CIRJE-F-832

# Mixed Effects Prediction under Benchmarking and Applications to Small Area Estimation

Tatsuya Kubokawa
University of Tokyo

January 2012

# Mixed Effects Prediction under Benchmarking and Applications to Small Area Estimation

Tatsuya Kubokawa*
*University of Tokyo*

January 3, 2012

**Abstract**

The empirical best linear unbiased predictor (EBLUP) in the linear mixed model (LMM) is useful for the small area estimation in the sense of increasing the precision of estimation of small area means. However, one potential difficulty of EBLUP is that when aggregated, the overall estimate for a larger geographical area may be quite different from the corresponding direct estimate like the overall sample mean. One way to solve this problem is the benchmarking approach, and the constrained EBLUP is a feasible solution which satisfies the constraints that the aggregated mean and variance are identical to the requested values of mean and variance. An interesting query is whether the constrained EBLUP may have a larger estimation error than EBLUP. In this paper, we address this issue by deriving asymptotic approximations of MSE of the constrained EBLUP. Also, we provide asymptotic unbiased estimators of the MSE of the constrained EBLUP based on the parametric bootstrap method, and establish their second-order justification. Finally, the performances of the suggested MSE estimators are numerically investigated.

*Key words and phrases:* Benchmarking, best linear unbiased predictor, constrained Bayes, empirical Bayes, linear mixed model, mean squared error, parametric bootstrap, second-order approximation, small area estimation.

## 1 Introduction

The linear mixed models (LMM) and the model-based estimates including empirical best linear unbiased predictor (EBLUP) or the empirical Bayes estimator (EB) have been recognized useful in small area estimation. The typical models used for the small area estimation are the Fay-Herriot model and the nested error regression model (NERM), and the usefulness of EBLUP is illustrated by Fay and Herriot (1979) and Battese, Harter and Fuller (1988). For a good review and account on this topic, see Ghosh and Rao (1994), Rao (2003) and Pfeffermann (2002).

*Department of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, JAPAN
E-Mail: tatsuya@e.u-tokyo.ac.jp

One potential difficulty of EBLUP is that when aggregated, the overall estimate for a larger geographical area may be quite different from the corresponding direct estimate like the overall sample mean. One way to solve this problem is the benchmarking approach, which modifies EBLUP so that one gets the same aggregate mean and/or variance for the larger geographical area. Ghosh (1992) suggested the constrained Bayes estimator or the constrained EBLUP which satisfy the constraints that the aggregated mean and variance are identical to the mean and variance of the posterior distribution, and Datta, Ghosh, Steorts and Maples (2011) gave some extensions. Since the sample variance of EBLUP is smaller than the posterior variance, the constrained EBLUP modifies EBLUP so that its sample variance is identical to the posterior variance. However, the usefulness and purpose of EBLUP is that EBLUP gives stable estimates with higher precision of estimation. We then have a concern whether the constrained EBLUP may be against this purpose. Thus, it is quite interesting and important to assess the mean squared error (MSE) of the constrained EBLUP. In this paper, we address this issue for the general constrained EBLUP in the general linear mixed model.

In Section 2, we consider the general constraint on the mean and variance of estimators and derive the general constrained estimators including the constrained EBLUP, which is an extension of the constrained empirical Bayes estimators given by Ghosh (1992) and Datta, *et al*. (2011). In Section 3, we derive the asymptotic approximations of MSE of the constrained EBLUP. When the variance constraint is the posterior variance, it is shown that MSE of the constrained EBLUP is larger than MSE of EBLUP in the first order approximation. To modify this property, we suggest some modification of the variance constraint. We also provide an asymptotically unbiased estimator of MSE of the constrained EBLUP based on the parametric bootstrap method, and establish the second-order justification. In Section 4, we investigate the performances of MSE of the constrained EBLUP and the MSE estimators. Section 5 gives a concluding remark. The proofs of the asymptotic approximations are given in the appendix.

# 2 Benchmarking EBLUP

## 2.1 Linear mixed model and constraints

Consider the general linear mixed model

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{v} + \boldsymbol{\epsilon}, \tag{2.1}$$

where $\boldsymbol{y}$ is an $N \times 1$ observation vector of the response variable, $\boldsymbol{X}$ and $\boldsymbol{Z}$ are $N \times p$ and $N \times M$ matrices, respectively, of the explanatory variables, $\boldsymbol{\beta}$ is a $p \times 1$ unknown vector of the regression coefficients, $\boldsymbol{v}$ is an $M \times 1$ vector of the random effects, and $\boldsymbol{\epsilon}$ is an $N \times 1$ vector of the random errors. Here, $\boldsymbol{v}$ and $\boldsymbol{\epsilon}$ are mutually independently distributed as $\boldsymbol{v} \sim \mathcal{N}_M(0, \boldsymbol{Q})$ and $\boldsymbol{\epsilon} \sim \mathcal{N}_N(0, \boldsymbol{R})$ where $\boldsymbol{Q}$ and $\boldsymbol{R}$ are positive definite matrices. Then, $\boldsymbol{y}$ has a marginal distribution $\mathcal{N}_N(\boldsymbol{X}\boldsymbol{\beta}, \boldsymbol{\Sigma})$ for $\boldsymbol{\Sigma} = \boldsymbol{R} + \boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'$. It is assumed that $\boldsymbol{Q}$ and $\boldsymbol{R}$ are functions of unknown parameters $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_q)'$, namely, $\boldsymbol{Q} = \boldsymbol{Q}(\boldsymbol{\psi})$, $\boldsymbol{R} = \boldsymbol{R}(\boldsymbol{\psi})$ and $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\psi})$. Denote unknown parameters by $\boldsymbol{\omega} = (\boldsymbol{\psi}', \boldsymbol{\beta}')'$.

Let $\boldsymbol{\theta} = \boldsymbol{X\beta} + \boldsymbol{Zv}$ and let $\widehat{\boldsymbol{\theta}} = \widehat{\boldsymbol{\theta}}(\boldsymbol{y})$ be a predictor of $\boldsymbol{\theta}$. Suppose that $\widehat{\boldsymbol{\theta}}$ is evaluated relative to the quadratic loss function $\|\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\|^2_{\boldsymbol{\Omega}} = (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})'\boldsymbol{\Omega}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ for known positive definite matrix $\boldsymbol{\Omega}$. In the Bayesian framework, $\boldsymbol{\theta}$ has a prior distribution $\mathcal{N}_N(\boldsymbol{X\beta}, \boldsymbol{ZQZ}')$, and the posterior distribution of $\boldsymbol{\theta}$ given $\boldsymbol{y}$ is

$$\boldsymbol{\theta}|\boldsymbol{y} \sim \mathcal{N}_N\big(\boldsymbol{X\beta} + \boldsymbol{ZQZ}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{X\beta}), (\boldsymbol{Q}^{-1} + \boldsymbol{Z}'\boldsymbol{R}^{-1}\boldsymbol{Z})^{-1}\big), \tag{2.2}$$

which gives the Bayes estimator

$$\widehat{\boldsymbol{\theta}}^B = \widehat{\boldsymbol{\theta}}^B(\boldsymbol{\omega}) = \boldsymbol{X\beta} + \boldsymbol{ZQZ}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{X\beta}). \tag{2.3}$$

As demonstrated by Louis (1984) and Ghosh (1992), the Bayes estimator is shrunken too much and we need to consider the constrained Bayes estimators. In this paper, we treat more general constraints on estimator $\widehat{\boldsymbol{\theta}}$.

$$\textbf{(C)}\begin{cases} \text{(C1)} \quad \boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}} = \boldsymbol{t}_1(\boldsymbol{y}) \text{ for } N \times L \text{ known matrix } \boldsymbol{W} \text{ and } L\text{-variate function } \boldsymbol{t}_1(\boldsymbol{y}), \\ \text{(C2)} \quad \|\widehat{\boldsymbol{\theta}} - \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}\|^2_{\boldsymbol{\Omega}} = t_2(\boldsymbol{y}) \text{ for function } t_2(\boldsymbol{y}). \end{cases}$$

Datta, *et al.* (2011) dealt with the general linear constraint (C1), but a simple constraint for (C2). The estimator $\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}$ corresponds to the generalized least square estimator of $\boldsymbol{\xi}$ under the loss $\|\widehat{\boldsymbol{\theta}} - \boldsymbol{W\xi}\|^2_{\boldsymbol{\Omega}}$, and $\|\widehat{\boldsymbol{\theta}} - \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}\|^2_{\boldsymbol{\Omega}}$ is the residual variance. It is noted that the constraint (C2) is also expressed as

$$\|\widehat{\boldsymbol{\theta}} - \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}\|^2_{\boldsymbol{\Omega}} = \widehat{\boldsymbol{\theta}}'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}} = t_2(\boldsymbol{y}), \tag{2.4}$$

where

$$\boldsymbol{P}_{\boldsymbol{\Omega}} = \boldsymbol{\Omega} - \boldsymbol{\Omega W}(\boldsymbol{W}'\boldsymbol{\Omega W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}.$$

For these general constraints, we derive the constrained Bayes estimators.

**Example 2.1** A typical example of $\boldsymbol{W}$ is $\boldsymbol{w} = (w_1, \ldots, w_N)'$ for nonnegative constants $w_i$'s, and in the case of $\boldsymbol{\Omega} = \boldsymbol{I}$, the linear combination in (C1) is the combined mean $\boldsymbol{w}'\widehat{\boldsymbol{\theta}} = \sum_{i=1}^N w_i\hat{\theta}_i$ for $\widehat{\boldsymbol{\theta}} = (\hat{\theta}_1, \ldots, \hat{\theta}_N)'$. For instance, let $t_1(\boldsymbol{y}) = \sum_{i=1}^N w_i y_i$ for $\boldsymbol{y} = (y_1, \ldots, y_N)'$, and the constraint (C1) is expressed as $\sum_{i=1}^N w_i\hat{\theta}_i = \sum_{i=1}^N w_i y_i$.

