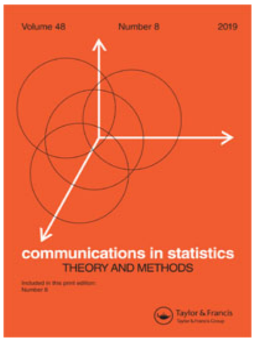

Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: https://www.tandfonline.com/loi/lsta20

Mean-Squared errors of small area estimators under a multivariate linear model for repeated measures data

Innocent Ngaruye, Dietrich Von Rosen & Martin Singull

To cite this article: Innocent Ngaruye, Dietrich Von Rosen & Martin Singull (2019) Mean-Squared errors of small area estimators under a multivariate linear model for repeated measures data, Communications in Statistics - Theory and Methods, 48:8, 2060-2073, DOI: 10.1080/03610926.2018.1444178

To link to this article: https://doi.org/10.1080/03610926.2018.1444178

Published online: 12 Mar 2018.

COMMUNICATIONS IN STATISTICS—THEORY AND METHODS
2019, VOL. 48, NO. 8, 2060–2073
https://doi.org/10.1080/03610926.2018.1444178

Mean-Squared errors of small area estimators under a multivariate linear model for repeated measures data

Innocent Ngaruye[a,c], Dietrich Von Rosen[a,b], and Martin Singull[a]

[a]Department of Mathematics, Linköping University, Linköping, Sweden; [b]Department of Mathematics, College of Science and Technology, University of Rwanda, Kigali, Rwanda; [c]Department of Energy and Technology, Swedish University of Agricultural Sciences, Uppsala, Sweden

ABSTRACT
In this paper, we discuss the derivation of the first and second moments for the proposed small area estimators under a multivariate linear model for repeated measures data. The aim is to use these moments to estimate the mean-squared errors (MSE) for the predicted small area means as a measure of precision. At the first stage, we derive the MSE when the covariance matrices are known. At the second stage, a method based on parametric bootstrap is proposed for bias correction and for prediction error that reflects the uncertainty when the unknown covariance is replaced by its suitable estimator.

ARTICLE HISTORY
Received 2 June 2017
Accepted 19 February 2018

KEYWORDS
Mean-squared errors; Multivariate linear model; Parametric bootstrap; Repeated measures data; Small area estimation.

MATHEMATICS SUBJECT CLASSIFICATION
62F12; 62F40; 62D05

1. Introduction

Reliable information about various population characteristics of interest are needed by policy and decision makers for planning. Therefore, there is a great need to estimate these characteristics of interest via survey sampling, not only for the total target population, but also for local sub-population units (domains). However, most sampling surveys are designed to target much larger populations. Then, the derived direct survey estimators obtained using data only from the target small domain of interest have been found to be with lack of precision due to small sample size connected to this domain. The development of estimation techniques that provide reliable estimates for such a small domain or small area and standard errors of estimates have been a big concern in recent years. These techniques are commonly known as Small Area Estimation (SAE) methods. For comprehensive reviews of SAE, one can refer to Rao (2003); Rao and Isabel (2015); Pfeffermann (2002, 2013).

Longitudinal surveys with repeated measures data over time are developed to study pattern of changes and trends over time. The demand for SAE statistics is not only for cross-sectional data, but also for repeated measures data. Ngaruye et al. (2017) have proposed a multivariate linear model for repeated measures data within small area estimation settings which accounts for grouped response units and random effects variations. One of the methods to ensure the precision of model-based estimators is the assessment of its mean-squared error (MSE). There is an extensive literature about the approaches used for the estimation of MSE of model-based small area estimators. The second-order approximation of asymptotically unbiased MSE based on Taylor series expansion has been considered by various authors such

CONTACT Innocent Ngaruye innocent.ngaruye@liu.se Department of Mathematics, Linköping University, SE-581 83 Linköping, Sweden.

as Kackar and Harville (1984); Datta et al. (1999); Das (Jiang); Baillo and Molina (2009), among others. However, as pointed out by Kubokawa and Nagashima (2012), the Taylor series expansion is sometimes complicated to implement for complicated models with many unknown parameters since it requires the computation of asymptotic bias and asymptotic variance and covariance for estimators of unknown parameters.

In this paper, we aim to derive first and second moments of the proposed estimators for unknown parameters in a special case of the model considered by Ngaruye et al. (2017) and we use these moments to derive the MSE for the predicted small area means. Further, following Butar and Lahiri (2003); Kubokawa and Nagashima (2012) we propose an unbiased estimator of MSE based on the parametric bootstrap method.

The article is organized as follows. In Section 2, the description of the considered model is reviewed. In Section 3, the approach used for estimation and prediction is presented. In Section 4, some preliminary basic results that are referred to in the next sections are provided. The first and second moments of the proposed estimators are presented in Section 5 and in Section 6 an unbiased estimator of MSE of predicted small area means under a multivariate linear model for repeated measures data is derived.

2. Description of the model

Consider the multivariate linear regression model for repeated measurements as defined in Ngaruye et al. (2017). Assume that a p-vector of measurements over time for a finite population of size N divided into m non-overlapping small areas of size N_i, $i = 1, \dots, m$ together with r-vector of auxiliary variables are available for all units in the population. Suppose also that the target population is composed of k group units and denote by N_{ig} the population size of the g-th group units, $g = 1, \dots, k$ such that $\sum_{g=1}^{k} N_{ig} = N_i$ and that the mean growth of the jth unit, $j = 1, \dots, N_i$, in area i for each group to be a polynomial in time of degree $q - 1$. Then, the unit level regression model for the j-th unit coming from the small area i at time t which applies for each one of all k group units can be expressed by

$$y_{ijt} = \beta_0 + \beta_1 t + \cdots + \beta_q t^{q-1} + \boldsymbol{\gamma}'\boldsymbol{x}_{ij} + u_{it} + e_{ijt},$$
$$i = 1, \dots, m;\ \ j = 1, \dots, N_i;\ \ t = t_1, \dots, t_p,$$

where the random errors e_{ijt} and random effects u_{it} are independent and assumed to be i.i.d. normal with mean zero and variance σ_e^2 and σ_{ut}^2, respectively. The $\boldsymbol{\gamma}$ is a vector of fixed regression coefficients representing the effects of auxiliary variables. The $\beta_0, \dots, \beta_q$ are unknown parameters. For all time points, the model can be written in matrix form as

$$\boldsymbol{y}_{ij} = \boldsymbol{A}\boldsymbol{\beta} + \mathbf{1}_p\boldsymbol{\gamma}'\boldsymbol{x}_{ij} + \boldsymbol{u}_i + \boldsymbol{e}_{ij}, \quad i = 1, \dots, m;\ \ j = 1, \dots, N_i;$$

where $\mathbf{1}_p$ is a p-vector of ones and $\boldsymbol{u}_i$ is assumed to be multivariate normally distributed with zero mean and unknown positive definite covariance matrix $\boldsymbol{\Sigma}_u$. In this article we assume $\boldsymbol{A} = \boldsymbol{I}_p$, meaning that we do not consider trends over time.

Hence, the associated multivariate linear regression model for all units coming from the i-th small area belonging to the g-th group units can be expressed by

$$\boldsymbol{Y}_{ig} = \boldsymbol{\beta}_g\mathbf{1}'_{N_{ig}} + \mathbf{1}_p\boldsymbol{\gamma}'\boldsymbol{X}_{ig} + \boldsymbol{u}_i\boldsymbol{z}'_{ig} + \boldsymbol{E}_{ig}, \quad i = 1, \dots, m,\ g = 1, \dots, k, \tag{1}$$

and the model at small area level for all k group units together, belonging to the i-th small area can be expressed as

$$\boldsymbol{Y}_i = \boldsymbol{B}\boldsymbol{C}_i + \mathbf{1}_p\boldsymbol{\gamma}'\boldsymbol{X}_i + \boldsymbol{u}_i\boldsymbol{z}'_i + \boldsymbol{E}_i, \quad \boldsymbol{u}_i \sim \mathcal{N}_p(\mathbf{0}, \boldsymbol{\Sigma}_u), \qquad \boldsymbol{E}_i \sim \mathcal{N}_{p,N_i}(\mathbf{0}, \boldsymbol{\Sigma}_e, \boldsymbol{I}_{N_i}), \tag{2}$$

where $\boldsymbol{Y}_i = (\boldsymbol{Y}_{i1}, \ldots, \boldsymbol{Y}_{ik})$; $\boldsymbol{B} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_k)$; $\boldsymbol{X}_i = (\boldsymbol{X}_{i1}, \ldots, \boldsymbol{X}_{ik})$; $\boldsymbol{z}_i = \frac{1}{\sqrt{N_i}}\boldsymbol{1}_{N_i}$; $\boldsymbol{E}_i = (\boldsymbol{e}_{i1}, \ldots, \boldsymbol{e}_{ik})$; and $\boldsymbol{C}_i = \text{blkdiag}(\boldsymbol{1}'_{N_{i1}}, \ldots, \boldsymbol{1}'_{N_{ik}})$, where the notation $\text{blkdiag}(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_k)$ is a block diagonal matrix with the given matrices $\boldsymbol{A}_i$ on the diagonal. The corresponding model combining all disjoint m small areas and all N units divided into k non-overlapping group units is given by

$$\boldsymbol{Y} = \boldsymbol{BC} + \boldsymbol{1}_p\boldsymbol{\gamma}'\boldsymbol{X} + \boldsymbol{UZ} + \boldsymbol{E}, \tag{3}$$

where $\boldsymbol{Y} = (\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_m)$; $\boldsymbol{C} = (\boldsymbol{1}'_m \otimes \boldsymbol{I}_k)\boldsymbol{C}_D = (\boldsymbol{C}_1, \ldots, \boldsymbol{C}_m)$; $\boldsymbol{C}_D = \text{blkdiag}(\boldsymbol{C}_1, \ldots, \boldsymbol{C}_m)$; $\boldsymbol{X} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m)$; $\boldsymbol{U} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m)$; $\boldsymbol{E} = (\boldsymbol{E}_1, \ldots, \boldsymbol{E}_m)$; $\boldsymbol{Z} = \text{blkdiag}(\boldsymbol{z}'_1, \ldots, \boldsymbol{z}'_m)$ and

$$\boldsymbol{E} \sim \mathcal{N}_{p,N}(\boldsymbol{0}, \boldsymbol{\Sigma}_e, \boldsymbol{I}_N), \qquad \boldsymbol{U} \sim \mathcal{N}_{p,m}(\boldsymbol{0}, \boldsymbol{\Sigma}_u, \boldsymbol{I}_m),$$

with $p \leq m$ and $\boldsymbol{\Sigma}_u$ is an arbitrary positive definite matrix. The symbol $\otimes$ denotes the Kronecker product. It can be worth to point out that the matrix $\boldsymbol{C}_D$ (index D for diagonal) is used later for an orthogonal transformation and partitioning of the model for estimation purpose.

The matrices $\boldsymbol{Z}, \boldsymbol{C}, \boldsymbol{C}_D$ are of full row rank, $\mathscr{C}(\boldsymbol{Z}') \subseteq \mathscr{C}(\boldsymbol{C}'_D)$ and $\boldsymbol{ZZ}' = \boldsymbol{I}_m$, where where $\mathscr{C}(\boldsymbol{A})$ denotes the column vector space generated by the columns of an arbitrary matrix $\boldsymbol{A}$.

In model (3), $\boldsymbol{Y} : p \times N$ is the data matrix, $\boldsymbol{B} : q \times k$ is unknown parameter matrix, $\boldsymbol{C}_D : mk \times N$ with $\text{rank}(\boldsymbol{C}_D) + p \leq N$ and $p \leq m$ is the between individual design matrix accounting for group effects, the matrix $\boldsymbol{U} : p \times m$ is a matrix of random effects, $\boldsymbol{Z} : m \times N$ is the design matrix for random effects and $\boldsymbol{E}$ is the error matrix. The matrix $\boldsymbol{C}$ is the between individual design matrix that captures all k group units. More details about model formulation can be found in Ngaruye et al. (2017).

