + \hat{\nu}_i^{(b)})]$, where

$$f_i = n_i/N_i,\ \bar{X}_i^* = \frac{N_i\bar{X}_i - n_i\bar{x}_i}{N_i - n_i},$$

and $\hat{v}_i^{(b)} = \hat{\Sigma}_v^{(b)} X_i' \hat{V}_i^{-1(b)}(y_i^{(b)} - X_i Z_i \hat{\beta}^{(b)})$.

The EBLUP estimator accounts for the finite population corrections f_i.

We computed the average mean squared error ($\overline{\text{MSE}}_2$) and average absolute relative error ($\overline{\text{ARE}}_2$) as

$$\overline{\text{MSE}}_2 = \frac{1}{m}\sum_i^m \text{MSE}_{2i},\quad \overline{\text{ARE}}_2 = \frac{1}{m}\sum_i^m \text{ARE}_{2i},$$

where

$$\text{MSE}_{2i} = \frac{1}{B}\sum_{b=1}^B (\hat{\bar{Y}}_i^{(b)} - \bar{Y}_i)^2,\quad \text{ARE}_{2i} = \frac{1}{B}\sum_{b=1}^B |\hat{\bar{Y}}_i^{(b)} - \bar{Y}_i| / \bar{Y}_i,$$

and $\hat{\bar{Y}}_i^{(b)}$ denotes $\hat{\bar{Y}}_i^{\text{LV}(b)}$, $\hat{\bar{Y}}_i^{G(b)}$ or $\hat{\bar{Y}}_i^{B(b)}$. We report the results in Tables 5 and 6 for $n_i = 5$ and $n_i = 10$ respectively. Both models with area level covariate z and without z are considered.

Table 5
Comparison of small area estimators: average MSE ($\overline{\text{MSE}}_2$) and average absolute relative error ($\overline{\text{ARE}}_2$) under a finite population framework ($n_i = 5$)

Quality Measure	Model without z			Model with z		
	GREG	New GREG	EBLUP	GREG	New GREG	EBLUP
$\overline{\text{MSE}}_2$	11.03	10.02	6.50	10.76	10.06	7.06
$\overline{\text{ARE}}_2$	0.27	0.24	0.18	0.25	0.23	0.22

Table 5 shows that for $n_i = 5$ the new GREG estimator exhibits slightly better performance relative to the GREG estimator in the sense of smaller $\overline{\text{MSE}}_2$ and $\overline{\text{ARE}}_2$. On the other hand, Table 6 reveals that with $n_i = 10$ the GREG estimator has slightly better performance than the new GREG estimator in terms of $\overline{\text{MSE}}_2$ but not $\overline{\text{ARE}}_2$. However, the EBLUP estimator gives substantially smaller $\overline{\text{MSE}}_2$ and $\overline{\text{ARE}}_2$ than the GREG and the new GREG in both cases due to borrowing strength from related small areas. For example, for the model without z and $n_i = 5$, $\overline{\text{MSE}}_2 = 10.02$, 11.03 and 6.50 for the new GREG, the GREG and the EBLUP, respectively.

Table 6
Comparison of small area estimators: average MSE ($\overline{\text{MSE}}_2$) and average absolute relative error ($\overline{\text{ARE}}_2$) under finite population framework (n_i = 10)

Quality Measure	Model without z			Model with z		
	GREG	New GREG	EBLUP	GREG	New GREG	EBLUP
$\overline{\text{MSE}}_2$	6.53	6.77	4.73	6.75	6.96	5.24
$\overline{\text{ARE}}_2$	0.20	0.18	0.15	0.19	0.18	0.19

5. Summary

In this paper, we derived the model mean squared error (MSE) of a two-level model-assisted new GREG estimator of a small area mean, proposed by Lehtonen and Veijanen (1999). In addition, we used a data set of Moura and Holt (1999) to demonstrate empirically that the BLUP estimator is substantially more efficient than the new GREG estimator in terms of model MSE, while the new GREG is only slightly more efficient than the customary GREG based on the regression model $y_i = X_i Z_i \beta + e_i, i = 1, ..., m.$ Moreover, using a simulation study under a model-based framework, we have shown that the new GREG estimator has consistently better performance relative to the GREG estimator in terms of average MSE, $\overline{\text{MSE}}$, and average absolute relative error, $\overline{\text{ARE}}$. However, due to borrowing strength from related small areas, EBLUP estimator exhibits significantly better performance relative to the GREG and the new GREG estimators. In addition, we conducted a simulation study under a finite population framework and showed that the EBLUP estimator outperforms the new GREG and the GREG estimators in terms of $\overline{\text{MSE}}$ and $\overline{\text{ARE}}$.