As another example, consider the case that the whole area is divided into $G$ groups and the $i$-th group consists of $m_i$ small areas. When we can consider the benchmarking for each group, $\boldsymbol{W}$ is given by

$$\boldsymbol{W} = \text{block diag}(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_G), \quad \boldsymbol{w}_i = (w_{i1}, \ldots, w_{im_i})',$$

where $w_{ij}$'s are nonnegative constants. Also, $\widehat{\boldsymbol{\theta}}$ is decomposed as $\widehat{\boldsymbol{\theta}} = (\widehat{\boldsymbol{\theta}}_1', \ldots, \widehat{\boldsymbol{\theta}}_G)'$ for $\widehat{\boldsymbol{\theta}}_i = (\hat{\theta}_{i1}, \ldots, \hat{\theta}_{im_i})'$. Then the linear combination in the constraint (C1) is

$$\boldsymbol{W}'\widehat{\boldsymbol{\theta}} = ((\boldsymbol{w}_1'\widehat{\boldsymbol{\theta}}_1)', \ldots, (\boldsymbol{w}_G'\widehat{\boldsymbol{\theta}}_G)')'.$$

For instance, let $\boldsymbol{t}_1(\boldsymbol{y}) = (\boldsymbol{w}_1'\boldsymbol{y}_1, \ldots, \boldsymbol{w}_G'\boldsymbol{y}_G)'$, and the constraint (C1) is expressed as $\boldsymbol{w}_i'\widehat{\boldsymbol{\theta}}_i = \boldsymbol{w}_i'\boldsymbol{y}_i$, $i = 1, \ldots, G$, where $\boldsymbol{y} = (\boldsymbol{y}_1', \ldots, \boldsymbol{y}_G')'$ is decomposed similarly to $\widehat{\boldsymbol{\theta}}$.

The above setup is applicable to a model for analysis of cross-sectional and time-series data. Let $y_{it}$ be an observation at the $i$-th area and $t$ time point for $i = 1, \ldots, K$ and $t = 1, \ldots, T$. Let $\boldsymbol{y} = (\boldsymbol{y}_1', \ldots, \boldsymbol{y}_T')'$ for $\boldsymbol{y}_t = (y_{1t}, \ldots, y_{kt})'$. Putting $m_i = K$ and $G = T$ in the above setup, we get the benchmarking at each time point. ■

We here verify that the same phenomenon as demonstrated in Louis (1984) and Ghosh (1992) holds for the above general situations.

**Proposition 2.1** *The following relationships hold between the Bayes estimator* $\widehat{\boldsymbol{\theta}}^B$ *and the posterior distribution: For the linear transformation,* $E[\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{\theta}|\boldsymbol{y}] = \boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}^B$, *but for the residual variance,*

$$\begin{aligned} &E[\|\boldsymbol{\theta} - \boldsymbol{W}(\boldsymbol{W}'\Omega\boldsymbol{W})^{-1}\boldsymbol{W}'\Omega\boldsymbol{\theta}\|_{\Omega}^2|\boldsymbol{y}] \\ &=\|\widehat{\boldsymbol{\theta}}^B - \boldsymbol{W}(\boldsymbol{W}'\Omega\boldsymbol{W})^{-1}\boldsymbol{W}'\Omega\widehat{\boldsymbol{\theta}}^B\|_{\Omega}^2 + \mathrm{tr}\left[\boldsymbol{Z}'\boldsymbol{P}_{\Omega}\boldsymbol{Z}\{\boldsymbol{Q}^{-1} + \boldsymbol{Z}'\boldsymbol{R}^{-1}\boldsymbol{Z}\}^{-1}\right] \\ &\geq\|\widehat{\boldsymbol{\theta}}^B - \boldsymbol{W}(\boldsymbol{W}'\Omega\boldsymbol{W})^{-1}\boldsymbol{W}'\Omega\widehat{\boldsymbol{\theta}}^B\|_{\Omega}^2, \end{aligned} \tag{2.5}$$

*where* $E[\cdot|\boldsymbol{y}]$ *denotes a posterior expectation.*

In fact, it is noted that

$$E[\boldsymbol{\theta}\boldsymbol{\theta}'|\boldsymbol{y}] = \widehat{\boldsymbol{\theta}}^B(\widehat{\boldsymbol{\theta}}^B)' + E\left[\boldsymbol{Z}(\boldsymbol{v} - E[\boldsymbol{v}|\boldsymbol{y}])(\boldsymbol{v} - E[\boldsymbol{v}|\boldsymbol{y}])'\boldsymbol{Z}'|\boldsymbol{y}\right],$$

since $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}^B + \boldsymbol{Z}(\boldsymbol{v} - E[\boldsymbol{v}|\boldsymbol{y}])$. From the posterior distribution (2.2), it follows that $E\left[(\boldsymbol{v} - E[\boldsymbol{v}|\boldsymbol{y}])(\boldsymbol{v} - E[\boldsymbol{v}|\boldsymbol{y}])'|\boldsymbol{y}\right] = (\boldsymbol{Q}^{-1} + \boldsymbol{Z}'\boldsymbol{R}^{-1}\boldsymbol{Z})^{-1}$. Noting the expression (2.4), we have

$$E[\boldsymbol{\theta}'\boldsymbol{P}_{\Omega}\boldsymbol{\theta}|\boldsymbol{y}] = \mathrm{tr}\,[\boldsymbol{P}_{\Omega}\boldsymbol{\theta}\boldsymbol{\theta}'|\boldsymbol{y}] = \mathrm{tr}\,[\boldsymbol{P}_{\Omega}\widehat{\boldsymbol{\theta}}^B(\widehat{\boldsymbol{\theta}}^B)'] + \mathrm{tr}\,[\boldsymbol{P}_{\Omega}(\boldsymbol{Q}^{-1} + \boldsymbol{Z}'\boldsymbol{R}^{-1}\boldsymbol{Z})^{-1}], \tag{2.6}$$

which implies the equality in (2.5).

Proposition 2.1 shows that

$$E[\|\boldsymbol{\theta} - \boldsymbol{W}(\boldsymbol{W}'\Omega\boldsymbol{W})^{-1}\boldsymbol{W}'\Omega\boldsymbol{\theta}\|_{\Omega}^2] \geq E[\|\widehat{\boldsymbol{\theta}}^B - \boldsymbol{W}(\boldsymbol{W}'\Omega\boldsymbol{W})^{-1}\boldsymbol{W}'\Omega\widehat{\boldsymbol{\theta}}^B\|_{\Omega}^2],$$

namely, for any linear transformation matrix $\boldsymbol{W}$, the expectation of the sample residual variance of the Bayes estimators is less than the expected residual variance of the unobserved parameters. Louis (1984) and Ghosh (1992) showed this fact in the case of $\boldsymbol{W} = \boldsymbol{j}_N = (1, \ldots, 1)'$, and pointed out that the Bayes estimator is shrunken too much since the variance of the Bayes estimator is smaller than the variance of the prior distribution. This gives a motivation of the constrained Bayes estimators under the constraints (C1) and (C2).

## 2.2 Unified constrained estimators

We now derive the constrained Bayes estimators under the constraints (C1) and (C2).

**Theorem 2.1** *The constrained Bayes estimator under the constraint* (C1) *is*

$$\widehat{\boldsymbol{\theta}}^{CB1} = \widehat{\boldsymbol{\theta}}^B - \boldsymbol{W}(\boldsymbol{W}'\Omega\boldsymbol{W})^{-1}\left\{\boldsymbol{W}'\Omega\widehat{\boldsymbol{\theta}}^B - \boldsymbol{t}_1(\boldsymbol{y})\right\}, \tag{2.7}$$

*where* $\widehat{\boldsymbol{\theta}}^B$ *is the Bayes estimator given in* (2.3)*. The constrained Bayes estimator under the constraints* (C1) *and* (C2) *is*

$$\widehat{\boldsymbol{\theta}}^{CB2} = a^B(\boldsymbol{y})\{\boldsymbol{I} - \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\}\widehat{\boldsymbol{\theta}}^B + \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{t}_1(\boldsymbol{y}), \tag{2.8}$$

*where*

$$\{a^B(\boldsymbol{y})\}^2 = \frac{t_2(\boldsymbol{y})}{(\widehat{\boldsymbol{\theta}}^B)'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^B}.$$

This theorem provides not only general constrained estimators, but also a unified expression for the constrained estimators. For an estimator $\widehat{\boldsymbol{\theta}}$, we can treat constrained estimators as the following unified form:

$$\widehat{\boldsymbol{\theta}}^{CG} = \frac{t_2(\boldsymbol{y})}{\widehat{\boldsymbol{\theta}}'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}}\{\boldsymbol{I} - \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\}\widehat{\boldsymbol{\theta}} + \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{t}_1(\boldsymbol{y}). \tag{2.9}$$

In fact, this class of the estimators includes $\widehat{\boldsymbol{\theta}}^{CB1}$ and $\widehat{\boldsymbol{\theta}}^{CB2}$. When $\boldsymbol{t}_1(\boldsymbol{y}) = \boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}$ and $t_2(\boldsymbol{y}) = \|\widehat{\boldsymbol{\theta}} - \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}\|_{\boldsymbol{\Omega}}^2 = \widehat{\boldsymbol{\theta}}'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}$, it is seen that $\widehat{\boldsymbol{\theta}}^{CG} = \widehat{\boldsymbol{\theta}}$. That is, the constraints do not give any change if the least squares statistic and the residual variance of the constrained estimator in (C1) and (C2) are the same to those of the original estimator.

**Proof of Theorem 2.1.** The constrained Bayes estimator $\widehat{\boldsymbol{\theta}}^{CB1}$ follows from Datta, *et al*. (2011). We shall derive the constrained Bayes estimator $\widehat{\boldsymbol{\theta}}^{CB2}$ under the constraints (C1) and (C2). Ghosh (1992) and Datta, *et al*. (2011) provided such a constrained Bayes estimator where they treated a constraint simpler than (C2). We here obtain $\widehat{\boldsymbol{\theta}}^{CB2}$ directly in a different way as well as in more general constraint (C2).