3. Estimation and prediction

Model (3) is considered as a random effects model with covariates. For a comprehensive review of different considerations of the random effects model, see for e.g., Yokoyama and Fujikoshi (1992); Yokoyama (1995); Nummi (1997); Pan, Fang, and Fang (2002). The estimation and prediction are performed with a likelihood based approach. In what follows, for an arbitrary matrix $\boldsymbol{A}$, $\boldsymbol{A}^o$ stands for any matrix of full rank spanning $\mathscr{C}(\boldsymbol{A})^\perp$, i.e., $\mathscr{C}(\boldsymbol{A}^o) = \mathscr{C}(\boldsymbol{A})^\perp$, where $\mathscr{C}(\boldsymbol{A})^\perp$ is an orthogonal complement to $\mathscr{C}(\boldsymbol{A})$. Moreover, $\boldsymbol{A}^-$ denotes an arbitrary generalized inverse of the matrix $\boldsymbol{A}$ such that $\boldsymbol{AA}^-\boldsymbol{A} = \boldsymbol{A}$. We also denote by $\boldsymbol{P}_A = \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^-\boldsymbol{A}'$ and $\boldsymbol{Q}_A = \boldsymbol{I} - \boldsymbol{P}_A$ the orthogonal projection matrices onto the column space $\mathscr{C}(\boldsymbol{A})$ and onto its orthogonal complement $\mathscr{C}(\boldsymbol{A})^\perp$, respectively. Derivation of estimators and predictors of model (3) are developed in Ngaruye et al. (2017).

We make an orthogonal transformation of model (3) and partition it into three independent models. This partition is based on orthogonal diagonalization of the idempotent matrix $(\boldsymbol{C}_D\boldsymbol{C}'_D)^{-1/2}\boldsymbol{C}_D\boldsymbol{Z}'\boldsymbol{Z}\boldsymbol{C}'_D(\boldsymbol{C}_D\boldsymbol{C}'_D)^{-1/2}$ by $\boldsymbol{\Gamma} = \begin{pmatrix}\boldsymbol{\Gamma}_1 & \boldsymbol{\Gamma}_2\end{pmatrix}$, the orthogonal matrix of eigenvectors for m and $N - m$ elements. We use the following notations to shorten matrix expressions:

$$\boldsymbol{K}_i = \boldsymbol{CR}_i, \qquad \boldsymbol{R}_i = \boldsymbol{C}'_D(\boldsymbol{C}_D\boldsymbol{C}'_D)^{-1/2}\boldsymbol{\Gamma}_i, \quad i = 1, 2. \tag{4}$$

With this orthogonal transformation, we obtain the three independently distributed models

$$\begin{aligned}
\boldsymbol{V}_1 &= \boldsymbol{YR}_1 \sim \mathcal{N}_{p,m}(\boldsymbol{BK}_1 + \boldsymbol{1}_p\boldsymbol{\gamma}'\boldsymbol{XR}_1, \boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_e, \boldsymbol{I}_m),\\
\boldsymbol{V}_2 &= \boldsymbol{YR}_2 \sim \mathcal{N}_{p,mk-m}(\boldsymbol{BK}_2 + \boldsymbol{1}_p\boldsymbol{\gamma}'\boldsymbol{XR}_2, \boldsymbol{\Sigma}_e, \boldsymbol{I}_{mk-m}),\\
\boldsymbol{V}_3 &= \boldsymbol{YC}'^o_D \sim \mathcal{N}_{p,N-mk}(\boldsymbol{1}_p\boldsymbol{\gamma}'\boldsymbol{XC}'^o_D, \boldsymbol{\Sigma}_e, \boldsymbol{I}_{N-mk}).
\end{aligned}$$

In the following the estimation of mean and covariance is performed using a likelihood based approach while the prediction of random effect is performed using Henderson's approach consisting of the maximization of the joint density $f(\boldsymbol{Y}, \boldsymbol{U})$ with respect to $\boldsymbol{U}$ under the assumption of known $\boldsymbol{\Sigma}_e$ and $\boldsymbol{\Sigma}_u$ (Henderson 1973). We now have the following theorem for the estimators and predictor.

Theorem 3.1 (Ngaruye et al. (2017)). *Consider the model (3). Assume that the matrices $\boldsymbol{X}$ and $\boldsymbol{K}_2^{o\prime}\boldsymbol{K}_1$ are of full rank. Then the estimators for $\boldsymbol{\gamma}$, $\boldsymbol{\Sigma}_u$, the linear combination $\boldsymbol{BC}$ and the predictor of $\boldsymbol{U}$ are given by*

$$\widehat{\boldsymbol{\gamma}} = \frac{1}{p}(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{Y}'\mathbf{1}_p,$$
$$\widehat{\boldsymbol{BC}} = \left(\boldsymbol{Y} - \frac{1}{p}\mathbf{1}_p\mathbf{1}_p'\boldsymbol{Y}\boldsymbol{P}_1\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\right)\boldsymbol{R}_2\boldsymbol{K}_2'(\boldsymbol{K}_2\boldsymbol{K}_2')^{-}\boldsymbol{C} + \boldsymbol{W}\boldsymbol{K}_1'\boldsymbol{P}_2\boldsymbol{C},$$

where

$$\boldsymbol{P}_1 = \boldsymbol{C}_D^{\prime o}(\boldsymbol{C}_D^{\prime o})' + \boldsymbol{R}_2\boldsymbol{Q}_{\boldsymbol{K}_2'}\boldsymbol{R}_2',$$
$$\boldsymbol{P}_2 = \boldsymbol{K}_2^o(\boldsymbol{K}_2^{o\prime}\boldsymbol{K}_1\boldsymbol{K}_1'\boldsymbol{K}_2^o)^{-1}\boldsymbol{K}_2^{o\prime},$$
$$\boldsymbol{W} = \left(\boldsymbol{Y} - \frac{1}{p}\mathbf{1}_p\mathbf{1}_p'\boldsymbol{Y}\boldsymbol{P}_1\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\right)\boldsymbol{G}\boldsymbol{R}_1,$$
$$\boldsymbol{G} = \boldsymbol{I}_N - \boldsymbol{R}_2\boldsymbol{K}_2'(\boldsymbol{K}_2\boldsymbol{K}_2')^{-}\boldsymbol{C}.$$

and

$$\widehat{\boldsymbol{\Sigma}}_u = \frac{1}{m-1}\boldsymbol{W}\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{W}' - \boldsymbol{\Sigma}_e, \text{ assumed to be positive definite},$$
$$\widehat{\boldsymbol{U}} = (\boldsymbol{\Sigma}_e + \widehat{\boldsymbol{\Sigma}}_u)^{-1}\widehat{\boldsymbol{\Sigma}}_u\boldsymbol{W}\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{R}_1'\boldsymbol{Z}'.$$

For the details of the proof we refer to Ngaruye et al. (2017). A particular case of model (3) with empirical data analysis for $p = q$ has been discussed by Ngaruye (von Rosen). Here we have corrected the estimator $\widehat{\boldsymbol{\Sigma}}_u$ to be an unbiased estimator for $\boldsymbol{\Sigma}_u$. We note that the estimators given in Theorem 3.1 are unique. The following two lemmas discuss the uniqueness of estimators $\widehat{\boldsymbol{BC}}$ and $\widehat{\boldsymbol{U}}$, the proof of the uniqueness of other estimators is straightforward.

Lemma 3.1. *The estimator $\widehat{\boldsymbol{BC}}$ given in Theorem 3.1 is invariant with respect to the choice of generalized inverse.*

Proof. By replacing $\boldsymbol{W}$ with its expression in the Theorem 3.1, we can rewrite $\widehat{\boldsymbol{BC}}$ as

$$\begin{aligned}\widehat{\boldsymbol{BC}} &= \left(\boldsymbol{Y} - \frac{1}{p}\mathbf{1}_p\mathbf{1}_p'\boldsymbol{Y}\boldsymbol{P}_1\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\right)(\boldsymbol{R}_2\boldsymbol{K}_2'(\boldsymbol{K}_2\boldsymbol{K}_2')^{-} + \boldsymbol{G}\boldsymbol{R}_1\boldsymbol{K}_1'\boldsymbol{P}_2)\boldsymbol{C}\\ &= \left(\boldsymbol{Y} - \frac{1}{p}\mathbf{1}_p\mathbf{1}_p'\boldsymbol{Y}\boldsymbol{P}_1\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\right)(\boldsymbol{R}_2\boldsymbol{K}_2'(\boldsymbol{K}_2\boldsymbol{K}_2')^{-}(\boldsymbol{I}_k - \boldsymbol{K}_1\boldsymbol{K}_1'\boldsymbol{P}_2) + \boldsymbol{R}_1\boldsymbol{K}_1'\boldsymbol{P}_2)\boldsymbol{C}.\end{aligned}$$

We can put

$$\begin{aligned}\boldsymbol{K}_1\boldsymbol{K}_1'\boldsymbol{P}_2 &= \boldsymbol{K}_1\boldsymbol{K}_1'\boldsymbol{K}_2^o(\boldsymbol{K}_2^{o\prime}\boldsymbol{K}_1\boldsymbol{K}_1'\boldsymbol{K}_2^o)^{-1}\boldsymbol{K}_2^{o\prime}\\ &= \boldsymbol{I}_k - \boldsymbol{K}_2(\boldsymbol{K}_2'(\boldsymbol{K}_1\boldsymbol{K}_1')^{-1}\boldsymbol{K}_2)^{-}\boldsymbol{K}_2'(\boldsymbol{K}_1\boldsymbol{K}_1')^{-1}.\end{aligned} \quad (5)$$

Therefore,

$$\widehat{BC} = \left(Y - \frac{1}{p} \mathbf{1}_p \mathbf{1}'_p Y P_1 X' (X P_1 X')^{-1} X \right) (R_2 K'_2 (K_2 K'_2)^- K_2 (K'_2 (K_1 K'_1)^{-1} K_2)^- \times K'_2 (K_1 K'_1)^{-1} + R_1 K'_1 P_2) C$$

is unique, which completes the proof of the lemma. □

Lemma 3.2. *The predictor* $\widehat{U}$ *given in Theorem 3.1 is invariant with respect to the choice of generalized inverse.*

Proof. From the expression of $\widehat{U}$ in Theorem 3.1, inserting the value of $\widehat{\Sigma}_u$ yields

$$\widehat{U} = (\Sigma_e + \widehat{\Sigma}_u)^{-1} \widehat{\Sigma}_u W Q_{K'_1 K_2^o} R'_1 Z' = (I_p - (m-1)\sigma_e^2 (W Q_{K'_1 K_2^o} W')^{-1}) W Q_{K'_1 K_2^o} R'_1 Z'.$$

First observe that

$$W Q_{K'_1 K_2^o} = \left(Y - \frac{1}{p} \mathbf{1}_p \mathbf{1}'_p Y P_1 X' (X P_1 X')^{-1} X \right) R_1 Q_{K'_1 K_2^o} - \left(Y - \frac{1}{p} \mathbf{1}_p \mathbf{1}'_p Y P_1 X' (X P_1 X')^{-1} X \right) R_2 K'_2 (K_2 K'_2)^- K_1 Q_{K'_1 K_2^o}$$

and also note that $K'_1 P_2 K_1 = P_{K'_1 K_2^o}$. From relation (5), it follows that

$$\begin{aligned} K_1 Q_{K'_1 K_2^o} &= K_1 - K_1 P_{K'_1 K_2^o} = (I_k - K_1 K'_1 P_2) K_1 \\ &= K_2 (K'_2 (K_1 K'_1)^{-1} K_2)^- K'_2 (K_1 K'_1)^{-1} K_1. \end{aligned}$$

Thus, $W Q_{K'_1 K_2^o}$ does not depend on a choice of a generalized inverse and the uniqueness of $W Q_{K'_1 K_2^o}$ implies the uniqueness of $\widehat{U}$. □

3.1. Prediction of small area means

The prediction of small area means is based on the prediction approach to finite population under model-based theory. By this approach, the target population under study is considered as a random sample from a larger population characterized by a suitable model and the predictive distribution of the values for non sampled units is obtained given the realized values of sampled units (Bolfarine and Zacks 1992). The model (3) considered in this paper belongs to the extensions of unit level model, often known as nested linear regression model which was originally proposed by Battese, Harter and Fuller (1988) for prediction of mean per-capital income in small geographical areas within counties in the United States. Following these authors, the estimation of a population mean from the sample returns to the prediction of a mean of non-sampled values.