Acknowledgements

This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada. It is based on a chapter in M.Torabi's Ph.D. thesis written under the supervision of J.N.K. Rao. The authors thank the referees and an associate editor for their helpful comments on the original version of this paper.

Appendix

Derivation of MSE($\tilde{\mu}_i^{\text{LV}}$):

$$\begin{aligned}\text{MSE}(\tilde{\mu}_i^{\text{LV}}) &= E(\tilde{\mu}_i^{\text{LV}} - \mu_i)^2 \\ &= E[\bar{X}_i'(\tilde{\nu}_i - \nu_i)]^2 + E[\bar{y}_i - \bar{x}_i'(Z_i\beta + \tilde{\nu}_i)]^2 \\ &\quad + 2E[(\bar{y}_i - \bar{x}_i'(Z_i\beta + \tilde{\nu}_i))\bar{X}_i'(\tilde{\nu}_i - \nu_i)], \qquad \text{(A.1)}\end{aligned}$$

where the first term on the right hand side of (A.1) is the MSE of the BLUP estimator under the two-level model, given by (9). Moreover, we may write

$$\begin{aligned}E[\bar{y}_i - \bar{x}_i'(Z_i\beta + \tilde{\nu}_i)]^2 &= E[\bar{y}_i - \bar{x}_i'Z_i\beta]^2 \\ &\quad + E(\bar{x}_i'\tilde{\nu}_i)^2 - 2E[(\bar{y}_i - \bar{x}_i'Z_i\beta)(\bar{x}_i'\tilde{\nu}_i)], \qquad \text{(A.2)}\end{aligned}$$

where

$$E[\bar{y}_i - \bar{x}_i'Z_i\beta]^2 = \text{Var}(\bar{y}_i) = \bar{x}_i'\Sigma_v\bar{x}_i + \frac{\sigma_e^2}{n_i},$$

and

$$\begin{aligned}E(\bar{x}_i'\tilde{\nu}_i)^2 &= \text{Var}(\bar{x}_i'\tilde{\nu}_i) + [E(\bar{x}_i'\tilde{\nu}_i)]^2 \\ &= \text{Var}\left[\bar{x}_i'\Sigma_v X_i'V_i^{-1}(\text{y}_i - X_iZ_i\beta)\right] \\ &\quad + \left[E\left(\bar{x}_i'\Sigma_v X_i'V_i^{-1}(\text{y}_i - X_iZ_i\beta)\right)\right]^2 \\ &= \bar{x}_i'\Sigma_v X_i'V_i^{-1}X_i\Sigma_v\bar{x}_i.\end{aligned}$$

In addition,

$$\begin{aligned}E[(\bar{y}_i - \bar{x}_i'Z_i\beta)(\bar{x}_i'\tilde{\nu}_i)] &= E\left[\bar{y}_i\bar{x}_i'\Sigma_v X_i'V_i^{-1}(\text{y}_i - X_iZ_i\beta)\right] \\ &\quad - E\left[\bar{x}_i'Z_i\beta\bar{x}_i'\Sigma_v X_i'V_i^{-1}(\text{y}_i - X_iZ_i\beta)\right],\end{aligned}$$

where the second term is zero. Therefore, we may write

$$\begin{aligned}E[(\bar{y}_i - \bar{x}_i'Z_i\beta)(\bar{x}_i'\tilde{\nu}_i)] &= E\left[\bar{y}_i\bar{x}_i'\Sigma_v X_i'V_i^{-1}\text{y}_i\right] \\ &\quad - \bar{x}_i'Z_i\beta\bar{x}_i'\Sigma_v X_i'V_i^{-1}X_iZ_i\beta,\end{aligned}$$

where the first term can be written as

$$\begin{aligned}&E\left[\frac{1'}{n_i}\text{y}_i\bar{x}_i'\Sigma_v X_i'V_i^{-1}\text{y}_i\right] = \\ &\qquad \frac{1'}{n_i}\left(X_i\Sigma_v\bar{x}_i + X_iZ_i\beta\bar{x}_i'\Sigma_v X_i'V_i^{-1}X_iZ_i\beta\right),\end{aligned}$$

using the following Lemma:

LEMMA 1 (Searle 1971). *If y is a $n \times 1$ vector with mean μ and variance-covariance matrix Σ and b is a $n \times 1$ vector, then* $E(yb'y) = \Sigma b + \mu b'\mu$.