Note that $E[\|\boldsymbol{\theta}-\widehat{\boldsymbol{\theta}}\|_{\boldsymbol{\Omega}}^2|\boldsymbol{y}] = E[\|\boldsymbol{\theta}-\widehat{\boldsymbol{\theta}}^B\|_{\boldsymbol{\Omega}}^2|\boldsymbol{y}] + \|\widehat{\boldsymbol{\theta}}-\widehat{\boldsymbol{\theta}}^B\|_{\boldsymbol{\Omega}}^2$. Then, the constrained Bayes estimator can be given as a solution on $\widehat{\boldsymbol{\theta}}$ which minimizes the Lagrangian multiplier

$$H(\widehat{\boldsymbol{\theta}}) = \|\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}^B\|_{\boldsymbol{\Omega}}^2 + \boldsymbol{\lambda}_1'\{\boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}} - \boldsymbol{t}_1(\boldsymbol{y})\} + \lambda_2\{\widehat{\boldsymbol{\theta}}'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}} - t_2(\boldsymbol{y})\},$$

where $\boldsymbol{\lambda}_1$ is an $L \times 1$ multiplier and $\lambda_2$ is a scalar multiplier. Differentiating $H(\widehat{\boldsymbol{\theta}})$ with respect to $\widehat{\boldsymbol{\theta}}$ gives $\boldsymbol{\Omega}(\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}^B) + \frac{1}{2}\boldsymbol{\Omega}\boldsymbol{W}\boldsymbol{\lambda}_1 + \lambda_2\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}} = 0$, which can be rewritten as $\widehat{\boldsymbol{\theta}} = \{(1+\lambda_2)\boldsymbol{\Omega} - \lambda_2\boldsymbol{P}_{\boldsymbol{\Omega}}\}^{-1}\boldsymbol{\Omega}(\widehat{\boldsymbol{\theta}}^B - \frac{1}{2}\boldsymbol{W}\boldsymbol{\lambda}_1)$. It is here noted that $\{(1+\lambda_2)\boldsymbol{\Omega} - \lambda_2\boldsymbol{P}_{\boldsymbol{\Omega}}\}^{-1} = \boldsymbol{\Omega}^{-1} - \lambda_2(1+\lambda_2)^{-1}\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\boldsymbol{\Omega}^{-1}$. Since $\boldsymbol{W}'\boldsymbol{P}_{\boldsymbol{\Omega}} = 0$, $\widehat{\boldsymbol{\theta}}$ can be expressed as

$$\widehat{\boldsymbol{\theta}} = \left(\boldsymbol{I} - \frac{\lambda_2}{1+\lambda_2}\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\right)\widehat{\boldsymbol{\theta}}^B - \frac{1}{2}\boldsymbol{W}\boldsymbol{\lambda}_1. \tag{2.10}$$

Substituting $\widehat{\boldsymbol{\theta}}$ given in (2.10) into the constraint (C1) or $\boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}} = \boldsymbol{t}_1(\boldsymbol{y})$, we get the equality $-0.5\boldsymbol{\lambda}_1 = (\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\{\boldsymbol{t}_1(\boldsymbol{y}) - \boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}^B\}$, which is again substituted into (2.10) to yield

$$\widehat{\boldsymbol{\theta}} = \left(\boldsymbol{I} - \frac{\lambda_2}{1+\lambda_2}\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\right)\widehat{\boldsymbol{\theta}}^B + \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\{\boldsymbol{t}_1(\boldsymbol{y}) - \boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}^B\}. \tag{2.11}$$

Substituting $\widehat{\boldsymbol{\theta}}$ given in (2.11) into the constraint (C2) or $\|\widehat{\boldsymbol{\theta}} - \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}\|_{\boldsymbol{\Omega}}^2 = t_2(\boldsymbol{y})$, we get

$$\Big\{(\boldsymbol{I} - \frac{\lambda_2}{1+\lambda_2}\boldsymbol{\Omega}^{-1}\boldsymbol{P_\Omega})\widehat{\boldsymbol{\theta}}^B + \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\{\boldsymbol{t}_1(\boldsymbol{y}) - \boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}^B\}\Big\}'\boldsymbol{P_\Omega}$$
$$\Big\{(\boldsymbol{I} - \frac{\lambda_2}{1+\lambda_2}\boldsymbol{\Omega}^{-1}\boldsymbol{P_\Omega})\widehat{\boldsymbol{\theta}}^B + \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\{\boldsymbol{t}_1(\boldsymbol{y}) - \boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}^B\}\Big\} = t_2(\boldsymbol{y}).$$

Noting again that $\boldsymbol{W}'\boldsymbol{P_\Omega} = 0$, this equality can be simplified as

$$(\widehat{\boldsymbol{\theta}}^B)'\left(\boldsymbol{I} - \frac{\lambda_2}{1+\lambda_2}\boldsymbol{P_\Omega}\boldsymbol{\Omega}^{-1}\right)\boldsymbol{P_\Omega}\left(\boldsymbol{I} - \frac{\lambda_2}{1+\lambda_2}\boldsymbol{\Omega}^{-1}\boldsymbol{P_\Omega}\right)\widehat{\boldsymbol{\theta}}^B = t_2(\boldsymbol{y}),$$

or $(1+\lambda_2)^2 = (\widehat{\boldsymbol{\theta}}^B)'\boldsymbol{P_\Omega}\widehat{\boldsymbol{\theta}}^B/t_2(\boldsymbol{y})$, since $(\boldsymbol{\Omega}^{-1/2}\boldsymbol{P_\Omega}\boldsymbol{\Omega}^{-1/2})^2 = \boldsymbol{\Omega}^{-1/2}\boldsymbol{P_\Omega}\boldsymbol{\Omega}^{-1/2}$. Let

$$a^B(\boldsymbol{y}) = \sqrt{t_2(\boldsymbol{y})/(\widehat{\boldsymbol{\theta}}^B)'\boldsymbol{P_\Omega}\widehat{\boldsymbol{\theta}}^B}.$$

Since $\widehat{\boldsymbol{\theta}}$ given in (2.11) can be rewritten as

$$\widehat{\boldsymbol{\theta}} = \frac{1}{1+\lambda_2}\{\boldsymbol{I} - \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{\Omega}\}\widehat{\boldsymbol{\theta}}^B + \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}\boldsymbol{t}_1(\boldsymbol{y}),$$

substituting $a^B(\boldsymbol{y})$ into this estimator gives the constrained Bayes estimator $\widehat{\boldsymbol{\theta}}^{CB2}$. ■

## 2.3 Benchmarked EBLUP

We shall treat the problem of predicting $\mu = \boldsymbol{c}'\boldsymbol{\theta} = \boldsymbol{c}'\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{c}'\boldsymbol{Z}\boldsymbol{v}$ where $\boldsymbol{c}$ is an $N \times 1$ vector. Typical example of $\boldsymbol{c}$ is that components of $\boldsymbol{c}$ are constants for an area of interest and zeros for other areas. Let $\widehat{\boldsymbol{\psi}} = (\widehat{\psi}_1, \ldots, \widehat{\psi}_q)'$ be a consistent estimator of $\boldsymbol{\psi}$. Also, let $\widehat{\boldsymbol{Q}} = \boldsymbol{Q}(\widehat{\boldsymbol{\psi}})$, $\widehat{\boldsymbol{R}} = \boldsymbol{R}(\widehat{\boldsymbol{\psi}})$ and $\widehat{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}(\widehat{\boldsymbol{\psi}}) = \widehat{\boldsymbol{R}} + \boldsymbol{Z}\widehat{\boldsymbol{Q}}\boldsymbol{Z}'$. The generalized least squares estimator of $\boldsymbol{\beta}$ is

$$\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\psi}}) = (\boldsymbol{X}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{y}. \tag{2.12}$$

Since the EBLUP or empirical Bayes estimator of $\boldsymbol{\theta}$ is

$$\widehat{\boldsymbol{\theta}}^{EB} = \boldsymbol{X}\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\psi}}) + \boldsymbol{Z}\widehat{\boldsymbol{Q}}\boldsymbol{Z}'\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\psi}})), \tag{2.13}$$

the benchmarked EBLUP is given by

$$\widehat{\mu}^{CEB} = \boldsymbol{c}'\widehat{\boldsymbol{\theta}}^{CG} = a(\boldsymbol{y})\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P_\Omega}\widehat{\boldsymbol{\theta}}^{EB} + \boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{t}_1(\boldsymbol{y}), \tag{2.14}$$

where $\boldsymbol{c}_{\boldsymbol{w}}' = \boldsymbol{c}'\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}$ and

$$\{a(\boldsymbol{y})\}^2 = \frac{t_2(\boldsymbol{y})}{(\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P_\Omega}\widehat{\boldsymbol{\theta}}^{EB}}. \tag{2.15}$$

The benchmarked EBLUP can be rewritten as

$$\widehat{\mu}^{CEB} = \widehat{\mu}^{EB} + \{a(\boldsymbol{y}) - 1\}\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P_\Omega}\widehat{\boldsymbol{\theta}}^{EB} + \boldsymbol{c}_{\boldsymbol{w}}'\{\boldsymbol{t}_1(\boldsymbol{y}) - \boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}^{EB}\}, \tag{2.16}$$

for $\widehat{\mu}^{EB} = \boldsymbol{c}'\widehat{\boldsymbol{\theta}}^{EB}$.

It is noted from (2.2) and (2.6) that the regressed posterior mean $E[\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{\theta}|\boldsymbol{y}]$ and the residual variance $E[\|\boldsymbol{\theta} - \boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{\theta}\|_{\boldsymbol{\Omega}}^2|\boldsymbol{y}]$ of $\boldsymbol{\theta}$ are estimated by

$$\begin{aligned} \boldsymbol{t}_1^{(0)}(\boldsymbol{y}) =& \boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}^{EB}, \\ t_2^{(0)}(\boldsymbol{y}) =& (\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB} + \mathrm{tr}\,[\boldsymbol{P}_{\boldsymbol{\Omega}}(\widehat{\boldsymbol{Q}}^{-1} + \boldsymbol{Z}'\widehat{\boldsymbol{R}}^{-1}\boldsymbol{Z})^{-1}]. \end{aligned} \tag{2.17}$$

Then, it may be reasonable to put $\boldsymbol{t}_1(\boldsymbol{y}) = \boldsymbol{t}_1^{(0)}(\boldsymbol{y})$ and $t_2(\boldsymbol{y}) = t_2^{(0)}(\boldsymbol{y})$ for the constraints (C1) and (C2). The resulting constrained estimator $\widehat{\mu}^{CEB}$ has more variability than $\widehat{\mu}^{EB}$. On the other hand, the aim of EBLUP $\widehat{\mu}^{EB}$ is to give an estimate with higher precision, and there is a trade-off between the benchmarking and the aim of EBLUP. Thus, it is important how to set up the variability $t_2(\boldsymbol{y})$ of $\widehat{\mu}^{CEB}$. One of reasonable requirement for variability is that the constrained estimator $\widehat{\mu}^{CEB}$ satisfies the property

$$\lim_{N\to\infty} MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB}) = \lim_{N\to\infty} MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB}), \tag{2.18}$$

namely, MSE of $\widehat{\mu}^{CEB}$ is equal to that of $\widehat{\mu}^{EB}$ in the first-order $O(1)$. It is here noted that MSE of $\widehat{\mu}^{EB}$ does not depend on $\boldsymbol{\beta}$ from the assumption (A3). Unfortunately, the constrained estimator with $t_2(\boldsymbol{y}) = t_2^{(0)}(\boldsymbol{y})$ does not satisfy the requirement (2.18), namely,

$$\lim_{N\to\infty} MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB}) > \lim_{N\to\infty} MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB}).$$

This fact is against the aim of EBLUP. For (C2), it is interesting to consider the constraint $t_2(\boldsymbol{y}) = t_2^{(r)}(\boldsymbol{y})$ where

$$t_2^{(r)}(\boldsymbol{y}) = (\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB} + \frac{1}{N^r}\mathrm{tr}\,[\boldsymbol{P}_{\boldsymbol{\Omega}}(\widehat{\boldsymbol{Q}}^{-1} + \boldsymbol{Z}'\widehat{\boldsymbol{R}}^{-1}\boldsymbol{Z})^{-1}], \tag{2.19}$$

for $0 \le r \le 1$. As shown in the next section, the constrained EBLUP for (2.19) satisfies the requirement (2.18) when $r \ge 1/2$.