For k group units in all small areas, we consider the partition of N_i units into N_{ig}, $g = 1, \ldots, k$, and n_i sampled units into n_{ig} such that $N_i = \sum_{g=1}^k N_{ig}$ and $n_i = \sum_{g=1}^k n_{ig}$ and similarly for Y_i such that $Y_i = [Y_{i1}, \ldots, Y_{ik}]$. Then the corresponding target small area means at each time point for each group unit are given by

$$\widehat{\mu}_{ig} = \frac{1}{N_{ig}} \left(Y_{ig}^{(s)} \mathbf{1}_{n_{ig}} + \widehat{Y}_{ig}^{(r)} \mathbf{1}_{N_{ig} - n_{ig}} \right), \tag{6}$$

where $Y_i^{(s)} = (y_{i1}, \ldots, y_{in_i}) : p \times n_i$, standing for the sampled n_i observations from the i-th small area and $\widehat{Y}_i^{(r)} = (y_{in_{i+1}}, \ldots, y_{iN_i}) : p \times (N_i - n_i)$, corresponds to the predicted values for non-sampled $(N_i - n_i)$ units from the i-th small area. The first term of the expression

(6) on the right side is known from the sampled observations and the second term is the prediction of non-sampled observations obtained using the considered model (Henderson 1975) and is given by

$$\widehat{\boldsymbol{Y}}_{ig}^{(r)}\mathbf{1}_{N_{ig}-n_{ig}} = \left(1-\frac{n_{ig}}{N_{ig}}\right)\widehat{\boldsymbol{\beta}}_g + \frac{1}{N_{ig}}\mathbf{1}_p\mathbf{1}'_{N_{ig}-n_{ig}}\boldsymbol{X}_{ig}^{(r)'}\widehat{\boldsymbol{\gamma}} + \frac{\sqrt{N_{ig}-n_{ig}}}{N_{ig}}\widehat{\boldsymbol{u}}_i, \tag{7}$$

where $\boldsymbol{X}_{ig}^{(r)}$ stands for the matrix of auxiliary information of non-sampled units in the i-th area belonging to the group units g.

Note that $\widehat{\boldsymbol{\beta}}_g$ is the estimator of $\boldsymbol{\beta}_g$ which is the g-th column of the estimated matrix $\widehat{\boldsymbol{B}}$ and $\widehat{\boldsymbol{u}}_i$ is the i-th column of the predicted matrix $\widehat{\boldsymbol{U}}$. Throughout this article, we are interested in the estimation of mean-squared errors (MSE) for the predicted small area means given in above relation (7). Therefore, we need to calculate the moments of the proposed estimators in order to derive an estimate of the MSE.

4. Preparations for moments calculation

In this section a lemma is given for some technical results that are referred to later on for the moment derivations and other calculations. The following standard definition of the covariance between two random matrices, will be used:

$$\text{cov}(\boldsymbol{X},\boldsymbol{Y}) = \mathbb{E}[\text{vec}\boldsymbol{X}\text{vec}'\boldsymbol{Y}] - \mathbb{E}[\text{vec}\boldsymbol{X}]\mathbb{E}[\text{vec}'\boldsymbol{Y}],$$

where vec is the usual columnwise vectorization. Note that the dispersion matrix $\mathbb{D}[\boldsymbol{X}]$ is defined by $\mathbb{D}[\boldsymbol{X}] = \text{cov}(\boldsymbol{X},\boldsymbol{X})$.

To simplify our calculations later we present the next technical results.

Lemma 4.1. *From the definition of the matrices $\boldsymbol{K}_1, \boldsymbol{K}_2, \boldsymbol{R}_1$ and $\boldsymbol{R}_2$ given in (4) together with the fact that $\mathscr{C}(\boldsymbol{Z}') \subseteq \mathscr{C}(\boldsymbol{C}'_D)$ and $\boldsymbol{Z}\boldsymbol{Z}' = \boldsymbol{I}_m$, the following identities hold:*

(i) $\boldsymbol{R}'_1\boldsymbol{R}_1\boldsymbol{Z} = \boldsymbol{Z}$,
(ii) $\boldsymbol{Z}'\boldsymbol{Z}\boldsymbol{R}_1 = \boldsymbol{R}_1$,
(iii) $\boldsymbol{K}_1\boldsymbol{K}'_1 = \boldsymbol{C}\boldsymbol{Z}'\boldsymbol{Z}\boldsymbol{C}'$,
(iv) $\boldsymbol{R}'_2\boldsymbol{R}_1 = \boldsymbol{R}'_2\boldsymbol{Z}' = \mathbf{0}$ *and hence* $\boldsymbol{P}_1\boldsymbol{Z}' = \mathbf{0}$, *where $\boldsymbol{P}_1$ is defined in Theorem 3.1, and*
(v) $\boldsymbol{C}\boldsymbol{P}_1 = \mathbf{0}$.

Proof. We prove the first three identities, the others are obtained by straightforward calculations. From the orthogonal diagonalization

$$(\boldsymbol{C}_D\boldsymbol{C}'_D)^{-1/2}\boldsymbol{C}_D\boldsymbol{Z}'\boldsymbol{Z}\boldsymbol{C}'_D(\boldsymbol{C}_D\boldsymbol{C}'_D)^{-1/2} = \boldsymbol{\Gamma}\boldsymbol{D}\boldsymbol{\Gamma}' = (\boldsymbol{\Gamma}_1 \quad \boldsymbol{\Gamma}_2)\begin{pmatrix}\boldsymbol{I}_m & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\boldsymbol{\Gamma}'_1\\ \boldsymbol{\Gamma}'_2\end{pmatrix}.$$

It follows that

$$\boldsymbol{\Gamma}_1(\boldsymbol{C}_D\boldsymbol{C}'_D)^{-1/2}\boldsymbol{C}_D\boldsymbol{Z}'\boldsymbol{Z}\boldsymbol{C}'_D(\boldsymbol{C}_D\boldsymbol{C}'_D)^{-1/2}\boldsymbol{\Gamma}'_1 = \boldsymbol{R}'_1\boldsymbol{Z}'\boldsymbol{Z}\boldsymbol{R}_1 = \boldsymbol{I}_m.$$

Furthermore, since $\boldsymbol{Z}\boldsymbol{Z}' = \boldsymbol{R}'_1\boldsymbol{R}_1 = \boldsymbol{I}_m$, it follows that $\boldsymbol{Z}\boldsymbol{Z}'\boldsymbol{Z} = \boldsymbol{R}'_1\boldsymbol{R}_1\boldsymbol{Z} = \boldsymbol{Z}$. Similarly, from $\boldsymbol{R}'_1\boldsymbol{Z}'\boldsymbol{Z}\boldsymbol{R}_1 = \boldsymbol{R}'_1\boldsymbol{R}_1$, we deduce that $\boldsymbol{Z}'\boldsymbol{Z}\boldsymbol{R}_1 = \boldsymbol{R}_1$. Moreover, since $\boldsymbol{K}_1$ is a full row rank, then $\boldsymbol{K}_1\boldsymbol{K}'_1 = \boldsymbol{K}_1\boldsymbol{R}'_1\boldsymbol{Z}'\boldsymbol{Z}\boldsymbol{R}_1\boldsymbol{K}'_1 = \boldsymbol{C}\boldsymbol{R}_1\boldsymbol{R}'_1\boldsymbol{Z}'\boldsymbol{Z}\boldsymbol{R}_1\boldsymbol{R}'_1\boldsymbol{C}' = \boldsymbol{C}\boldsymbol{Z}'\boldsymbol{Z}\boldsymbol{C}'$. □

5. Moments of proposed estimators

In this section, we present the first and second moments of the proposed estimators. All proofs will be given in the end of this article in an Appendix.

5.1. Theoretical moments with known Σ_u

For the purpose of moment calculations, from now on, we suppose that we have complete knowledge on the covariance matrix $\boldsymbol{\Sigma}_u$ or it has been estimated previously so that it is taken to be as known. In that regard, we use the following predictor of $\boldsymbol{U}$

$$\widetilde{\boldsymbol{U}} = (\boldsymbol{\Sigma}_e + \boldsymbol{\Sigma}_u)^{-1}\boldsymbol{\Sigma}_u \boldsymbol{W}\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{R}_1'\boldsymbol{Z}'. \tag{8}$$

Theorem 5.1. *Given the estimators in Theorem 3.1. Then, $\widehat{\boldsymbol{Y}} = \widehat{\boldsymbol{B}}\boldsymbol{C} + \boldsymbol{1}_p\widehat{\boldsymbol{\gamma}}'\boldsymbol{X} + \widetilde{\boldsymbol{U}}\boldsymbol{Z}$ is an unbiased predictor, i.e., $\mathbb{E}[\widehat{\boldsymbol{Y}}] = \mathbb{E}[\boldsymbol{Y}]$.*

The following two theorems give the main results of the paper about moments of the proposed estimators.

Theorem 5.2. *Given the estimators in Theorem 3.1. The dispersion matrices $\mathbb{D}[\widehat{\boldsymbol{\gamma}}]$, $\mathbb{D}[\widehat{\boldsymbol{B}}\boldsymbol{C}]$ and $\mathbb{D}[\widetilde{\boldsymbol{U}}]$ are given by*

$$\begin{aligned}
\mathbb{D}[\widehat{\boldsymbol{\gamma}}] &= \frac{\sigma_e^2}{p}(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1},\\
\mathbb{D}[\widehat{\boldsymbol{B}}\boldsymbol{C}] &= \boldsymbol{C}'\boldsymbol{P}_2\boldsymbol{C}\otimes(\boldsymbol{\Sigma}_u+\boldsymbol{\Sigma}_e) + (\boldsymbol{C}'(\boldsymbol{I}_k-\boldsymbol{P}_2\boldsymbol{K}_1\boldsymbol{K}_1')(\boldsymbol{K}_2\boldsymbol{K}_2')^-\\
&\quad\times(\boldsymbol{I}_k-\boldsymbol{K}_1\boldsymbol{K}_1'\boldsymbol{P}_2)\boldsymbol{C})\otimes\boldsymbol{\Sigma}_e + \frac{\sigma_e^2}{p}\boldsymbol{C}'((\boldsymbol{K}_2\boldsymbol{K}_2')^-\boldsymbol{K}_2\boldsymbol{R}_2' + \boldsymbol{P}_2\boldsymbol{K}_1\boldsymbol{R}_1'\boldsymbol{G}')\\
&\quad\times\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}(\boldsymbol{R}_2\boldsymbol{K}_2'(\boldsymbol{K}_2\boldsymbol{K}_2')^- + \boldsymbol{G}\boldsymbol{R}_1\boldsymbol{K}_1'\boldsymbol{P}_2)\boldsymbol{C}\otimes\boldsymbol{1}_p\boldsymbol{1}_p',\\
\mathbb{D}[\widetilde{\boldsymbol{U}}] &= \boldsymbol{Z}\boldsymbol{R}_1\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}(\boldsymbol{I}_m+\boldsymbol{K}_1(\boldsymbol{K}_2\boldsymbol{K}_2')^-\boldsymbol{K}_1)\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{R}_1'\boldsymbol{Z}'\otimes\boldsymbol{\Sigma}_u(\boldsymbol{\Sigma}_e+\boldsymbol{\Sigma}_u)^{-1}\boldsymbol{\Sigma}_u\\
&\quad+\frac{\sigma_e^2}{p}\boldsymbol{Z}\boldsymbol{R}_1\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{R}_1'\boldsymbol{G}'\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{G}\boldsymbol{R}_1\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{R}_1'\boldsymbol{Z}'\\
&\quad\otimes\boldsymbol{\Sigma}_u(\boldsymbol{\Sigma}_e+\boldsymbol{\Sigma}_u)^{-1}\boldsymbol{1}_p\boldsymbol{1}_p'(\boldsymbol{\Sigma}_e+\boldsymbol{\Sigma}_u)^{-1}\boldsymbol{\Sigma}_u.
\end{aligned}$$

In Theorem 5.2 the dispersion matrices for the estimators and predictor are given. However, for prediction purposes, it is also of interest to derive the covariances between them.