Hence,

$$\begin{aligned} E[(\bar{y}_i - \bar{x}_i' Z_i \beta)(\bar{x}_i' \tilde{\nu}_i)] &= \bar{x}_i' \Sigma_v \bar{x}_i + \bar{x}_i' Z_i \beta \bar{x}_i' \Sigma_v X_i' V_i^{-1} X_i Z_i \beta \\ &\quad - \bar{x}_i' Z_i \beta \bar{x}_i' \Sigma_v X_i' V_i^{-1} X_i Z_i \beta \\ &= \bar{x}_i' \Sigma_v \bar{x}_i. \end{aligned}$$

Then we may write (A.2) as follows:

$$\begin{aligned} E[\bar{y}_i - \bar{x}_i'(Z_i\beta + \tilde{\nu}_i)]^2 &= \bar{x}_i' \Sigma_v \bar{x}_i + \frac{\sigma_e^2}{n_i} \\ &\quad + \bar{x}_i' \Sigma_v X_i' V_i^{-1} X_i \Sigma_v \bar{x}_i - 2\bar{x}_i' \Sigma_v \bar{x}_i \\ &= \bar{x}_i' \Sigma_v X_i' V_i^{-1} X_i \Sigma_v \bar{x}_i - \bar{x}_i' \Sigma_v \bar{x}_i + \frac{\sigma_e^2}{n_i}. \end{aligned} \quad \text{(A.3)}$$

Finally, we need to find the cross-product term of (A.1). We have

$$\begin{aligned} E[(\bar{y}_i - \bar{x}_i'(Z_i\beta + \tilde{\nu}_i))\bar{X}_i'(\tilde{\nu}_i - \nu_i)] &= E[\bar{y}_i \bar{X}_i'(\tilde{\nu}_i - \nu_i)] \\ &\quad - E[\bar{x}_i'(Z_i\beta + \tilde{\nu}_i)\bar{X}_i'(\tilde{\nu}_i - \nu_i)], \end{aligned} \quad \text{(A.4)}$$

where the first term on the right side of (A.4) may be written as

$$\begin{aligned} E[\bar{y}_i \bar{X}_i'(\tilde{\nu}_i - \nu_i)] &= E\left[\bar{y}_i \bar{X}_i'\left(\Sigma_v X_i' V_i^{-1}(y_i - X_i Z_i \beta) - \nu_i\right)\right] \\ &= E\left[\bar{y}_i \bar{X}_i' \Sigma_v X_i' V_i^{-1} y_i\right] \\ &\quad - E\left[\bar{y}_i \bar{X}_i' \Sigma_v X_i' V_i^{-1} X_i Z_i \beta\right] \\ &\quad - E[\bar{y}_i \bar{X}_i' \nu_i]. \end{aligned} \quad \text{(A.5)}$$

The first term of (A.5) is

$$\begin{aligned} E\left[\bar{y}_i \bar{X}_i' \Sigma_v X_i' V_i^{-1} y_i\right] &= E\left[\frac{1'}{n_i} y_i \bar{X}_i' \Sigma_v X_i' V_i^{-1} y_i\right] \\ &= \frac{1'}{n_i} X_i\left(\Sigma_v \bar{X}_i + Z_i \beta \bar{X}_i' \Sigma_v X_i' V_i^{-1} X_i Z_i \beta\right), \end{aligned} \quad \text{(A.6)}$$

the second term of (A.5) can be written as

$$\begin{aligned} E\left[\bar{y}_i \bar{X}_i' \Sigma_v X_i' V_i^{-1} X_i Z_i \beta\right] &= \bar{x}_i' Z_i \beta \bar{X}_i' \Sigma_v X_i' V_i^{-1} X_i Z_i \beta, \end{aligned} \quad \text{(A.7)}$$

and the third term can be expressed as

$$\begin{aligned} E[\bar{y}_i \bar{X}_i' \nu_i] &= E[(\bar{x}_i' Z_i \beta + \bar{x}_i' \nu_i + e_i)\bar{X}_i' \nu_i] \\ &= E[\bar{x}_i' \nu_i \bar{X}_i' \nu_i] = \bar{x}_i' \Sigma_v \bar{X}_i. \end{aligned} \quad \text{(A.8)}$$

Therefore, substituting (A.6), (A.7) and (A.8) in (A.5), we have

$$E[\bar{y}_i \bar{X}_i'(\tilde{\nu}_i - \nu_i)] = 0. \quad \text{(A.9)}$$

We now turn to the second term of (A.4). We have

$$\begin{aligned} E[\bar{x}_i'(Z_i\beta + \tilde{\nu}_i)\bar{X}_i'(\tilde{\nu}_i - \nu_i)] &= E[\bar{x}_i' Z_i \beta \bar{X}_i' \tilde{\nu}_i] + E[\bar{x}_i' \tilde{\nu}_i \bar{X}_i' \tilde{\nu}_i] \\ &\quad - E[\bar{x}_i' Z_i \beta \bar{X}_i' \nu_i] - E[\bar{x}_i' \tilde{\nu}_i \bar{X}_i' \nu_i]. \end{aligned} \quad \text{(A.10)}$$