# 3 Approximation and Estimation of MSE

## 3.1 Approximation of MSE

Constrained Bayes estimators or benchmarked empirical Bayes estimators have been suggested to modify the phenomena of shrinking the original data too much. However, this raises a question about the cost that the benchmarking makes variability of EBLUP higher. Thus, it is important to examine the mean squared error (MSE) of the benchmarked estimators. However, it is difficult to investigate the MSE analytically for the general constraint (C). In this section, we focus on the following constraint which includes typical constraints:

$$(\mathbf{C}_{(r)}) \begin{cases} \text{(C1)} & \boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}} = \boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}^{EB} + \boldsymbol{m}(\boldsymbol{y}), \\ \text{(C2)} & \widehat{\boldsymbol{\theta}}'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}} = (\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB} + h(\widehat{\boldsymbol{\psi}}), \end{cases} \tag{3.1}$$

where $\boldsymbol{m}(\boldsymbol{y}) = O_p(N^{1/2})$ and $h(\boldsymbol{\psi}) = O(N^{1-r})$ for $0 \leq r \leq 1$. Since $\boldsymbol{t}_1(\boldsymbol{y}) - \boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}^{EB} = \boldsymbol{m}(\boldsymbol{y})$, from (2.16), $\widehat{\mu}^{CEB}$ is expressed as

$$\widehat{\mu}^{CEB} = \widehat{\mu}^{EB} + \{a(\boldsymbol{y}) - 1\}\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\Omega}\widehat{\boldsymbol{\theta}}^{EB} + \boldsymbol{c}_w'\boldsymbol{m}(\boldsymbol{y}), \tag{3.2}$$

where $\boldsymbol{c}_w' = \boldsymbol{c}'\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W})^{-1}$ and

$$\{a(\boldsymbol{y})\}^2 = \frac{h(\widehat{\boldsymbol{\psi}})}{(\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\Omega}\widehat{\boldsymbol{\theta}}^{EB}} + 1. \tag{3.3}$$

Then, MSE of $\widehat{\mu}^{CEB}$ is written as $MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB}) = I_1 + I_2 + 2I_3$, where

$$\begin{aligned} I_1 =& E[(\widehat{\mu}^{EB} - \mu)^2], \\ I_2 =& E\big[\big\{\{a(\boldsymbol{y}) - 1\}\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\Omega}\widehat{\boldsymbol{\theta}}^{EB} + \boldsymbol{c}_w'\boldsymbol{m}(\boldsymbol{y})\big\}^2\big], \\ I_3 =& E\big[(\widehat{\mu}^{EB} - \widehat{\mu}^{B})\big\{\{a(\boldsymbol{y}) - 1\}\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\Omega}\widehat{\boldsymbol{\theta}}^{EB} + \boldsymbol{c}_w'\boldsymbol{m}(\boldsymbol{y})\big\}\big]. \end{aligned} \tag{3.4}$$

The asymptotic property of $MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB})$ depends on the value of $r$, so that we consider the three cases of $r$ below.

**[Scenario 1] Case of $r = 0$.** From Theorem A.1, it follows that the MSE of $\widehat{\mu}^{CEB}$ is approximated as

$$MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB}) = MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB}) + \{A(\boldsymbol{\omega}) - 1\}^2 B(\boldsymbol{\omega}) + O(N^{-1/2}), \tag{3.5}$$

where

$$\{A(\boldsymbol{\omega})\}^2 = \frac{h(\boldsymbol{\psi})}{\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{P}_{\Omega}\boldsymbol{X}\boldsymbol{\beta} + \mathrm{tr}\,[\boldsymbol{Z}'\boldsymbol{P}_{\Omega}\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z}]} + 1, \tag{3.6}$$

$$B(\boldsymbol{\omega}) = (\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\Omega}\boldsymbol{X}\boldsymbol{\beta})^2 + \boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\Omega}\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{P}_{\Omega}\boldsymbol{\Omega}^{-1}\boldsymbol{c}. \tag{3.7}$$

Concerning $MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB})$, it can be easily shown that $MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB}) = g_1(\boldsymbol{\psi}) + O(N^{-1/2})$, where

$$g_1(\boldsymbol{\psi}) = \boldsymbol{c}'\boldsymbol{Z}\boldsymbol{Z}(\boldsymbol{Q}^{-1} + \boldsymbol{Z}'\boldsymbol{R}^{-1}\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{Z}'\boldsymbol{c} = \boldsymbol{c}'\boldsymbol{Q}\boldsymbol{c} - \boldsymbol{s}'\boldsymbol{\Sigma}\boldsymbol{s}. \tag{3.8}$$

for $\boldsymbol{s} = \boldsymbol{s}(\boldsymbol{\psi}) = \boldsymbol{\Sigma}^{-1}\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{c}$. Hence, from (3.5),

$$\lim_{N\to\infty} MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB}) = \lim_{N\to\infty} g_1(\boldsymbol{\psi}) + \lim_{N\to\infty} \{A(\boldsymbol{\omega}) - 1\}^2 B(\boldsymbol{\omega}). \tag{3.9}$$

Since $\lim_{N\to\infty} MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB}) = \lim_{N\to\infty} g_1(\boldsymbol{\psi})$, it is seen that

$$\lim_{N\to\infty} MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB}) > \lim_{N\to\infty} MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB}),$$

that is, the constrained EBLUP $\widehat{\mu}^{CEB}$ has a larger MSE than the EBLUP $\widehat{\mu}^{EB}$ in the first order approximation. Since a good property of EBLUP is that the MSE of the EBLUP is small, the above fact gives a criticism for the constrained EBLUP.

Although we give the first-order approximation of the MSE in the case of $r = 0$, it may be possible to derive a second-order approximation. However, the second-order term probably consists of many terms, it may be hard to describe. Also, the difference between $\widehat{\mu}^{CEB}$ and $\widehat{\mu}^{EB}$ is recognized in the first-order as explained above. Thus, we do not give any further study about the second-order approximation in the case of $r = 0$.

**[Scenario 2] Case of $r = 1/2$.** In this case, a second order approximation for the MSE of $\widehat{\mu}^{CEB}$ is given by Theorem A.2, and it is given by

$$MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB}) = MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB}) + \overline{I}_2(\boldsymbol{\omega}) + 2\overline{I}_3(\boldsymbol{\omega}) + O(N^{-3/2}), \tag{3.10}$$

where

$$\begin{aligned}\overline{I}_2(\boldsymbol{\omega}) =& \frac{1}{4}\big[\{A(\boldsymbol{\omega})\}^2 - 1\big]^2 B(\boldsymbol{\omega}) + E\big[\{\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})\}^2\big] \\ &+ \big[\{A(\boldsymbol{\omega})\}^2 - 1\big]\big\{\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\Omega}\boldsymbol{X}\boldsymbol{\beta}E[\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})] + \boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\Omega}\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'E[\boldsymbol{\nabla}_{\boldsymbol{y}}\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})]\big\},\end{aligned} \tag{3.11}$$

$$\begin{aligned}\overline{I}_3(\boldsymbol{\omega}) =& \frac{1}{2}\big[\{A(\boldsymbol{\omega})\}^2 - 1\big](\boldsymbol{c} - \boldsymbol{s})'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{P}_{\Omega}\boldsymbol{\Omega}^{-1}\boldsymbol{c} \\ &+ (\boldsymbol{c} - \boldsymbol{s})'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'E[\boldsymbol{\nabla}_{\boldsymbol{y}}\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})] + \sum_{a=1}^{q} E[\widehat{\psi}_a^{\dagger}\boldsymbol{s}_{(a)}'\boldsymbol{\nabla}_{\boldsymbol{y}}\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})],\end{aligned} \tag{3.12}$$

for the differential operator $\boldsymbol{\nabla}_{\boldsymbol{y}} = \partial/\partial\boldsymbol{y} = (\partial/\partial y_1, \ldots, \partial/\partial y_N)'$.

Concerning $MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB})$, on the other hand, it follows from Prasad and Rao (1990), Datta and Lahiri (2000) and Kubokawa (2011) that

$$MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB}) = g_1(\boldsymbol{\psi}) + g_2(\boldsymbol{\psi}) + g_3(\boldsymbol{\psi}) + O(N^{-3/2}), \tag{3.13}$$

where $g_1(\boldsymbol{\psi})$ is given in (3.8), and

$$\begin{aligned}g_2(\boldsymbol{\psi}) =& (\boldsymbol{c} - \boldsymbol{s})'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{c} - \boldsymbol{s}), \\ g_3(\boldsymbol{\psi}) =& \mathrm{tr}\Big[\big(\frac{\partial\boldsymbol{s}'}{\partial\boldsymbol{\psi}}\big)\boldsymbol{\Sigma}\big(\frac{\partial\boldsymbol{s}'}{\partial\boldsymbol{\psi}}\big)'\mathbf{Cov}\,(\widehat{\boldsymbol{\psi}}^{\dagger})\Big],\end{aligned} \tag{3.14}$$

for $\boldsymbol{s} = \boldsymbol{\Sigma}^{-1}\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{c}$ and $\mathbf{Cov}\,(\widehat{\boldsymbol{\psi}}^{\dagger}) = E[(\widehat{\boldsymbol{\psi}}^{\dagger} - E[\widehat{\boldsymbol{\psi}}^{\dagger}])(\widehat{\boldsymbol{\psi}}^{\dagger} - E[\widehat{\boldsymbol{\psi}}^{\dagger}])']$, $\widehat{\boldsymbol{\psi}}^{\dagger}$ being given in (A4) in the appendix. Hence, from (3.10), the MSE of the constrained EBLUP $\widehat{\mu}^{CEB}$ is approximated as

$$MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB}) = g_1(\boldsymbol{\psi}) + g_2(\boldsymbol{\psi}) + g_3(\boldsymbol{\psi}) + \overline{I}_2(\boldsymbol{\omega}) + 2\overline{I}_3(\boldsymbol{\omega}) + O(N^{-3/2}), \tag{3.15}$$

It is noted that $g_1(\boldsymbol{\psi}) = O(1)$, $g_2(\boldsymbol{\psi}) = O(N^{-1})$, $g_3(\boldsymbol{\psi}) = O(N^{-1})$, $\overline{I}_2(\boldsymbol{\omega}) = O(N^{-1})$ and $\overline{I}_3(\boldsymbol{\omega}) = O(N^{-1})$. This implies that in the case of $r = 1/2$,

$$\lim_{N\to\infty} MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB}) = \lim_{N\to\infty} MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB}),$$

that is, the MSE of $\widehat{\mu}^{CEB}$ is equal to that of $\widehat{\mu}^{EB}$ in the first order, and their difference appears in the second order terms.