Theorem 5.3. *Consider the estimators given in Theorem 3.1. Then, the covariances $\mathrm{cov}[\widetilde{\boldsymbol{U}},\widehat{\boldsymbol{\gamma}}]$, $\mathrm{cov}[\widehat{\boldsymbol{B}}\boldsymbol{C},\widetilde{\boldsymbol{U}}]$ and $\mathrm{cov}[\widehat{\boldsymbol{B}}\boldsymbol{C},\widehat{\boldsymbol{\gamma}}]$ are given by*

$$\begin{aligned}
\mathrm{cov}[\widetilde{\boldsymbol{U}},\widehat{\boldsymbol{\gamma}}] &= -\frac{\sigma_e^2}{p}\boldsymbol{Z}\boldsymbol{R}_1\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{R}_1'\boldsymbol{G}'\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\otimes\boldsymbol{\Sigma}_u(\boldsymbol{\Sigma}_e+\boldsymbol{\Sigma}_u)^{-1}\boldsymbol{1}_p,\\
\mathrm{cov}[\widehat{\boldsymbol{B}}\boldsymbol{C},\widetilde{\boldsymbol{U}}] &= \frac{\sigma_e^2}{p}\boldsymbol{C}'((\boldsymbol{K}_2\boldsymbol{K}_2')^-\boldsymbol{K}_2\boldsymbol{R}_2'+\boldsymbol{P}_2\boldsymbol{K}_1\boldsymbol{R}_1'\boldsymbol{G}')\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{G}\boldsymbol{R}_1\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{R}_1'\boldsymbol{Z}'\\
&\quad\otimes\boldsymbol{1}_p\boldsymbol{1}_p'(\boldsymbol{\Sigma}_e+\boldsymbol{\Sigma}_u)^{-1}\boldsymbol{\Sigma}_u,\\
\mathrm{cov}[\widehat{\boldsymbol{B}}\boldsymbol{C},\widehat{\boldsymbol{\gamma}}] &= -\frac{\sigma_e^2}{p}\boldsymbol{C}'((\boldsymbol{K}_2\boldsymbol{K}_2')^-\boldsymbol{K}_2\boldsymbol{R}_2'+\boldsymbol{P}_2\boldsymbol{K}_1\boldsymbol{R}_1'\boldsymbol{G}')\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\otimes\boldsymbol{1}_p.
\end{aligned}$$

In the next section we will use Theorem 5.2 and Theorem 5.3 to derive the mean-squared errors of predicted small area means.

5.2. Simulation study with the empirical moments

To put some light on the moment expressions in Section 5.1 we provide a simulation study comparing the first and second moments of $\widehat{\boldsymbol{B}}$ and $\widehat{\boldsymbol{\gamma}}$. Assume we have 8 small areas, i.e, $m = 8$

with $k = 3$ groups and given sample sizes

$$\begin{aligned} n_{11} &= 2, \quad n_{12} = 3, \quad n_{13} = 6,\\ n_{21} &= 7, \quad n_{22} = 9, \quad n_{23} = 13,\\ n_{31} &= 10, \quad n_{32} = 4, \quad n_{33} = 5,\\ n_{41} &= 3, \quad n_{42} = 9, \quad n_{43} = 7,\\ n_{51} &= 4, \quad n_{52} = 6, \quad n_{53} = 2,\\ n_{61} &= 10, \quad n_{62} = 11, \quad n_{63} = 18,\\ n_{71} &= 7, \quad n_{72} = 4, \quad n_{73} = 4,\\ n_{81} &= 6, \quad n_{82} = 9, \quad n_{83} = 5. \end{aligned}$$

Furthermore, let $p = q = 2$ with

$$\boldsymbol{B} = \begin{pmatrix} 8 & 10 & 12 \\ 9 & 11 & 13 \end{pmatrix}, \quad \boldsymbol{\gamma} = (1 \quad 2 \quad 3)', \quad \boldsymbol{\Sigma}_u = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

$\boldsymbol{\Sigma}_e = \boldsymbol{I}_p$, i.e., $\sigma_e^2 = 1$, and the design matrix $\boldsymbol{X}$ is chosen randomly as $x_{ij} \sim U[0, 1]$. Given model (3) we simulate the observation matrix $\boldsymbol{Y}$ for different sample sizes, where we multiply the sample sizes above with the factors 1, 2, 4, 8, 12, 16, 20, i.e., we have $n_{11} = 2, 4, 8, 16, 24, 32, 40$, and so on. For 10 000 replicates, for each set up of sample sizes, we calculate the estimates $\widehat{\boldsymbol{B}}$ and $\widehat{\boldsymbol{\gamma}}$, respectively. Based on these simulated estimates we derive the empirical moments and the Frobenius norm of the difference between the empirical ($\widehat{\mathbb{E}}[\cdot]$ and $\widehat{\mathbb{D}}[\cdot]$) and true moments ($\mathbb{E}[\cdot]$ and $\mathbb{D}[\cdot]$) given in Section 5.1. As we can see in Table 1, the estimates are close to the true values and getting better when the sample sizes increases (i.e., estimators seems to be consistent).

6. Mean-squared errors of predicted small area means

6.1. Derivation of MSE($\widehat{\tau}_{ig}$)

In this section, following Butar and Lahiri (2003) we estimate the mean squared error for predicted small area means in two steps. At the first step, under the assumption of known covariance matrices $\boldsymbol{\Sigma}_e$ and $\boldsymbol{\Sigma}_u$, the derivation of MSE is presented. At the second step, a parametric bootstrap approach is proposed for bias correction and approximation of the uncertainty due to the estimation of $\boldsymbol{\Sigma}_u$. Put

$$\boldsymbol{K} = \left(1 - \frac{n_{ig}}{N_{ig}}\right)\boldsymbol{I}, \;\; \boldsymbol{L} = \frac{1}{N_{ig}}\boldsymbol{1}_p\boldsymbol{1}'_{N_{ig}-n_{ig}}\boldsymbol{X}_{ig}^{(r)'}, \;\; \boldsymbol{M} = \frac{\sqrt{N_{ig}-n_{ig}}}{N_{ig}}\boldsymbol{I}_p,$$
$$i = 1, \ldots, m, \; g = 1, \ldots, k. \tag{9}$$

Table 1. Frobenius norm of the difference between the empirical ($\widehat{\mathbb{E}}[\cdot]$ and $\widehat{\mathbb{D}}[\cdot]$) and true moments ($\mathbb{E}[\cdot]$ and $\mathbb{D}[\cdot]$).

Factor	$\lVert\boldsymbol{B} - \widehat{\mathbb{E}}[\widehat{\boldsymbol{B}}]\rVert_F$	$\lVert\mathbb{D}[\widehat{\boldsymbol{B}}] - \widehat{\mathbb{D}}[\widehat{\boldsymbol{B}}]\rVert_F$	$\lVert\boldsymbol{\gamma} - \widehat{\mathbb{E}}[\widehat{\boldsymbol{\gamma}}]\rVert_F$	$\lVert\mathbb{D}[\widehat{\boldsymbol{\gamma}}] - \widehat{\mathbb{D}}[\widehat{\boldsymbol{\gamma}}]\rVert_F$
1	0.0106	0.0176	0.0072	0.0039
2	0.0202	0.0037	0.0074	0.0025
4	0.0052	0.0026	0.0045	0.0008
8	0.0084	0.0012	0.0038	0.0005
12	0.0035	0.0005	0.0034	0.0002
16	0.0036	0.0005	0.0020	0.0002
20	0.0031	0.0005	0.0016	0.0001

Then, the linear prediction and empirical linear prediction quantities from small area means given in (7) can be written by

$$\begin{aligned}
\widetilde{\boldsymbol{\tau}}_{ig} &= \boldsymbol{K}\widehat{\boldsymbol{\beta}}_g + \boldsymbol{L}\widehat{\boldsymbol{\gamma}} + \boldsymbol{M}\widetilde{\boldsymbol{u}}_i, \quad i = 1, \ldots, m,\ g = 1, \ldots, k \\
\widehat{\boldsymbol{\tau}}_{ig} &= \boldsymbol{K}\widehat{\boldsymbol{\beta}}_g + \boldsymbol{L}\widehat{\boldsymbol{\gamma}} + \boldsymbol{M}\widehat{\boldsymbol{u}}_i, \quad i = 1, \ldots, m,\ g = 1, \ldots, k.
\end{aligned} \tag{10}$$

Let $\boldsymbol{e}_g : k \times 1$ and $\boldsymbol{f}_i : m \times 1$ be the unit basis vectors, i.e., k and m vectors with 1 in the gth and ith position, respectively, and 0 elsewhere. Put $\boldsymbol{\alpha}_g = \boldsymbol{C}'(\boldsymbol{C}\boldsymbol{C}')^{-1}\boldsymbol{e}_g$ so that $\widehat{\boldsymbol{\beta}}_g = \widehat{\boldsymbol{B}}\boldsymbol{C}\boldsymbol{\alpha}_g$, $\widetilde{\boldsymbol{u}}_i = \widetilde{\boldsymbol{U}}\boldsymbol{f}_i$ and $\widehat{\boldsymbol{u}}_i = \widehat{\boldsymbol{U}}\boldsymbol{f}_i$.

Furthermore, let us write $\widehat{\boldsymbol{\tau}}_{ig} - \boldsymbol{\tau}_{ig} = (\widehat{\boldsymbol{\tau}}_{ig} - \widetilde{\boldsymbol{\tau}}_{ig}) + (\widetilde{\boldsymbol{\tau}}_{ig} - \boldsymbol{\tau}_{ig})$ and the MSE of $\widehat{\boldsymbol{\tau}}_{ig}$ can be obtained by

$$\mathrm{MSE}(\widehat{\boldsymbol{\tau}}_{ig}) = \mathrm{MSE}(\widetilde{\boldsymbol{\tau}}_{ig}) + 2\mathbb{E}[(\widehat{\boldsymbol{\tau}}_{ig} - \widetilde{\boldsymbol{\tau}}_{ig})(\widetilde{\boldsymbol{\tau}}_{ig} - \boldsymbol{\tau}_{ig})'] + \mathbb{E}[(\widehat{\boldsymbol{\tau}}_{ig} - \widetilde{\boldsymbol{\tau}}_{ig})(\widehat{\boldsymbol{\tau}}_{ig} - \widetilde{\boldsymbol{\tau}}_{ig})']. \tag{11}$$