Then we obtain the following expression for the four terms on the right side of (A.10):

$$\begin{aligned} E[\bar{x}_i' Z_i \beta \bar{X}_i' \tilde{\nu}_i] &= E\left[\bar{x}_i' Z_i \beta \bar{X}_i' \Sigma_v X_i' V_i^{-1}(y_i - X_i Z_i \beta)\right] = 0, \end{aligned} \quad \text{(A.11)}$$

$$\begin{aligned} E[\bar{x}_i' \tilde{\nu}_i \bar{X}_i' \tilde{\nu}_i] &= \bar{x}_i' \Sigma_{\tilde{\nu}} \bar{X}_i + \bar{x}_i' E(\tilde{\nu}_i)\bar{X}_i' E(\tilde{\nu}_i) \\ &= \bar{x}_i' \mathrm{Var}\left[\Sigma_v X_i' V_i^{-1}(y_i - X_i Z_i \beta)\right]\bar{X}_i \\ &\quad + \bar{x}_i' E\left(\Sigma_v X_i' V_i^{-1}(y_i - X_i Z_i \beta)\right) \\ &\quad \times \bar{X}_i' E\left(\Sigma_v X_i' V_i^{-1}(y_i - X_i Z_i \beta)\right) \\ &= \bar{x}_i' \Sigma_v X_i' V_i^{-1} X_i \Sigma_v \bar{X}_i, \end{aligned} \quad \text{(A.12)}$$

$$E[\bar{x}_i' Z_i \beta \bar{X}_i' \nu_i] = 0 \quad \text{(A.13)}$$

and

$$\begin{aligned} E[\bar{x}_i' \tilde{\nu}_i \bar{X}_i' \nu_i] &= E\left[\bar{x}_i' \Sigma_v X_i' V_i^{-1}(y_i - X_i Z_i \beta)\bar{X}_i' \nu_i\right] \\ &= E\left[\bar{x}_i' \Sigma_v X_i' V_i^{-1} y_i \bar{X}_i' \nu_i\right] \\ &\quad - E\left[\bar{x}_i' \Sigma_v X_i' V_i^{-1} X_i Z_i \beta \bar{X}_i' \nu_i\right] \\ &= E\left[\bar{x}_i' \Sigma_v X_i' V_i^{-1}(X_i Z_i \beta + X_i \nu_i + e_i)\bar{X}_i' \nu_i\right] \\ &= E\left[\bar{x}_i' \Sigma_v X_i' V_i^{-1} X_i \nu_i \bar{X}_i' \nu_i\right] \\ &= \bar{x}_i' \Sigma_v X_i' V_i^{-1} X_i \Sigma_v \bar{X}_i. \end{aligned} \quad \text{(A.14)}$$

Therefore, substituting (A.11)-(A.14) in (A.10), we get

$$E[\bar{x}_i'(Z_i\beta + \tilde{\nu}_i)\bar{X}_i'(\tilde{\nu}_i - \nu_i)] = 0. \quad \text{(A.15)}$$

Hence, it follows from (A.4), (A.9) and (A.15) that

$$E[(\bar{y}_i - \bar{x}_i'(Z_i\beta + \tilde{\nu}_i))\bar{X}_i'(\tilde{\nu}_i - \nu_i)] = 0. \quad \text{(A.16)}$$

It now follows from (A.1), (A.3) and (A.16) that

$$\mathrm{MSE}(\tilde{\mu}_i^{\mathrm{LV}}) = \mathrm{MSE}(\tilde{\mu}_i^{B}) + \left\{ \bar{x}_i' \Sigma_v X_i' V_i^{-1} X_i \Sigma_v \bar{x}_i - \bar{x}_i' \Sigma_v \bar{x}_i + \frac{\sigma_e^2}{n_i} \right\},$$

as stated in Theorem 3.

References

Lehtonen, R., and Veijanen, A. (1999). Domain estimation with logistic generalized regression and related estimators. *IASS Satellite Conference on Small Area Estimation*, Riga: Latvian Council of Science, 121- 128.

Lehtonen, R., Särndal, C.-E. and Veijanen, A. (2003). The effect of model choice in estimation for domains, including small domains. *Survey Methodology*, 29, 33-44.

Moura, F.A.S., and Holt, D. (1999). Small area estimation using multilevel models. *Survey Methodology*, 25, 73-80.

Rao, J.N.K. (2003). *Small Area Estimation*. Hoboken: New York: John Wiley & Sons, Inc.

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

Searle, S.R. (1971). *Linear Models*. New York: John Wiley & Sons, Inc.