**[Scenario 3] Case of $r = 1$.** From Theorem A.3, it follows that

$$MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB}) = MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB}) + E[\{\boldsymbol{c}_{\boldsymbol{w}}' \boldsymbol{m}(\boldsymbol{y})\}^2] + 2\overline{I}_{32}(\boldsymbol{\omega}) + O(N^{-3/2}), \quad (3.16)$$

where

$$\overline{I}_{32}(\boldsymbol{\omega}) = (\boldsymbol{c} - \boldsymbol{s})' \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}' E[\boldsymbol{\nabla}_{\boldsymbol{y}} \boldsymbol{c}_{\boldsymbol{w}}' \boldsymbol{m}(\boldsymbol{y})] + \sum_{a=1}^{q} E[\widehat{\psi}_a^{\dagger} \boldsymbol{s}_{(a)}' \boldsymbol{\nabla}_{\boldsymbol{y}} \boldsymbol{c}_{\boldsymbol{w}}' \boldsymbol{m}(\boldsymbol{y})]. \quad (3.17)$$

In the case that $\boldsymbol{m}(\boldsymbol{y}) = 0$, it is simplified as

$$MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB}) = MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB}) + O(N^{-3/2}),$$

that is, the constrained EBLUP has the same MSE to EBLUP up to $O(N^{-1})$.

**Example 3.1 (variance constraint)** Consider the following variance constraint which benchmarks the variance of estimates:

$$(\mathbf{V}_{(r)}) \begin{cases} \text{(V1)} \quad \boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}} = \boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}^{EB}, \\ \text{(V2)} \quad \widehat{\boldsymbol{\theta}}' \boldsymbol{P}_{\boldsymbol{\Omega}} \widehat{\boldsymbol{\theta}} = t_2^{(r)}(\boldsymbol{y}) \text{ for the function } t_2^{(r)}(\boldsymbol{y}) \text{ given in (2.19).} \end{cases} \quad (3.18)$$

This corresponds to the case that $\boldsymbol{m}(\boldsymbol{y}) = 0$ and

$$h(\boldsymbol{\psi}) = \frac{1}{N^r} \mathrm{tr}\, [\boldsymbol{P}_{\boldsymbol{\Omega}}(\boldsymbol{Q}^{-1} + \boldsymbol{Z}'\boldsymbol{R}^{-1}\boldsymbol{Z})^{-1}]. \quad (3.19)$$

Then, the constrained EBLUP is

$$\widehat{\mu}^{CEB} = \widehat{\mu}^{EB} + \{a^{(r)}(\boldsymbol{y}) - 1\} \boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB}, \quad (3.20)$$

where

$$\{a^{(r)}(\boldsymbol{y})\}^2 = 1 + \frac{1}{N^r} \frac{\mathrm{tr}\, [\boldsymbol{P}_{\boldsymbol{\Omega}}(\widehat{\boldsymbol{Q}}^{-1} + \boldsymbol{Z}'\widehat{\boldsymbol{R}}^{-1}\boldsymbol{Z})^{-1}]}{(\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB}}. \quad (3.21)$$

The corresponding approximations of MSE of the constrained EBLUP can be provided from the above results. ■

**Example 3.2 (mean constraint)** Consider the following mean constraint which benchmarks the mean of estimates:

$$(\mathbf{M}) \begin{cases} \text{(M1)} \quad \boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}} = \boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{y}, \\ \text{(M2)} \quad \widehat{\boldsymbol{\theta}}' \boldsymbol{P}_{\boldsymbol{\Omega}} \widehat{\boldsymbol{\theta}} = (\widehat{\boldsymbol{\theta}}^{EB})' \boldsymbol{P}_{\boldsymbol{\Omega}} \widehat{\boldsymbol{\theta}}^{EB}. \end{cases} \quad (3.22)$$

In this case, $h(\boldsymbol{\psi}) = 0$ or $a(\boldsymbol{y}) = 1$, and $\boldsymbol{m}(\boldsymbol{y}) = \boldsymbol{W}'\Omega(\boldsymbol{y} - \widehat{\boldsymbol{\theta}}^{EB}) = \boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{R}}\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\psi}}))$ in (3.1). From Proposition A.1, it follows that

$$\begin{aligned} MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB}) =& MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB}) + \boldsymbol{c}_{\boldsymbol{w}}' \boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{R}\boldsymbol{\Sigma}^{-1}\boldsymbol{R}\boldsymbol{\Omega}\boldsymbol{W}\boldsymbol{c}_{\boldsymbol{w}} \\ & + 2(\boldsymbol{c} - \boldsymbol{s})'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{R}\boldsymbol{\Omega}\boldsymbol{W}\boldsymbol{c}_{\boldsymbol{w}} + O(N^{-3/2}). \end{aligned} \quad (3.23)$$

It is easily seen that MSE of $\widehat{\mu}^{CEB}$ is equal to MSE of $\widehat{\mu}^{EB}$ in the first order. ■

## 3.2 Approximated unbiased estimator of MSE

We here provide an asymptotically unbiased estimator of $MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB})$ with second-order accuracy for $\widehat{\mu}^{CEB}$ given in (3.2).

The parametric bootstrap method is useful for estimating second-order unbiased estimators. Based on Butar and Lahiri (2003), Hall and Maiti (2006) and Kubokawa and Nagashima (2012) and others, we consider the following conditional linear mixed model given $\boldsymbol{y}$:

$$\boldsymbol{y}^* = \boldsymbol{X}\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\psi}}) + \boldsymbol{Z}\boldsymbol{v}^* + \boldsymbol{\epsilon}^*, \tag{3.24}$$

where $\boldsymbol{X}$ and $\boldsymbol{Z}$ are the same matrices as given in (2.1), and given $\boldsymbol{y}$, $\boldsymbol{v}^*$ and $\boldsymbol{\epsilon}^*$ are conditionally mutually independently distributed as $\boldsymbol{v}^*|\boldsymbol{y} \sim \mathcal{N}_M(0, \boldsymbol{Q}(\widehat{\boldsymbol{\psi}}))$ and $\boldsymbol{\epsilon}^*|\boldsymbol{y} \sim \mathcal{N}_N(0, \boldsymbol{R}(\widehat{\boldsymbol{\psi}}))$. Then, Kubokawa and Nagashima (2012) suggested the second-order unbiased estimator of $MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB})$, given by

$$mse^*_{EB} = 2\{g_1(\widehat{\boldsymbol{\psi}}) + g_2(\widehat{\boldsymbol{\psi}})\} - E_*[g_1(\widehat{\boldsymbol{\psi}}^*) + g_2(\widehat{\boldsymbol{\psi}}^*)|\boldsymbol{y}] + g^*_3(\widehat{\boldsymbol{\psi}}), \tag{3.25}$$

where

$$g^*_3(\widehat{\boldsymbol{\psi}}) = E_*[\{\boldsymbol{s}(\widehat{\boldsymbol{\psi}}^*) - \boldsymbol{s}(\widehat{\boldsymbol{\psi}})\}'\widehat{\boldsymbol{\Sigma}}(\widehat{\boldsymbol{\psi}})\{\boldsymbol{s}(\widehat{\boldsymbol{\psi}}^*) - \boldsymbol{s}(\widehat{\boldsymbol{\psi}})\}|\boldsymbol{y}].$$

**[1] General case of $r \geq 0$.** We can get a second-order unbiased estimator by estimating $I_2$ and $I_3$ given in (3.4) directly. Based on the model (3.24), define $I^*_3$ by

$$I^*_3 = E_*\big[(\widehat{\mu}^{EB*} - \widehat{\mu}^{B*})\big\{\{a(\boldsymbol{y}^*) - 1\}\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_\Omega\widehat{\boldsymbol{\theta}}^{EB*} + \boldsymbol{c}'_w\boldsymbol{m}(\boldsymbol{y}^*)\big\}|\boldsymbol{y}\big], \tag{3.26}$$

where for $\widehat{\boldsymbol{\beta}}^*(\widehat{\boldsymbol{\psi}}^*) = (\boldsymbol{X}'\widehat{\boldsymbol{\Sigma}}(\widehat{\boldsymbol{\psi}}^*)^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\widehat{\boldsymbol{\Sigma}}(\widehat{\boldsymbol{\psi}}^*)^{-1}\boldsymbol{y}^*$,

$$\begin{aligned}\widehat{\mu}^{EB*} =& \boldsymbol{c}'\boldsymbol{X}\widehat{\boldsymbol{\beta}}^*(\widehat{\boldsymbol{\psi}}^*) + \boldsymbol{Z}\boldsymbol{Q}(\widehat{\psi}^*)\boldsymbol{Z}'\boldsymbol{\Sigma}(\widehat{\boldsymbol{\psi}}^*)^{-1}(\boldsymbol{y}^* - \boldsymbol{X}\widehat{\boldsymbol{\beta}}^*(\widehat{\boldsymbol{\psi}}^*)),\\ \widehat{\mu}^{B*} =& \boldsymbol{c}'\boldsymbol{X}\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\psi}}) + \boldsymbol{Z}\boldsymbol{Q}(\widehat{\psi})\boldsymbol{Z}'\boldsymbol{\Sigma}(\widehat{\boldsymbol{\psi}})^{-1}(\boldsymbol{y}^* - \boldsymbol{X}\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\psi}})),\end{aligned} \tag{3.27}$$

Then, from Theorem A.4, a second-order unbiased estimator of $MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB})$ is given by

$$mse^*_{CEB} = mse^*_{EB} + \widehat{I}_2 + 2I^*_3, \tag{3.28}$$

where

$$\widehat{I}_2 = \big\{\{a(\boldsymbol{y}) - 1\}\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_\Omega\widehat{\boldsymbol{\theta}}^{EB} + \boldsymbol{c}'_w\boldsymbol{m}(\boldsymbol{y})\big\}^2. \tag{3.29}$$

**[2] Case of $r = 1/2$.** In this case, we can consider two other estimators using the second-order approximation of MSE given in the previous section. One is a procedure based on the Taylor series expansion. For the MSE $MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB})$ of the EBLUP $\widehat{\mu}^{EB}$, Prasad and Rao (1990), Datta and Lahiri (2000) and Kubokawa (2010) derived a second-order unbiased estimator, denoted by $mse_{EB}$, based on Taylor series approximation. Together with Theorem A.2, we get the second order unbiased estimator

$$mse_{CEB} = mse_{EB} + \overline{I}_2(\widehat{\boldsymbol{\psi}}) + 2\overline{I}_3(\widehat{\boldsymbol{\psi}}), \tag{3.30}$$

which satisfies that $E[mse_{CEB}] = MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB}) + O(N^{-3/2})$.