The first term of the right hand side of (11) has the form

$$\begin{aligned}
\mathrm{MSE}(\widetilde{\boldsymbol{\tau}}_{ig}) &= \mathbb{E}[(\widetilde{\boldsymbol{\tau}}_{ig} - \boldsymbol{\tau}_{ig})(\widetilde{\boldsymbol{\tau}}_{ig} - \boldsymbol{\tau}_{ig})'] \\
&= \mathbb{E}[(\boldsymbol{K}(\widehat{\boldsymbol{\beta}}_g - \boldsymbol{\beta}_g) + \boldsymbol{L}(\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}) + \boldsymbol{M}(\widetilde{\boldsymbol{u}}_i - \boldsymbol{u}_i)) \\
&\quad \times (\boldsymbol{K}(\widehat{\boldsymbol{\beta}}_g - \boldsymbol{\beta}_g) + \boldsymbol{L}(\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}) + \boldsymbol{M}(\widetilde{\boldsymbol{u}}_i - \boldsymbol{u}_i))'] \\
&= \boldsymbol{K}\mathbb{D}[\widehat{\boldsymbol{\beta}}_g]\boldsymbol{K}' + \boldsymbol{K}\mathrm{cov}[\widehat{\boldsymbol{\beta}}_g, \widehat{\boldsymbol{\gamma}}]\boldsymbol{L}' + \boldsymbol{K}\mathrm{cov}[\widehat{\boldsymbol{\beta}}_g, \widetilde{\boldsymbol{u}}_i]\boldsymbol{M}' + \boldsymbol{L}\mathbb{D}[\widehat{\boldsymbol{\gamma}}]\boldsymbol{L}' \\
&\quad + \boldsymbol{L}\mathrm{cov}[\widehat{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{u}}_i]\boldsymbol{M}' + \boldsymbol{M}\mathbb{D}[\widetilde{\boldsymbol{u}}_i, \widehat{\boldsymbol{\beta}}_g]\boldsymbol{K}' + \boldsymbol{L}\mathrm{cov}[\widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{B}}]\boldsymbol{K}' \\
&\quad + \boldsymbol{M}\mathrm{cov}[\widetilde{\boldsymbol{u}}_i, \widehat{\boldsymbol{\gamma}}]\boldsymbol{L}' + \boldsymbol{M}\mathbb{D}[\widetilde{\boldsymbol{u}}_i - \boldsymbol{u}_i]\boldsymbol{M}'.
\end{aligned} \tag{12}$$

Observe that, from the definitions given to $\widehat{\boldsymbol{\beta}}_g = \widehat{\boldsymbol{B}}\boldsymbol{C}\boldsymbol{\alpha}_g$, $\widetilde{\boldsymbol{u}}_i = \widetilde{\boldsymbol{U}}\boldsymbol{f}_i$ and $\widehat{\boldsymbol{u}}_i = \widehat{\boldsymbol{U}}\boldsymbol{f}_i$ the covariances presented in Equation (12) are expressed by

$$\begin{aligned}
\mathrm{cov}[\widehat{\boldsymbol{\beta}}_g, \widetilde{\boldsymbol{u}}_i] &= \mathrm{cov}[\widehat{\boldsymbol{B}}\boldsymbol{C}\boldsymbol{\alpha}_g, \widetilde{\boldsymbol{U}}\boldsymbol{f}_i] = (\boldsymbol{\alpha}_g' \otimes \boldsymbol{I}_p)\mathrm{cov}[\widehat{\boldsymbol{B}}\boldsymbol{C}, \widetilde{\boldsymbol{U}}](\boldsymbol{f}_i \otimes \boldsymbol{I}_p), \\
\mathrm{cov}[\widetilde{\boldsymbol{u}}_i, \widehat{\boldsymbol{\gamma}}] &= \mathrm{cov}[\widetilde{\boldsymbol{U}}\boldsymbol{f}_i, \widehat{\boldsymbol{\gamma}}] = (\boldsymbol{f}_i' \otimes \boldsymbol{I}_p)\mathrm{cov}[\widetilde{\boldsymbol{U}}, \widehat{\boldsymbol{\gamma}}], \\
\mathrm{cov}[\widehat{\boldsymbol{\beta}}_g, \widehat{\boldsymbol{\gamma}}] &= \mathrm{cov}[\widehat{\boldsymbol{B}}\boldsymbol{C}\boldsymbol{\alpha}_g, \widehat{\boldsymbol{\gamma}}] = (\boldsymbol{\alpha}_g' \otimes \boldsymbol{I}_p)\mathrm{cov}[\widehat{\boldsymbol{B}}\boldsymbol{C}, \widehat{\boldsymbol{\gamma}}].
\end{aligned}$$

Similarly, the dispersion matrices presented in Equation (12) are expressed by

$$\begin{aligned}
\mathbb{D}[\widehat{\boldsymbol{\beta}}_g] &= \mathbb{D}[\widehat{\boldsymbol{B}}\boldsymbol{C}\boldsymbol{\alpha}_g] = (\boldsymbol{\alpha}_g' \otimes \boldsymbol{I}_p)\mathbb{D}[\widehat{\boldsymbol{B}}\boldsymbol{C}](\boldsymbol{\alpha}_g \otimes \boldsymbol{I}_p), \\
\mathbb{D}[\widetilde{\boldsymbol{u}}_i] &= \mathbb{D}[\widetilde{\boldsymbol{U}}\boldsymbol{f}_i] = (\boldsymbol{f}_i' \otimes \boldsymbol{I}_p)\mathbb{D}[\widetilde{\boldsymbol{U}}](\boldsymbol{f}_i \otimes \boldsymbol{I}_p),
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{D}[\widetilde{\boldsymbol{u}}_i - \boldsymbol{u}_i] &= (\boldsymbol{f}_i' \otimes \boldsymbol{I}_p)\mathbb{D}[\widetilde{\boldsymbol{U}} - \boldsymbol{U}](\boldsymbol{f}_i \otimes \boldsymbol{I}_p) \\
&= (\boldsymbol{f}_i' \otimes \boldsymbol{I}_p)(\mathbb{D}[\widetilde{\boldsymbol{U}}] - 2\mathrm{cov}(\widetilde{\boldsymbol{U}}, \boldsymbol{U}) + \mathbb{D}[\boldsymbol{U}])(\boldsymbol{f}_i \otimes \boldsymbol{I}_p) \\
&= (\boldsymbol{f}_i' \otimes \boldsymbol{I}_p)(\mathbb{D}[\boldsymbol{U}] - \mathbb{D}[\widetilde{\boldsymbol{U}}])(\boldsymbol{f}_i \otimes \boldsymbol{I}_p).
\end{aligned}$$

Altogether we can give the following theorem.

Theorem 6.1. *The MSE of the linear prediction from small area means* $\widetilde{\boldsymbol{\tau}}_{ig}$ *is given as*

$$\begin{aligned}
\mathrm{MSE}(\widetilde{\boldsymbol{\tau}}_{ig}) &= \boldsymbol{K}(\boldsymbol{\alpha}_g' \otimes \boldsymbol{I}_p)\mathbb{D}[\widehat{\boldsymbol{B}}\boldsymbol{C}](\boldsymbol{\alpha}_g \otimes \boldsymbol{I}_p)\boldsymbol{K}' + \boldsymbol{K}(\boldsymbol{\alpha}_g' \otimes \boldsymbol{I}_p)\mathrm{cov}[\widehat{\boldsymbol{B}}\boldsymbol{C}, \widehat{\boldsymbol{\gamma}}]\boldsymbol{L}' \\
&\quad + \boldsymbol{K}(\boldsymbol{\alpha}_g' \otimes \boldsymbol{I}_p)\mathrm{cov}[\widehat{\boldsymbol{B}}\boldsymbol{C}, \widetilde{\boldsymbol{U}}](\boldsymbol{f}_i \otimes \boldsymbol{I}_p)\boldsymbol{M}' + \boldsymbol{L}\mathbb{D}[\widehat{\boldsymbol{\gamma}}]\boldsymbol{L}' \\
&\quad + \boldsymbol{L}\mathrm{cov}[\widetilde{\boldsymbol{U}}, \widehat{\boldsymbol{\gamma}}]'(\boldsymbol{f}_i \otimes \boldsymbol{I}_p)\boldsymbol{M}' + \boldsymbol{M}(\boldsymbol{f}_i' \otimes \boldsymbol{I}_p)\mathrm{cov}[\widehat{\boldsymbol{B}}\boldsymbol{C}, \widetilde{\boldsymbol{U}}]'(\boldsymbol{\alpha}_g \otimes \boldsymbol{I}_p)\boldsymbol{K}' \\
&\quad + \boldsymbol{L}\mathrm{cov}[\widehat{\boldsymbol{B}}\boldsymbol{C}, \widehat{\boldsymbol{\gamma}}]'(\boldsymbol{\alpha}_g \otimes \boldsymbol{I}_p)\boldsymbol{K}' + \boldsymbol{M}(\boldsymbol{f}_i' \otimes \boldsymbol{I}_p)\mathrm{cov}[\widetilde{\boldsymbol{U}}, \widehat{\boldsymbol{\gamma}}]\boldsymbol{L}' \\
&\quad + \boldsymbol{M}\boldsymbol{\Sigma}_u\boldsymbol{M}' - \boldsymbol{M}(\boldsymbol{f}_i' \otimes \boldsymbol{I}_p)\mathbb{D}[\widetilde{\boldsymbol{U}}](\boldsymbol{f}_i \otimes \boldsymbol{I}_p)\boldsymbol{M}',
\end{aligned} \tag{13}$$

where the dispersion matrices $\mathbb{D}[\widehat{\boldsymbol{\gamma}}]$, $\mathbb{D}[\widehat{\boldsymbol{BC}}]$, $\mathbb{D}[\widetilde{\boldsymbol{U}}]$ *and the covariance matrices* $\operatorname{cov}[\widetilde{\boldsymbol{U}}, \widehat{\boldsymbol{\gamma}}]$, $\operatorname{cov}[\widehat{\boldsymbol{BC}}, \widetilde{\boldsymbol{U}}]$, $\operatorname{cov}[\widehat{\boldsymbol{BC}}, \widehat{\boldsymbol{\gamma}}]$ *are presented in Theorem 5.2 and Theorem 5.3, respectively.*

6.2. Estimation of MSE($\widehat{\boldsymbol{\tau}}_{ig}$)

The second two terms of the right hand side of Equation (11) are intractable and need to be approximated. It is important to note that in practice, the covariance matrix $\boldsymbol{\Sigma}_u$ is unknown. As pointed out by different authors (see for example Das (Jiang)), a naive estimator of MSE obtained by replacing the unknown covariance matrix $\boldsymbol{\Sigma}_u$ by its estimator $\widehat{\boldsymbol{\Sigma}}_u$ in (13) that ignores the variability associated with $\widehat{\boldsymbol{\Sigma}}_u$ can lead to underestimation of the true MSE. Therefore, we propose a parametric bootstrap method to estimate the MSE when $\boldsymbol{\Sigma}_u$ is replaced by its estimator.

The approximate estimator of MSE($\widehat{\boldsymbol{\tau}}_{ig}$) given in Equation (11) can be decomposed as

$$\operatorname{MSE}(\widehat{\boldsymbol{\tau}}_{ig}) = \boldsymbol{G}_{1i}(\boldsymbol{\Sigma}_u) + \boldsymbol{G}_{2i}(\boldsymbol{\Sigma}_u) + \boldsymbol{G}_{3i}(\boldsymbol{\Sigma}_u), \tag{14}$$

where

$$\boldsymbol{G}_{1i}(\boldsymbol{\Sigma}_u) + \boldsymbol{G}_{2i}(\boldsymbol{\Sigma}_u) = \operatorname{MSE}(\widetilde{\boldsymbol{\tau}}_{ig})$$
$$\boldsymbol{G}_{3i}(\boldsymbol{\Sigma}_u) = 2\mathbb{E}[(\widehat{\boldsymbol{\tau}}_{ig} - \widetilde{\boldsymbol{\tau}}_{ig})(\widetilde{\boldsymbol{\tau}}_{ig} - \boldsymbol{\tau}_{ig})'] + \mathbb{E}[(\widehat{\boldsymbol{\tau}}_{ig} - \widetilde{\boldsymbol{\tau}}_{ig})(\widehat{\boldsymbol{\tau}}_{ig} - \widetilde{\boldsymbol{\tau}}_{ig})'].$$

For known $\boldsymbol{\Sigma}_u$, the quantity $\boldsymbol{G}_{1i}(\boldsymbol{\Sigma}_u) + \boldsymbol{G}_{2i}(\boldsymbol{\Sigma}_u)$ is given in Equation (13). When $\boldsymbol{\Sigma}_u$ is replaced by its estimator, the quantity $\boldsymbol{G}_{1i}(\widehat{\boldsymbol{\Sigma}}_u) + \boldsymbol{G}_{2i}(\widehat{\boldsymbol{\Sigma}}_u)$ introduces an additional bias related to $\widehat{\boldsymbol{\Sigma}}_u$, i.e., $\mathbb{E}[\widehat{\boldsymbol{\Sigma}}_u] - \boldsymbol{\Sigma}_u$ (see Datta and Lahiri (2000)). Following Butar and Lahiri (2003); Kubokawa and Nagashima (2012), we propose a parametric bootstrap method to estimate the first two terms of (14) by correcting the bias of $\boldsymbol{G}_{1i}(\widehat{\boldsymbol{\Sigma}}_u) + \boldsymbol{G}_{2i}(\widehat{\boldsymbol{\Sigma}}_u)$ when $\boldsymbol{\Sigma}_u$ is replaced by its estimator $\widehat{\boldsymbol{\Sigma}}_u$ in Equation (13) and secondly for estimating the third term $\boldsymbol{G}_{3i}(\boldsymbol{\Sigma}_u)$ of Equation (14).