Another second-order unbiased estimator for $r = 1/2$ can be derived from Theorem A.5, and it is given by

$$mse_{CEB}^{**} = mse_{EB}^{*} + \overline{I}_2^{*} + 2\overline{I}_3^{*}, \tag{3.31}$$

where for $\widehat{\boldsymbol{\omega}} = (\widehat{\boldsymbol{\psi}}, \widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\psi}}))$ and $\widehat{\boldsymbol{\beta}} = \widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\psi}})$,

$$\begin{aligned}\overline{I}_2^{*} =& \frac{1}{4}\big[\{A(\widehat{\boldsymbol{\omega}})\}^2 - 1\big]^2 B(\widehat{\boldsymbol{\omega}}) + \big\{\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})\big\}^2 \\ &+ \big[\{A(\widehat{\boldsymbol{\omega}})\}^2 - 1\big]\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\big\{\boldsymbol{X}\widehat{\boldsymbol{\beta}}\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y}) + \boldsymbol{Z}\widehat{\boldsymbol{Q}}\boldsymbol{Z}'\widehat{\boldsymbol{\Sigma}}^{-1}E_*\big[(\boldsymbol{y}^* - \boldsymbol{X}\widehat{\boldsymbol{\beta}})\}\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y}^*)|\boldsymbol{y}\big]\big\},\end{aligned} \tag{3.32}$$

$$\overline{I}_3^{*} = \frac{1}{2}\big[\{A(\widehat{\boldsymbol{\omega}})\}^2 - 1\big]E_*\Big[(\widehat{\mu}^{EB*} - \widehat{\mu}^{B*})\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{B*}|\boldsymbol{y}\Big] + E_*\Big[(\widehat{\mu}^{EB*} - \widehat{\mu}^{B*})\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y}^*)|\boldsymbol{y}\Big], \tag{3.33}$$

**[3] Case of $r = 1$.** The estimator given in (3.31) is applicable to the case of $r = 1$, and the second order unbiased estimator can be provided by putting $A(\widehat{\boldsymbol{\omega}}) = 1$ in $mse_{CEB}^{**}$. It can be seen that the resulting estimator is identical to $mse_{CEB}^{*}$ given by

$$mse_{CEB}^{*} = mse_{EB}^{*} + \big\{\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})\big\}^2 + 2E_*\Big[(\widehat{\mu}^{EB*} - \widehat{\mu}^{B*})\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y}^*)|\boldsymbol{y}\Big], \tag{3.34}$$

which is identical to $mse(\widehat{\mu}^{CEB})$ given in (3.28) for $a(\boldsymbol{y}) = 1$.

# 4 Simulation and Empirical Studies in the Fay-Herriot model

In this section, we apply the proposed estimator to the Fay-Herriot model, and investigate the performances by simulation and an empirical example.

## 4.1 Constrained EBLUP and MSE estimation in the Fay-Herriot model

The basic area level model proposed by Fay and Herriot (1979) is described by

$$y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + v_i + \varepsilon_i, \quad i = 1, \ldots, k, \tag{4.1}$$

where $k$ is the number of small areas, $\boldsymbol{x}_i$ is a $p \times 1$ vector of explanatory variables, $\boldsymbol{\beta}$ is a $p \times 1$ unknown common vector of regression coefficients, and $v_i$'s and $\varepsilon_i$'s are mutually independently distributed random errors such that $v_i \sim \mathcal{N}(0, \psi)$ and $\varepsilon_i \sim \mathcal{N}(0, d_i)$ for known $d_i$'s. Let $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k)'$, $\boldsymbol{y} = (y_1, \ldots, y_k)'$, and let $\boldsymbol{v}$ and $\boldsymbol{\epsilon}$ be similarly defined. Then, the model is expressed in vector notations as $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{v} + \boldsymbol{\epsilon}$ and $\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{X}\boldsymbol{\beta}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\psi) = \psi\boldsymbol{I}_k + \boldsymbol{D}$ for $\boldsymbol{D} = \mathrm{diag}\,(d_1, \ldots, d_k)$. In this model, $\boldsymbol{Z} = \boldsymbol{I}_k$, $\boldsymbol{R} = \boldsymbol{D}$ and $\boldsymbol{Q} = \psi\boldsymbol{I}_k$.

As estimators of $\psi$, we here use the truncated Prasad-Rao estimator and the truncated Fay-Herriot estimator. The truncated Prasad-Rao estimator is $\widehat{\psi}^{PR} = \{(k-p)^{-1}(\boldsymbol{y}'\boldsymbol{E}_0\boldsymbol{y} -$

$\mathrm{tr}\,[\boldsymbol{D}\boldsymbol{E}_0]), k^{-1/2}\}$ for $\boldsymbol{E}_0 = \boldsymbol{I}_k - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$. Then, $Bias(\widehat{\psi}^{PR}) = E[\widehat{\psi}^{PR} - \psi] = O(k^{-3/2})$ and $Var(\widehat{\psi}^{PR}) = E[(\widehat{\psi}^{PR} - \psi)^2] = 2k^{-2}\mathrm{tr}\,\boldsymbol{\Sigma}^2 + O(k^{-3/2})$. The truncated Fay-Herriot estimator is $\widehat{\psi}^{FH} = \max\{\psi_0, k^{-1/2}\}$, where $\psi_0$ is the solution of the equation $L^{FH}(\psi_0) = 0$, where

$$L^{FH}(\psi_0) = \boldsymbol{y}'\left\{\boldsymbol{\Sigma}(\psi_0)^{-1} - \boldsymbol{\Sigma}(\psi_0)^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}(\psi_0)^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}(\psi_0)^{-1}\right\}\boldsymbol{y} - (k-p).$$

Then, $Bias(\widehat{\psi}^{FH}) = 2\{k\mathrm{tr}\,[\boldsymbol{\Sigma}^{-2}] - (\mathrm{tr}\,[\boldsymbol{\Sigma}^{-1}])^2\}/(\mathrm{tr}\,[\boldsymbol{\Sigma}^{-1}])^3 + O(k^{-3/2})$ and $Var(\widehat{\psi}^{FH}) = 2k/(\mathrm{tr}\,[\boldsymbol{\Sigma}^{-1}])^2 + O(k^{-3/2})$.

As the constrained EBLUP, we consider the case that $L = 1$, $\boldsymbol{W} = \boldsymbol{j} = (1, \ldots, 1)' \in \boldsymbol{R}^k$, $\boldsymbol{\Omega} = \boldsymbol{D}^{-1}$ and $\boldsymbol{c} = \boldsymbol{e}_i$, where the $i$-th element of $\boldsymbol{e}_i$ is one and the other elements are zeros. Note that $\boldsymbol{P}_{\boldsymbol{\Omega}} = \boldsymbol{D}^{-1} - (\boldsymbol{j}'\boldsymbol{d}_*)^{-1}\boldsymbol{d}_*\boldsymbol{d}_*'$ and $\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}} = \boldsymbol{I}_k - (\boldsymbol{j}'\boldsymbol{d}_*)^{-1}\boldsymbol{j}\boldsymbol{d}_*'$ for $\boldsymbol{d}_* = \boldsymbol{\Omega}\boldsymbol{j} = (d_1^{-1}, \ldots, d_k^{-1})'$. Then from (3.2), the constrained EBLUP under the constraint (3.1) is written as

$$\hat{\theta}_i^{CEB} = \hat{\theta}_i^{EB} + \{a(\boldsymbol{y}) - 1\}\left\{\hat{\theta}_i^{EB} - (\boldsymbol{j}'\boldsymbol{d}_*)^{-1}\boldsymbol{d}_*'\widehat{\boldsymbol{\theta}}^{EB}\right\} + (\boldsymbol{j}'\boldsymbol{d}_*)^{-1}m(\boldsymbol{y}), \tag{4.2}$$

where $\widehat{\boldsymbol{\theta}}^{EB} = \boldsymbol{X}\widehat{\boldsymbol{\beta}}(\widehat{\psi}) + \widehat{\psi}\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}(\widehat{\psi}))$ and $\boldsymbol{a}(\boldsymbol{y})$ is given in (3.3).

As derived by Prasad and Rao (1990) and Datta, Rao and Smith (2005), EBLUP $\hat{\theta}_i^{EB}$ is $\hat{\theta}_i^{EB} = \boldsymbol{x}_i'\widehat{\boldsymbol{\beta}}(\widehat{\psi}) + \{1 - \gamma_i(\widehat{\psi})\}(y_i - \boldsymbol{x}_i'\widehat{\boldsymbol{\beta}}(\widehat{\psi}))$ for $\gamma_i(\psi) = d_i/(\psi + d_i)$, and the second-order approximation of the MSE is $MSE(\psi, \hat{\theta}_i^{EB}) = g_1(\psi) + g_2(\psi) + g_3(\psi) + O(k^{-3/2})$ where $g_1(\psi) = d_i\{1 - \gamma_i(\psi)\}$, $g_2(\psi) = \gamma_i(\psi)^2\boldsymbol{x}_i'(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{x}_i$ and $g_3(\psi) = d_i^{-1}\gamma_i(\psi)^3 Var(\widehat{\psi})$ for $\boldsymbol{s}(\psi) = \{1 - \gamma_i(\psi)\}\boldsymbol{e}_i$. Also, the second-order unbiased estimators of $MSE(\psi, \hat{\theta}_i^{EB})$ are given by

$$mse_{EB} = g_1(\widehat{\psi}) + g_2(\widehat{\psi}) + 2g_3(\widehat{\psi}) - g_{11}(\widehat{\psi}), \tag{4.3}$$

$$mse_{EB}^* = 2\{g_1(\widehat{\psi}) + g_2(\widehat{\psi})\} - E_*\left[g_1(\widehat{\psi}^*) + g_2(\widehat{\psi}^*)\middle|\boldsymbol{y}\right] + g_3{}^*(\widehat{\psi}), \tag{4.4}$$

where $g_{11}(\psi) = \gamma_i(\psi)^2 Bias(\widehat{\psi})$ and $g_3{}^*(\widehat{\psi}) = E_*\left[\{\gamma_i(\widehat{\psi}^*) - \gamma_i(\widehat{\psi})\}^2(\widehat{\psi} + d_i)|\boldsymbol{y}\right]$. It is noted that $mse_{EB}$ is based on the Taylor series approximation and $mse_{EB}^*$ is the parametric bootstrap procedure given in (3.25).

In this section, we treat the variance constraint (3.18) and the mean constraint (3.22) explained in Examples 3.1 and 3.2. For the variance constraint (3.18), $m(\boldsymbol{y})$ and $h(\psi)$ are given by $m(\boldsymbol{y}) = 0$ and $h(\psi) = N^{-r}\mathrm{tr}\,[\boldsymbol{P}_{\boldsymbol{\Omega}}\psi\boldsymbol{D}(\psi\boldsymbol{I}_k + \boldsymbol{D})^{-1}]$. For simplicity, the constrained EBLUP $\hat{\theta}_i^{CEB}$ under the variance constraint for $r = 0, 0.5, 1$ are denoted by $cEB_{V0}$, $cEB_{V0.5}$ and $cEB_{V1}$, respectively. For the mean constraint (3.22), $m(\boldsymbol{y})$ and $h(\psi)$ are given by $m(\boldsymbol{y}) = \boldsymbol{j}'\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}(\widehat{\psi}))$ and $h(\psi) = 0$, and the constrained EBLUP $\hat{\theta}_i^{CEB}$ under the mean constraint is denoted by $cEB_M$.

For the constrained EBLUP $cEB_{V0}$, $cEB_{V0.5}$, $cEB_{V1}$ and $cEB_M$, the second-order unbiased estimator of MSE is given in (3.28), which is expressed in this model as

$$m_{CEB}^* = mse_{EB}^* + \widehat{I}_2 + 2I_3^*, \tag{4.5}$$

where $\widehat{I}_2 = \big\{\{a^{(r)}(\boldsymbol{y})-1\}\{\hat{\theta}_i^{EB} - (\boldsymbol{j}'\boldsymbol{d}_*)^{-1}\boldsymbol{d}_*'\widehat{\boldsymbol{\theta}}^{EB}\} + (\boldsymbol{j}'\boldsymbol{d}_*)^{-1}m(\boldsymbol{y})\big\}^2$ and

$$I_3^* = E_*\big[(\hat{\theta}_i^{EB*} - \hat{\theta}_i^{B*})\big\{\{a^{(r)}(\boldsymbol{y}^*)-1\}\{\hat{\theta}_i^{EB*} - (\boldsymbol{j}'\boldsymbol{d}_*)^{-1}\boldsymbol{d}_*'\widehat{\boldsymbol{\theta}}^{EB}\} + (\boldsymbol{j}'\boldsymbol{d}_*)^{-1}m^{EB}(\boldsymbol{y}^*)\big\}\big],$$

for $\hat{\theta}_i^{EB*} = \boldsymbol{x}_i'\widehat{\boldsymbol{\beta}}^*(\widehat{\psi}^*)+\{1-\gamma_i(\widehat{\psi}^*)\}(y_i^*-\boldsymbol{x}_i'\widehat{\boldsymbol{\beta}}^*(\widehat{\psi}^*))$, $\hat{\theta}_i^{B*} = \boldsymbol{x}_i'\widehat{\boldsymbol{\beta}}(\widehat{\psi})+\{1-\gamma_i(\widehat{\psi})\}(y_i^*-\boldsymbol{x}_i'\widehat{\boldsymbol{\beta}}(\widehat{\psi}))$ and $a^{(r)}(\boldsymbol{y})$ given in (3.21).

For the estimator $cEB_{V0.5}$, we can provide two other second order unbiased estimators given in (3.30) and (3.31). Thus, we have three estimators $m_{CEB}^*$,

$$m_{V0.5} = mse_{EB} + \overline{I}_2(\widehat{\boldsymbol{\omega}}) + 2\overline{I}_3(\widehat{\boldsymbol{\omega}}), \quad (4.6)$$
$$m_{V05}^{**} = mse_{EB}^* + \overline{I}_2(\widehat{\boldsymbol{\omega}}) + 2\overline{I}_3^*, \quad (4.7)$$

where

$$\begin{aligned}\overline{I}_2(\boldsymbol{\omega}) =& 4^{-1}\big[\{A(\boldsymbol{\omega})\}^2-1\big]^2 B(\boldsymbol{\omega}),\\ \overline{I}_3(\boldsymbol{\omega}) =& 2^{-1}\big[\{A(\boldsymbol{\omega})\}^2-1\big]\gamma_i(\psi)\boldsymbol{x}_i'(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\psi(\boldsymbol{I}_k - (\boldsymbol{j}'\boldsymbol{d}_*)^{-1}\boldsymbol{j}\boldsymbol{d}_*')\boldsymbol{e}_i,\\ \overline{I}_3^* =& 2^{-1}\big[\{A(\widehat{\psi})\}^2-1\big]E_*\big[(\hat{\theta}_i^{EB*} - \hat{\theta}_i^{B*})\{\hat{\theta}_i^{B*} - (\boldsymbol{j}'\boldsymbol{d}_*)^{-1}\boldsymbol{d}_*'\widehat{\boldsymbol{\theta}}^{B*}\}\big].\end{aligned}$$

Here, $A(\boldsymbol{\omega})$ and $B(\boldsymbol{\omega})$ are given by

$$\begin{aligned}A(\boldsymbol{\omega})^2 =& \frac{h(\psi)}{\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{P}_{\boldsymbol{\Omega}}\boldsymbol{X}\boldsymbol{\beta} + \psi\mathrm{tr}\,[\boldsymbol{P}_{\boldsymbol{\Omega}}\boldsymbol{\Sigma}^{-1}]} + 1,\\ B(\boldsymbol{\omega}) =& \{\boldsymbol{x}_i'\boldsymbol{\beta} - (\boldsymbol{j}'\boldsymbol{d}_*)^{-1}\boldsymbol{d}_*'\boldsymbol{X}\boldsymbol{\beta}\}^2 + \psi^2\{\boldsymbol{e}_i - (\boldsymbol{j}'\boldsymbol{d}_*)^{-1}\boldsymbol{d}_*\}'\boldsymbol{\Sigma}^{-1}\{\boldsymbol{e}_i - (\boldsymbol{j}'\boldsymbol{d}_*)^{-1}\boldsymbol{d}_*\}.\end{aligned}$$

For the estimator $cEB_{V1}$, we have three types of second-order unbiased estimators, given by $m_{CEB}^*$, $m_{V1} = mse_{EB}$ and $m_{V1}^* = mse_{EB}^*$.

For second-order unbiased estimator of MSE of the estimator $cEB_M$, from (3.34) and (3.23), we have

$$m_{CEB}^* = mse_{EB}^* + \big\{(\boldsymbol{j}'\boldsymbol{d}_*)^{-1}m(\boldsymbol{y})\big\}^2 + 2E_*\Big[(\hat{\theta}_i^{EB*} - \hat{\theta}_i^{B*})(\boldsymbol{j}'\boldsymbol{d}_*)^{-1}m(\boldsymbol{y}^*)\Big], \quad (4.8)$$
$$m_M = mse_{EB} + (\boldsymbol{j}'\boldsymbol{d}_*)^{-2}\boldsymbol{j}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{j} + 2(\boldsymbol{j}'\boldsymbol{d}_*)^{-1}\gamma_i(\widehat{\psi})\boldsymbol{x}_i'(\boldsymbol{X}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{j}, \quad (4.9)$$

for $m(\boldsymbol{y}^*) = \boldsymbol{j}'\widehat{\boldsymbol{\Sigma}}(\widehat{\psi}^*)^{-1}(\boldsymbol{y}^* - \boldsymbol{X}\widehat{\boldsymbol{\beta}}^*(\widehat{\psi}^*))$.

## 4.2 Simulation results

We now investigate the performances of MSE of the constrained EBLUP and the performances of the MSE estimators. For the purpose, we adopt a part of the simulation framework of Datta, *et al*. (2005) for our study. We consider the Fay-Herriot model (4.1) with $k = 15$, $\psi = 1$ and two $d_i$-patterns: (a) 0.7, 0.6, 0.5, 0.4, 0.3; (b) 4.0, 0.6, 0.5, 0.4, 0.1, which correspond to patterns (a) and (c) of Datta, *et al*. (2005). Pattern (a) is less variable in $d_i$-values, while pattern (b) has larger variability. There are five groups $G_1, \ldots, G_5$ and three small areas in each group. The sampling variances $d_i$ are the same for area within the same group. Let us consider the case that $\boldsymbol{X}'\boldsymbol{\beta} = 0$ for simplicity as handled in Chatterjee, *et al.* (2008). Then, $\theta_i = v_i$, $\hat{\theta}_i^{EB} = \{1 - \gamma_i(\widehat{\psi})\}y_i$ for $\gamma_i(\psi) = d_i/(\psi + d_i)$.

We first investigate the values of MSE of the EBLUP and the constrained EBLUP by simulation based on 100,000 replications. We compare the EBLUP $\hat{\theta}_i^{EB}$, the variance constrained EBLUPs $cEB_{V0}$, $cEB_{V0.5}$ and $cEB_{V1}$ for $r = 0$, 0.5 and 1, and the mean constrained EBLUP $cEB_M$, whose MSEs are denoted by $M_{EB}$, $M_{V0}$, $M_{V0.5}$, $M_{V1}$ and $M_M$. Those numerical values are reported in Table 1. From this table, it is seen that $M_{V0}$ is slightly larger than $M_{EB}$, while MSEs of the other constrained EBLUPs are close to $M_{EB}$, the MSE of EBLUP. The relative fluctuations of MSE $M$ over $M_{EB}$, defined by

$$100 \times (M - M_{EB})/M_{EB},$$

are also given in parentheses in the table. The relative fluctuations of $M_{V0}$ range from 1.4% $\sim$ 6.3% and they are about 5% for the Prasad-Rao estimator in the pattern (a). Although the variance constrained EBLUP $cEB_{V0.5}$ is made so as to increase the sampling variance, the resulting MSEs are close to $M_{EB}$ and are smaller than $M_{EB}$ in some cases. The relative fluctuations of $M_M$ range from 0.4% $\sim$ 3.4%.