Consider the bootstrap model

$$\boldsymbol{Y}_i^* \mid \boldsymbol{u}_i^* \sim \mathcal{N}_{p,n_i}\left(\widehat{\boldsymbol{BC}}_i + \boldsymbol{1}_p\widehat{\boldsymbol{\gamma}}'\boldsymbol{X}_i + \boldsymbol{u}_i^*\boldsymbol{z}_i', \boldsymbol{\Sigma}_e, \boldsymbol{I}_{n_i}\right), \quad i = 1, \ldots, m, \tag{15}$$

where $\boldsymbol{u}_i^* \sim \mathcal{N}_p(\boldsymbol{0}, \widehat{\boldsymbol{\Sigma}}_u)$. If we put

$$\widehat{\boldsymbol{\tau}}_{ig}(\boldsymbol{Y}_i; \widehat{\boldsymbol{\beta}}_g, \widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{\Sigma}}_u) = \boldsymbol{K}\widehat{\boldsymbol{\beta}}_g + \boldsymbol{L}\widehat{\boldsymbol{\gamma}} + \boldsymbol{M}\widetilde{\boldsymbol{u}}_i, \tag{16}$$
$$\widehat{\boldsymbol{\tau}}_{ig}(\boldsymbol{Y}_i; \widehat{\boldsymbol{\beta}}_g^*, \widehat{\boldsymbol{\gamma}}^*, \widehat{\boldsymbol{\Sigma}}_u) = \boldsymbol{K}\widehat{\boldsymbol{\beta}}_g^* + \boldsymbol{L}\widehat{\boldsymbol{\gamma}}^* + \boldsymbol{M}\widetilde{\boldsymbol{u}}_i^*, \tag{17}$$
$$\widehat{\boldsymbol{\tau}}_{ig}(\boldsymbol{Y}_i; \widehat{\boldsymbol{\beta}}_g^*, \widehat{\boldsymbol{\gamma}}^*, \widehat{\boldsymbol{\Sigma}}_u^*) = \boldsymbol{K}\widehat{\boldsymbol{\beta}}_g^* + \boldsymbol{L}\widehat{\boldsymbol{\gamma}}^* + \boldsymbol{M}\widehat{\boldsymbol{u}}_i^*, \quad i = 1, \ldots, m, \ g = 1, \ldots, k. \tag{18}$$

then the quantity $\boldsymbol{G}_{1i}(\boldsymbol{\Sigma}_u) + \boldsymbol{G}_{2i}(\boldsymbol{\Sigma}_u)$ can be estimated by

$$\boldsymbol{G}_{1i}(\widehat{\boldsymbol{\Sigma}}_u) + \boldsymbol{G}_{2i}(\widehat{\boldsymbol{\Sigma}}_u) - \mathbb{E}_*\left[\boldsymbol{G}_{1i}(\widehat{\boldsymbol{\Sigma}}_u^*) + \boldsymbol{G}_{2i}(\widehat{\boldsymbol{\Sigma}}_u^*) - \boldsymbol{G}_{1i}(\widehat{\boldsymbol{\Sigma}}_u) - \boldsymbol{G}_{2i}(\widehat{\boldsymbol{\Sigma}}_u)\right]$$

and the quantity $\boldsymbol{G}_{3i}(\boldsymbol{\Sigma}_u)$ estimated by

$$2\mathbb{E}_*[(\widehat{\boldsymbol{\tau}}_{ig}(\boldsymbol{Y}_i; \widehat{\boldsymbol{\beta}}_g^*, \widehat{\boldsymbol{\gamma}}^*, \widehat{\boldsymbol{\Sigma}}_u^*) - \widehat{\boldsymbol{\tau}}_{ig}(\boldsymbol{Y}_i^*; \widehat{\boldsymbol{\beta}}_g^*, \widehat{\boldsymbol{\gamma}}^*, \widehat{\boldsymbol{\Sigma}}_u))(\widehat{\boldsymbol{\tau}}_{ig}(\boldsymbol{Y}_i; \widehat{\boldsymbol{\beta}}_g^*, \widehat{\boldsymbol{\gamma}}^*, \widehat{\boldsymbol{\Sigma}}_u) - \widehat{\boldsymbol{\tau}}_{ig}(\boldsymbol{Y}_i; \widehat{\boldsymbol{\beta}}_g, \widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{\Sigma}}_u))']$$
$$+ \mathbb{E}_*[(\widehat{\boldsymbol{\tau}}_{ig}(\boldsymbol{Y}_i; \widehat{\boldsymbol{\beta}}_g^*, \widehat{\boldsymbol{\gamma}}^*, \widehat{\boldsymbol{\Sigma}}_u^*) - \widehat{\boldsymbol{\tau}}_{ig}(\boldsymbol{Y}_i; \widehat{\boldsymbol{\beta}}_g^*, \widehat{\boldsymbol{\gamma}}^*, \widehat{\boldsymbol{\Sigma}}_u))()'],$$

where $\mathbb{E}_*$ is the expectation with respect to model (15), and the calculation of $\widehat{\boldsymbol{\gamma}}^*$, $\widehat{\boldsymbol{\beta}}_g^*$, $\widehat{\boldsymbol{\Sigma}}_u^*$ is performed similarly to that of $\widehat{\boldsymbol{\gamma}}$, $\widehat{\boldsymbol{\beta}}_g$, $\widehat{\boldsymbol{\Sigma}}_u$ except that $\widehat{\boldsymbol{\gamma}}^*$, $\widehat{\boldsymbol{\beta}}_g^*$, $\widehat{\boldsymbol{\Sigma}}_u^*$ are calculated based on $\boldsymbol{Y}_i^*$ instead of $\boldsymbol{Y}_i$.

Thus, when the unknown covariance $\boldsymbol{\Sigma}_u$ is replaced by its estimator in MSE($\widehat{\boldsymbol{\tau}}_{ig}$) of Equation (14), we obtain the proposed estimator of MSE whose results are summarized in the following theorem.

Theorem 6.2. *Consider the model (3) and assume that the matrices $\boldsymbol{X}$ and $\boldsymbol{K}_2^{o\prime}\boldsymbol{K}_1$ are of full rank. Consider the bootstrap model (15) and the estimated quantities (16)-(18). Then the estimator of MSE for predicted small area means given in (10) can be expressed by*

$$\begin{aligned}\widehat{\mathrm{MSE}}(\widehat{\boldsymbol{\tau}}_{ig}) = {} & 2[\boldsymbol{G}_{1i}(\widehat{\boldsymbol{\Sigma}}_u) + \boldsymbol{G}_{2i}(\widehat{\boldsymbol{\Sigma}}_u)] - \mathbb{E}_*[\boldsymbol{G}_{1i}(\widehat{\boldsymbol{\Sigma}}_u^*) + \boldsymbol{G}_{2i}(\widehat{\boldsymbol{\Sigma}}_u^*)] \\ & + 2\mathbb{E}_*[(\widehat{\boldsymbol{\tau}}_{ig}(\boldsymbol{Y}_i; \widehat{\boldsymbol{\beta}}_g^*, \widehat{\boldsymbol{\gamma}}^*, \widehat{\boldsymbol{\Sigma}}_u^*) - \widehat{\boldsymbol{\tau}}_{ig}(\boldsymbol{Y}_i^*; \widehat{\boldsymbol{\beta}}_g^*, \widehat{\boldsymbol{\gamma}}^*, \widehat{\boldsymbol{\Sigma}}_u)) \\ & \times (\widehat{\boldsymbol{\tau}}_{ig}(\boldsymbol{Y}_i; \widehat{\boldsymbol{\beta}}_g^*, \widehat{\boldsymbol{\gamma}}^*, \widehat{\boldsymbol{\Sigma}}_u) - \widehat{\boldsymbol{\tau}}_{ig}(\boldsymbol{Y}_i; \widehat{\boldsymbol{\beta}}_g, \widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{\Sigma}}_u))'] \\ & + \mathbb{E}_*[(\widehat{\boldsymbol{\tau}}_{ig}(\boldsymbol{Y}_i; \widehat{\boldsymbol{\beta}}_g^*, \widehat{\boldsymbol{\gamma}}^*, \widehat{\boldsymbol{\Sigma}}_u^*) - \widehat{\boldsymbol{\tau}}_{ig}(\boldsymbol{Y}_i; \widehat{\boldsymbol{\beta}}_g^*, \widehat{\boldsymbol{\gamma}}^*, \widehat{\boldsymbol{\Sigma}}_u))()'], \end{aligned} \tag{19}$$

where $\boldsymbol{G}_{1i}(\widehat{\boldsymbol{\Sigma}}_u) + \boldsymbol{G}_{2i}(\widehat{\boldsymbol{\Sigma}}_u)$ and $\boldsymbol{G}_{1i}(\widehat{\boldsymbol{\Sigma}}_u^*) + \boldsymbol{G}_{2i}(\widehat{\boldsymbol{\Sigma}}_u^*)$ are given by Equation (13) with $\boldsymbol{\Sigma}_u$ replaced by $\widehat{\boldsymbol{\Sigma}}_u$ and $\widehat{\boldsymbol{\Sigma}}_u^*$, respectively. In addition, the dispersion matrices $\mathbb{D}[\widehat{\boldsymbol{\gamma}}]$, $\mathbb{D}[\widehat{\boldsymbol{B}\boldsymbol{C}}]$, $\mathbb{D}[\widetilde{\boldsymbol{U}}]$ and the covariance matrices $\mathrm{cov}[\widetilde{\boldsymbol{U}}, \widehat{\boldsymbol{\gamma}}]$, $\mathrm{cov}[\widehat{\boldsymbol{B}\boldsymbol{C}}, \widetilde{\boldsymbol{U}}]$, $\mathrm{cov}[\widehat{\boldsymbol{B}\boldsymbol{C}}, \widehat{\boldsymbol{\gamma}}]$ involved in Equation (13) are presented in Theorem 5.2 and Theorem 5.3, respectively.

References

Baillo, A., and I. Molina. 2009. Mean-squared errors of small-area estimators under a unit-level multivariate model. *Statistics* 43 (6):553–69.

Battese, G. E., R. M. Harter, and W. A. Fuller. 1988. An error-components model for prediction of county crop areas using survey and satellite data. *Journal of the American Statistical Association* 83 (401):28–36.

Bolfarine, H., and S. Zacks, 1992. *Prediction theory for finite populations*. New York: Springer.

Butar, F. B., and P. Lahiri. 2003. On measures of uncertainty of empirical bayes small-area estimators. *Journal of Statistical Planning and Inference* 112 (1):63–76.

Das, K., J. Jiang, and J. Rao. 2004. Mean squared error of empirical predictor. *The Annals of Statistics* 32 (2):818–40.

Datta, G. S., and P. Lahiri. 2000. A unified measure of uncertainty of estimated best linear unbiased predictors in small area estimation problems. *Statistica Sinica* 10:613–27.

Datta, G. S., P. Lahiri, T. Maiti, and K. L. Lu. 1999. Hierarchical Bayes estimation of unemployment rates for the states of the US. *Journal of the American Statistical Association* 94 (448):1074–82.

Henderson, C. R. 1973. Sire evaluation and genetic trends. *Journal of Animal Science* 1973 (Symposium):10–41.

Henderson, C. R. 1975. Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31 (2):423–47.

Kackar, R. N., and D. A. Harville. 1984. Approximations for standard errors of estimators of fixed and random effects in mixed linear models. *Journal of the American Statistical Association* 79 (388): 853–62.

Kubokawa, T., and B. Nagashima. 2012. Parametric bootstrap methods for bias correction in linear mixed models. *Journal of Multivariate Analysis* 106:1–16.

Ngaruye, I., J. Nzabanita, D. von Rosen, and M. Singull. 2017. Small area estimation under a multivariate linear model for repeated measures data. *Communications in Statistics—Theory and Methods* 46 (21):10835–50.

Ngaruye, I., D. von Rosen, and M. Singull. 2016. Crop yield estimation at district level for agricultural seasons 2014 in Rwanda. *African Journal of Applied Statistics* 3 (1):69–90.

Nummi, T. 1997. Estimation in a random effects growth curve model. *Journal of Applied Statistics* 24 (2):157–68.

Pan, J.-X., K. Fang, and K.-T. Fang. 2002. *Growth curve models and statistical diagnostics*. New York: Springer Science & Business Media.

Pfeffermann, D. 2002. Small area estimation - new developments and directions. *International Statistical Review* 70 (1):125–43.

Pfeffermann, D. 2013. New important developments in small area estimation. *Statistical Science* 28 (1):40–68.

Rao, J. N. K. 2003. *Small area estimation*. New York: John Wiley and Sons.

Rao, J. N. K., and M. Isabel. 2015. *Small area estimation*. 2nd ed. New York: John Wiley and Sons.
Yokoyama, T. 1995. Statistical inference on some mixed manova-gmanova models with random effects. *Hiroshima Mathematical Journal* 25 (3):441–74.
Yokoyama, T., and Y. Fujikoshi. 1992. Tests for random-effects covariance structures in the growth curve model with covariates. *Hiroshima Mathematical Journal* 22:195–202.

Appendix

In this Appendix the proofs for the theorems in Section 5 are presented.

Proof of Theorem 5.1. First we show that $\widehat{\boldsymbol{\gamma}}$ is an unbiased estimator. We have

$$\begin{aligned}\mathbb{E}[\widehat{\boldsymbol{\gamma}}] &= \frac{1}{p}(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{P}_1\mathbb{E}[\boldsymbol{Y}']\mathbf{1}_p = \frac{1}{p}(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{P}_1(\boldsymbol{C}'\boldsymbol{B}' + \boldsymbol{X}'\boldsymbol{\gamma}\mathbf{1}_p')\mathbf{1}_p \\ &= \frac{1}{p}(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}'\boldsymbol{\gamma}\mathbf{1}_p'\mathbf{1}_p = \boldsymbol{\gamma}.\end{aligned}$$

Therefore, $\mathbb{E}[\mathbf{1}_p\widehat{\boldsymbol{\gamma}}'\boldsymbol{X}] = \mathbf{1}_p\boldsymbol{\gamma}'\boldsymbol{X}$. Moreover, from the expression of $\widehat{\boldsymbol{BC}}$ given in Theorem 3.1 and (5), it follows that

$$\begin{aligned}\mathbb{E}[\widehat{\boldsymbol{BC}}] &= \left(\mathbb{E}[\boldsymbol{Y}] - \frac{1}{p}\mathbf{1}_p\mathbf{1}_p'\mathbb{E}[\boldsymbol{Y}]\boldsymbol{P}_1\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\right)(\boldsymbol{R}_2\boldsymbol{K}_2'(\boldsymbol{K}_2\boldsymbol{K}_2')^- + \boldsymbol{G}\boldsymbol{R}_1\boldsymbol{K}_1'\boldsymbol{P}_2)\boldsymbol{C} \\ &= \left(\boldsymbol{BC} + \mathbf{1}_p\boldsymbol{\gamma}'\boldsymbol{X} - \frac{1}{p}\mathbf{1}_p\mathbf{1}_p'(\boldsymbol{BC} + \mathbf{1}_p\boldsymbol{\gamma}'\boldsymbol{X})\boldsymbol{P}_1\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\right) \\ &\quad \times (\boldsymbol{R}_2\boldsymbol{K}_2'(\boldsymbol{K}_2\boldsymbol{K}_2')^- + \boldsymbol{G}\boldsymbol{R}_1\boldsymbol{K}_1'\boldsymbol{P}_2)\boldsymbol{C} \\ &= \boldsymbol{BC}\big(\boldsymbol{R}_2\boldsymbol{K}_2'(\boldsymbol{K}_2\boldsymbol{K}_2')^- + \boldsymbol{G}\boldsymbol{R}_1\boldsymbol{K}_1'\boldsymbol{P}_2\big)\boldsymbol{C} \\ &= \boldsymbol{B}(\boldsymbol{K}_2\boldsymbol{K}_2'(\boldsymbol{K}_2\boldsymbol{K}_2')^- + \boldsymbol{K}_1\boldsymbol{K}_1'\boldsymbol{P}_2 - \boldsymbol{K}_2\boldsymbol{K}_2'(\boldsymbol{K}_2\boldsymbol{K}_2')^-\boldsymbol{K}_1\boldsymbol{K}_1'\boldsymbol{P}_2)\boldsymbol{C} \\ &= \boldsymbol{B}(\boldsymbol{K}_2\boldsymbol{K}_2'(\boldsymbol{K}_2\boldsymbol{K}_2')^- + \boldsymbol{I}_k - \boldsymbol{K}_2(\boldsymbol{K}_2'(\boldsymbol{K}_1\boldsymbol{K}_1')^{-1}\boldsymbol{K}_2)^-\boldsymbol{K}_2'(\boldsymbol{K}_1\boldsymbol{K}_1')^{-1} \\ &\quad - \boldsymbol{K}_2\boldsymbol{K}_2'(\boldsymbol{K}_2\boldsymbol{K}_2')^- + \boldsymbol{K}_2\boldsymbol{K}_2'(\boldsymbol{K}_2\boldsymbol{K}_2')^-\boldsymbol{K}_2(\boldsymbol{K}_2'(\boldsymbol{K}_1\boldsymbol{K}_1')^{-1}\boldsymbol{K}_2)^-\boldsymbol{K}_2'(\boldsymbol{K}_1\boldsymbol{K}_1')^{-1})\boldsymbol{C} \\ &= \boldsymbol{BC}.\end{aligned}$$

Hence, also $\widehat{\boldsymbol{BC}}$ is an unbiased estimator. Finally, we have for $\widetilde{\boldsymbol{U}}$ given in (8) the mean

$$\begin{aligned}\mathbb{E}[\widetilde{\boldsymbol{U}}\boldsymbol{Z}] &= (\boldsymbol{\Sigma}_e + \boldsymbol{\Sigma}_u)^{-1}\boldsymbol{\Sigma}_u\mathbb{E}[\boldsymbol{W}]\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{R}_1'\boldsymbol{Z}'\boldsymbol{Z} \\ &= (\boldsymbol{\Sigma}_e + \boldsymbol{\Sigma}_u)^{-1}\boldsymbol{\Sigma}_u\boldsymbol{B}(\boldsymbol{I}_k - \boldsymbol{K}_2\boldsymbol{K}_2'(\boldsymbol{K}_2\boldsymbol{K}_2')^-)\boldsymbol{K}_1\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{R}_1'\boldsymbol{Z}'\boldsymbol{Z},\end{aligned} \tag{A.1}$$

since from the expression of $\boldsymbol{W}$, it follows that

$$\mathbb{E}[\boldsymbol{W}] = \boldsymbol{BCGR}_1 = \boldsymbol{B}(\boldsymbol{I}_k - \boldsymbol{K}_2\boldsymbol{K}_2'(\boldsymbol{K}_2\boldsymbol{K}_2')^-)\boldsymbol{K}_1.$$

Recall that from the proof of Lemma 3.2, we have

$$\boldsymbol{K}_1\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o} = \boldsymbol{K}_2(\boldsymbol{K}_2'(\boldsymbol{K}_1\boldsymbol{K}_1')^{-1}\boldsymbol{K}_2)^-\boldsymbol{K}_2'(\boldsymbol{K}_1\boldsymbol{K}_1')^{-1}\boldsymbol{K}_1$$

and plugin this expression into $\mathbb{E}[\widetilde{\boldsymbol{U}}\boldsymbol{Z}]$, given in (A.1), leads to

$$\begin{aligned}\mathbb{E}[\widetilde{\boldsymbol{U}}\boldsymbol{Z}] &= (\boldsymbol{\Sigma}_e + \boldsymbol{\Sigma}_u)^{-1}\boldsymbol{\Sigma}_u\boldsymbol{B}(\boldsymbol{I}_k - \boldsymbol{K}_2\boldsymbol{K}_2'(\boldsymbol{K}_2\boldsymbol{K}_2')^-) \\ &\quad \times \boldsymbol{K}_2(\boldsymbol{K}_2'(\boldsymbol{K}_1\boldsymbol{K}_1')^{-1}\boldsymbol{K}_2)^-\boldsymbol{K}_2'(\boldsymbol{K}_1\boldsymbol{K}_1')^{-1}\boldsymbol{K}_1\boldsymbol{R}_1'\boldsymbol{Z}'\boldsymbol{Z} = \mathbf{0}.\end{aligned}$$

Hence, it is straightforward to see that

$$\mathbb{E}[\widehat{\boldsymbol{Y}}] = \mathbb{E}[\widehat{\boldsymbol{BC}} + \mathbf{1}_p\widehat{\boldsymbol{\gamma}}'\boldsymbol{X} + \widetilde{\boldsymbol{U}}\boldsymbol{Z}] = \boldsymbol{BC} + \mathbf{1}_p\boldsymbol{\gamma}'\boldsymbol{X} = \mathbb{E}[\boldsymbol{Y}],$$

which completes the proof of the theorem. □

In what follows, we use the notation $(\boldsymbol{A})\boldsymbol{Q}()'$ instead of $(\boldsymbol{A})\boldsymbol{Q}(\boldsymbol{A})'$ when it is possible and no confusions, in order to shorten the matrix expressions.

Proof of Theorem 5.2. Recall that $\widehat{\boldsymbol{\gamma}} = \frac{1}{p}(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{Y}'\mathbf{1}_p$. Hence, it follows that

$$\begin{aligned}\mathbb{D}[\widehat{\boldsymbol{\gamma}}] &= \frac{1}{p^2}(\mathbf{1}_p' \otimes (\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{P}_1)\mathbb{D}[\boldsymbol{Y}']()' \\ &= \frac{1}{p^2}(\mathbf{1}_p' \otimes (\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{P}_1)(\boldsymbol{\Sigma}_u \otimes \boldsymbol{Z}'\boldsymbol{Z} + \boldsymbol{\Sigma}_e \otimes \boldsymbol{I}_N)()' \\ &= \frac{\sigma_e^2}{p} \otimes ((\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}) = \frac{\sigma_e^2}{p}(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1},\end{aligned}$$

since $\boldsymbol{P}_1\boldsymbol{Z}' = \mathbf{0}$. Moreover, replacing $\boldsymbol{W}$ by its expression in the Theorem 3.1, we can rewrite $\widehat{\boldsymbol{BC}}$ by

$$\widehat{\boldsymbol{BC}} = \left(\boldsymbol{Y} - \frac{1}{p}\mathbf{1}_p\mathbf{1}_p'\boldsymbol{Y}\boldsymbol{P}_1\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\right)(\boldsymbol{R}_2\boldsymbol{K}_2'(\boldsymbol{K}_2\boldsymbol{K}_2')^{-} + \boldsymbol{G}\boldsymbol{R}_1\boldsymbol{K}_1'\boldsymbol{P}_2)\boldsymbol{C}$$

and therefore, the dispersion matrix $\mathbb{D}[\widehat{\boldsymbol{BC}}]$ is given by

$$\begin{aligned}\mathbb{D}[\widehat{\boldsymbol{BC}}] = \Big(&\boldsymbol{C}'((\boldsymbol{K}_2\boldsymbol{K}_2')^{-}\boldsymbol{K}_2\boldsymbol{R}_2' + \boldsymbol{P}_2\boldsymbol{K}_1\boldsymbol{R}_1'\boldsymbol{G}') \otimes \boldsymbol{I}_p \\ &- \frac{1}{p}\boldsymbol{C}'((\boldsymbol{K}_2\boldsymbol{K}_2')^{-}\boldsymbol{K}_2\boldsymbol{R}_2' + \boldsymbol{P}_2\boldsymbol{K}_1\boldsymbol{R}_1'\boldsymbol{G}')\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{P}_1 \otimes \mathbf{1}_p\mathbf{1}_p'\Big) \\ &\times (\boldsymbol{Z}'\boldsymbol{Z} \otimes \boldsymbol{\Sigma}_u + \boldsymbol{I}_N \otimes \boldsymbol{\Sigma}_e)()'.\end{aligned}$$

Using the results from Lemma 4.1, we obtain

$$\begin{aligned}\mathbb{D}[\widehat{\boldsymbol{BC}}] = \boldsymbol{C}'\boldsymbol{P}_2\boldsymbol{C} \otimes (\boldsymbol{\Sigma}_u + \boldsymbol{\Sigma}_e) &+ (\boldsymbol{C}'(\boldsymbol{I}_k - \boldsymbol{P}_2\boldsymbol{K}_1\boldsymbol{K}_1')(\boldsymbol{K}_2\boldsymbol{K}_2')^{-} \\ &\times (\boldsymbol{I}_k - \boldsymbol{K}_1\boldsymbol{K}_1'\boldsymbol{P}_2)\boldsymbol{C}) \otimes \boldsymbol{\Sigma}_e + \frac{\sigma_e^2}{p}\boldsymbol{C}'((\boldsymbol{K}_2\boldsymbol{K}_2')^{-}\boldsymbol{K}_2\boldsymbol{R}_2' + \boldsymbol{P}_2\boldsymbol{K}_1\boldsymbol{R}_1'\boldsymbol{G}') \\ &\times \boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}(\boldsymbol{R}_2\boldsymbol{K}_2'(\boldsymbol{K}_2\boldsymbol{K}_2')^{-} + \boldsymbol{G}\boldsymbol{R}_1\boldsymbol{K}_1'\boldsymbol{P}_2)\boldsymbol{C} \otimes \mathbf{1}_p\mathbf{1}_p'.\end{aligned}$$

Moreover, from $\widetilde{\boldsymbol{U}} = \boldsymbol{\Sigma}_u(\boldsymbol{\Sigma}_e + \boldsymbol{\Sigma}_u)^{-1}\boldsymbol{W}\boldsymbol{Q}_{K_1'K_2^o}\boldsymbol{R}_1'\boldsymbol{Z}'$, and replacing $\boldsymbol{W}$ by its expression, we can rewrite

$$\widetilde{\boldsymbol{U}} = \boldsymbol{\Sigma}_u(\boldsymbol{\Sigma}_e + \boldsymbol{\Sigma}_u)^{-1}\left(\boldsymbol{Y} - \frac{1}{p}\mathbf{1}_p\mathbf{1}_p'\boldsymbol{Y}\boldsymbol{P}_1\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\right)\boldsymbol{G}\boldsymbol{R}_1\boldsymbol{Q}_{K_1'K_2^o}\boldsymbol{R}_1'\boldsymbol{Z}'.$$

Given this, the dispersion matrix $\mathbb{D}[\widetilde{\boldsymbol{U}}]$ has the form

$$\begin{aligned}
\mathbb{D}[\widetilde{\boldsymbol{U}}] = &\Big(\boldsymbol{Z}\boldsymbol{R}_1\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{R}_1'\boldsymbol{G}' \otimes \boldsymbol{\Sigma}_u(\boldsymbol{\Sigma}_e + \boldsymbol{\Sigma}_u)^{-1} - \frac{1}{p}\boldsymbol{Z}\boldsymbol{R}_1\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{R}_1'\boldsymbol{G}'\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{P}_1 \\
&\otimes \boldsymbol{\Sigma}_u(\boldsymbol{\Sigma}_e + \boldsymbol{\Sigma}_u)^{-1}\boldsymbol{1}_p\boldsymbol{1}_p'\Big)(\boldsymbol{Z}'\boldsymbol{Z} \otimes \boldsymbol{\Sigma}_u + \boldsymbol{I}_N \otimes \boldsymbol{\Sigma}_e)()' \\
= &\boldsymbol{Z}\boldsymbol{R}_1\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}(\boldsymbol{I}_m + \boldsymbol{K}_1(\boldsymbol{K}_2\boldsymbol{K}_2')^{-}\boldsymbol{K}_1)\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{R}_1'\boldsymbol{Z}' \otimes \boldsymbol{\Sigma}_u(\boldsymbol{\Sigma}_e + \boldsymbol{\Sigma}_u)^{-1}\boldsymbol{\Sigma}_u \\
&+ \frac{\sigma_e^2}{p}\boldsymbol{Z}\boldsymbol{R}_1\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{R}_1'\boldsymbol{G}'\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{G}\boldsymbol{R}_1\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{R}_1'\boldsymbol{Z}' \\
&\otimes \boldsymbol{\Sigma}_u(\boldsymbol{\Sigma}_e + \boldsymbol{\Sigma}_u)^{-1}\boldsymbol{1}_p\boldsymbol{1}_p'(\boldsymbol{\Sigma}_e + \boldsymbol{\Sigma}_u)^{-1}\boldsymbol{\Sigma}_u.
\end{aligned}$$

□

Proof of Theorem 5.3. By using the results from Lemma 4.1, we get

$$\begin{aligned}
\mathrm{cov}[\widetilde{\boldsymbol{U}}, \widehat{\boldsymbol{\gamma}}] = &\Big(\boldsymbol{Z}\boldsymbol{R}_1\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{R}_1'\boldsymbol{G}' \otimes \boldsymbol{\Sigma}_u(\boldsymbol{\Sigma}_e + \boldsymbol{\Sigma}_u)^{-1} - \frac{1}{p}\boldsymbol{Z}\boldsymbol{R}_1\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{R}_1'\boldsymbol{G}'\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{P}_1 \\
&\otimes \boldsymbol{\Sigma}_u(\boldsymbol{\Sigma}_e + \boldsymbol{\Sigma}_u)^{-1}\boldsymbol{1}_p\boldsymbol{1}_p'\Big)(\boldsymbol{Z}'\boldsymbol{Z} \otimes \boldsymbol{\Sigma}_u + \boldsymbol{I}_N \otimes \boldsymbol{\Sigma}_e)\Big(\frac{1}{p}\boldsymbol{P}_1\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1} \otimes \boldsymbol{1}_p\Big) \\
= &-\frac{\sigma_e^2}{p}\boldsymbol{Z}\boldsymbol{R}_1\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{R}_1'\boldsymbol{G}'\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1} \otimes \boldsymbol{\Sigma}_u(\boldsymbol{\Sigma}_e + \boldsymbol{\Sigma}_u)^{-1}\boldsymbol{1}_p,
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{cov}[\widehat{\boldsymbol{B}}\boldsymbol{C}, \widetilde{\boldsymbol{U}}] = &\Big(\boldsymbol{C}'((\boldsymbol{K}_2\boldsymbol{K}_2')^{-}\boldsymbol{K}_2\boldsymbol{R}_2' + \boldsymbol{P}_2\boldsymbol{K}_1\boldsymbol{R}_1'\boldsymbol{G}') \otimes \boldsymbol{I}_p \\
&- \frac{1}{p}\boldsymbol{C}'((\boldsymbol{K}_2\boldsymbol{K}_2')^{-}\boldsymbol{K}_2\boldsymbol{R}_2' + \boldsymbol{P}_2\boldsymbol{K}_1\boldsymbol{R}_1'\boldsymbol{G}')\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{P}_1 \otimes \boldsymbol{1}_p\boldsymbol{1}_p'\Big) \\
&\times (\boldsymbol{Z}'\boldsymbol{Z} \otimes \boldsymbol{\Sigma}_u + \boldsymbol{I}_N \otimes \boldsymbol{\Sigma}_e)\Big(\boldsymbol{G}\boldsymbol{R}_1\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{R}_1'\boldsymbol{Z}' \otimes \boldsymbol{\Sigma}_u(\boldsymbol{\Sigma}_e + \boldsymbol{\Sigma}_u)^{-1} \\
&- \frac{1}{p}\boldsymbol{P}_1\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{G}\boldsymbol{R}_1\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{R}_1'\boldsymbol{Z}' \otimes \boldsymbol{1}_p\boldsymbol{1}_p'(\boldsymbol{\Sigma}_e + \boldsymbol{\Sigma}_u)^{-1}\boldsymbol{\Sigma}_u\Big) \\
= &\frac{\sigma_e^2}{p}\boldsymbol{C}'((\boldsymbol{K}_2\boldsymbol{K}_2')^{-}\boldsymbol{K}_2\boldsymbol{R}_2' + \boldsymbol{P}_2\boldsymbol{K}_1\boldsymbol{R}_1'\boldsymbol{G}')\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{G}\boldsymbol{R}_1\boldsymbol{Q}_{\boldsymbol{K}_1'\boldsymbol{K}_2^o}\boldsymbol{R}_1'\boldsymbol{Z}' \\
&\otimes \boldsymbol{1}_p\boldsymbol{1}_p'(\boldsymbol{\Sigma}_e + \boldsymbol{\Sigma}_u)^{-1}\boldsymbol{\Sigma}_u.
\end{aligned}$$

Similarly we have the last covariance as

$$\begin{aligned}
\mathrm{cov}[\widehat{\boldsymbol{B}}\boldsymbol{C}, \widehat{\boldsymbol{\gamma}}] = &\Big(\boldsymbol{C}'((\boldsymbol{K}_2\boldsymbol{K}_2')^{-}\boldsymbol{K}_2\boldsymbol{R}_2' + \boldsymbol{P}_2\boldsymbol{K}_1\boldsymbol{R}_1'\boldsymbol{G}') \otimes \boldsymbol{I}_p \\
&- \frac{1}{p}\boldsymbol{C}'((\boldsymbol{K}_2\boldsymbol{K}_2')^{-}\boldsymbol{K}_2\boldsymbol{R}_2' + \boldsymbol{P}_2\boldsymbol{K}_1\boldsymbol{R}_1'\boldsymbol{G}')\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{P}_1 \otimes \boldsymbol{1}_p\boldsymbol{1}_p'\Big) \\
&\times (\boldsymbol{Z}'\boldsymbol{Z} \otimes \boldsymbol{\Sigma}_u + \boldsymbol{I}_N \otimes \boldsymbol{\Sigma}_e)\Big(\frac{1}{p}\boldsymbol{P}_1\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1} \otimes \boldsymbol{1}_p\Big) \\
= &-\frac{\sigma_e^2}{p}\boldsymbol{C}'((\boldsymbol{K}_2\boldsymbol{K}_2')^{-}\boldsymbol{K}_2\boldsymbol{R}_2' + \boldsymbol{P}_2\boldsymbol{K}_1\boldsymbol{R}_1'\boldsymbol{G}')\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{P}_1\boldsymbol{X}')^{-1} \otimes \boldsymbol{1}_p.
\end{aligned}$$

□