We next investigate the performances of the proposed estimators of the MSE. We treat the estimator $m^*_{CEB}$ for the MSE $M_{V0}$, the estimators $m^*_{CEB}$, $m^*_{V0.5}$ and $m_{V0.5}$ for $M_{V0.5}$, the estimators $m^*_{CEB}$, $m^*_{V1}$ and $m_{V1}$ for $M_{V1}$, and the estimators $m^*_{CEB}$ and $m_M$ for $M_M$. The relative bias and the risk functions of MSE estimator $mse_i$ for MSE $MSE(\boldsymbol{\omega}, \hat{\theta}_i)$ are given by

$$\begin{aligned}
B_i(\boldsymbol{\omega}, mse_i) =& 100 \times E\Big[mse_i - MSE(\boldsymbol{\omega}, \hat{\theta}_i)\Big]/MSE(\boldsymbol{\omega}, \hat{\theta}_i), \\
R_i(\boldsymbol{\omega}, mse_i) =& 100 \times E\Big[\big\{mse_i - MSE(\boldsymbol{\omega}, \hat{\theta}_i)\big\}^2\Big]/\{MSE(\boldsymbol{\omega}, \hat{\theta}_i)\}^2.
\end{aligned}$$

These values are computed as average values based on 10,000 simulation runs where the size of the bootstrap sample is 1,000. Further, those values are averaged over areas within groups $G_i$, $i = 1, \ldots, 5$, and they are reported in Tables 2 and 3, respectively, for the Prasad-Rao and the Fay-Herriot estimators. Through these tables, it is seen that the estimators $m^*_{CEB}$, $m^*_{V0.5}$ and $m^*_{V1}$ have smaller biases but larger risks in the pattern (a) than the estimators $m_{V0.5}$, $m_{V1}$ and $m_M$ based on the Taylor series expansion. For the pattern (b), the MSE estimators with the Prasad-Rao estimator are not good in terms of biases and risks, but the MSE estimators with the Fay-Herriot estimator give appropriate biases and risks except $G_1$. The MSE estimator $m^*_{CEB}$ with the Fay-Herriot estimator is recommendable as a simple and useful procedure.

## 4.3 An example

We apply the benchmarked estimates and the estimates of the MSE to the data in the Survey of Family Income and Expenditure (SFIE) in Japan.

In this study, we use the data of the disbursement item 'Education' in the survey in November, 2011. The average disbursement (scaled by 1,000 Yen) at each capital city of 47 prefectures in Japan is denoted by $y_i$ for $i = 1, \ldots, 47$, and each variance $d_i$ is calculated

Table 1: Values of MSE $M_{EB}$ of EBLUP and MSEs $M_{V0}$, $M_{V0.5}$, $M_{V1}$ and $M_M$ of the constrained EBLUPs in patterns (a) and (b) for the Prasad-Rao estimator $\widehat{\psi}^{PR}$ and Fay-Herriot estimator $\widehat{\psi}^{FH}$ and $\psi = 1$ where the values in the parentheses denote the relative fluctuations over $M_{EB}$ in percentage

| | **pattern (a)** | | | | | | **pattern (b)** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $d_i$ | $M_{EB}$ | $M_{V0}$ | $M_{V0.5}$ | $M_{V1}$ | $M_M$ | $d_i$ | $M_{EB}$ | $M_{V0}$ | $M_{V0.5}$ | $M_{V1}$ | $M_M$ |
| | **Prasad-Rao estimator $\widehat{\psi}^{PR}$** | | | | | | | | | | | |
| $G_1$ | 0.7 | 0.438 | 0.460 | 0.438 | 0.437 | 0.447 | 4.0 | 0.909 | 0.963 | 0.921 | 0.912 | 0.917 |
| | | | (5.1) | (0.0) | (−0.1) | (2.1) | | | (6.0) | (1.3) | (0.3) | (1.0) |
| $G_2$ | 0.6 | 0.398 | 0.418 | 0.397 | 0.397 | 0.407 | 0.6 | 0.425 | 0.435 | 0.424 | 0.424 | 0.427 |
| | | | (5.0) | (−0.2) | (−0.2) | (2.2) | | | (2.4) | (−0.1) | (−0.1) | (0.6) |
| $G_3$ | 0.5 | 0.354 | 0.371 | 0.352 | 0.353 | 0.362 | 0.5 | 0.378 | 0.386 | 0.377 | 0.378 | 0.380 |
| | | | (5.0) | (−0.4) | (−0.2) | (2.3) | | | (2.0) | (−0.3) | (−0.2) | (0.5) |
| $G_4$ | 0.4 | 0.303 | 0.318 | 0.301 | 0.302 | 0.310 | 0.4 | 0.325 | 0.329 | 0.323 | 0.324 | 0.326 |
| | | | (5.1) | (−0.6) | (−0.3) | (2.6) | | | (1.5) | (−0.6) | (−0.2) | (0.4) |
| $G_5$ | 0.3 | 0.244 | 0.257 | 0.242 | 0.243 | 0.251 | 0.1 | 0.100 | 0.101 | 0.098 | 0.099 | 0.101 |
| | | | (5.5) | (−0.8) | (−0.4) | (3.1) | | | (1.4) | (−1.4) | (−0.5) | (1.9) |
| | **Fay-Herriot estimator $\widehat{\psi}^{FH}$** | | | | | | | | | | | |
| $G_1$ | 0.7 | 0.438 | 0.457 | 0.437 | 0.437 | 0.446 | 4.0 | 0.853 | 0.875 | 0.857 | 0.854 | 0.858 |
| | | | (4.4) | (−0.2) | (−0.2) | (2.0) | | | (2.5) | (0.5) | (0.1) | (0.6) |
| $G_2$ | 0.6 | 0.398 | 0.417 | 0.397 | 0.397 | 0.406 | 0.6 | 0.401 | 0.413 | 0.401 | 0.400 | 0.404 |
| | | | (4.6) | (−0.3) | (−0.2) | (2.1) | | | (3.0) | (0.0) | −0.1) | (0.8) |
| $G_3$ | 0.5 | 0.353 | 0.371 | 0.352 | 0.352 | 0.361 | 0.5 | 0.355 | 0.366 | 0.355 | 0.355 | 0.358 |
| | | | (5.0) | (−0.4) | (−0.2) | (2.3) | | | (3.1) | (−0.1) | (−0.1) | (0.8) |
| $G_4$ | 0.4 | 0.302 | 0.318 | 0.300 | 0.301 | 0.310 | 0.4 | 0.303 | 0.313 | 0.303 | 0.303 | 0.306 |
| | | | (5.5) | (−0.5) | (−0.3) | (2.6) | | | (3.3) | (−0.2) | (−0.1) | (0.9) |
| $G_5$ | 0.3 | 0.242 | 0.258 | 0.241 | 0.242 | 0.250 | 0.1 | 0.093 | 0.099 | 0.093 | 0.093 | 0.097 |
| | | | (6.3) | (−0.6) | (−0.4) | (3.2) | | | (5.8) | (−0.4) | (−0.2) | (3.4) |

Table 2: Values of relative biases and risks of the MSE estimators where $\psi = 1$ and it is estimated by the Prasad-Rao estimator $\widehat{\psi}^{PR}$

| | | $M_{V0}$ | $M_{V0.5}$ | | | $M_{V1}$ | | | $M_M$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $d_i$ | $m^*_{CEB}$ | $m^*_{CEB}$ | $m^*_{V0.5}$ | $m_{V0.5}$ | $m^*_{CEB}$ | $m^*_{V1}$ | $m_{V1}$ | $m^*_{CEB}$ | $m_M$ |
| | | | | | **bias in pattern (a)** | | | | | |
| $G_1$ | 0.7 | −0.85 | −1.86 | −1.39 | 2.66 | −2.15 | −2.14 | 2.31 | −1.95 | 2.49 |
| $G_2$ | 0.6 | −0.49 | −1.59 | −1.02 | 3.10 | −1.92 | −1.87 | 2.56 | −1.71 | 2.81 |
| $G_3$ | 0.5 | −0.11 | −1.31 | −0.60 | 3.57 | −1.67 | −1.57 | 2.81 | −1.42 | 3.16 |
| $G_4$ | 0.4 | 0.36 | −0.97 | −0.10 | 4.10 | −1.38 | −1.21 | 3.07 | −1.08 | 3.58 |
| $G_5$ | 0.3 | 0.81 | −0.60 | 0.45 | 4.67 | −1.06 | −0.82 | 3.31 | −0.67 | 4.07 |
| | | | | | **risk in pattern (a)** | | | | | |
| $G_1$ | 0.7 | 6.44 | 7.17 | 7.67 | 3.69 | 7.51 | 7.62 | 3.56 | 6.78 | 3.11 |
| $G_2$ | 0.6 | 5.55 | 6.26 | 6.76 | 2.87 | 6.60 | 6.71 | 2.72 | 5.89 | 2.32 |
| $G_3$ | 0.5 | 4.68 | 5.28 | 5.77 | 2.06 | 5.61 | 5.71 | 1.90 | 4.94 | 1.56 |
| $G_4$ | 0.4 | 3.78 | 4.20 | 4.69 | 1.31 | 4.53 | 4.63 | 1.13 | 3.93 | 0.88 |
| $G_5$ | 0.3 | 2.73 | 3.03 | 3.51 | 0.70 | 3.35 | 3.44 | 0.51 | 2.89 | 0.38 |
| | | | | | **bias in pattern (b)** | | | | | |
| $G_1$ | 4.0 | −9.91 | −10.93 | −11.99 | 7.43 | −11.23 | −11.63 | 8.46 | −11.18 | 8.32 |
| $G_2$ | 0.6 | −5.81 | −8.10 | −8.76 | 44.07 | −8.73 | −8.87 | 43.76 | −8.25 | 43.95 |
| $G_3$ | 0.5 | −5.35 | −7.88 | −8.49 | 48.66 | −8.59 | −8.69 | 48.14 | −7.99 | 48.47 |
| $G_4$ | 0.4 | −4.63 | −7.53 | −8.07 | 54.16 | −8.32 | −8.39 | 53.30 | −7.57 | 53.82 |
| $G_5$ | 0.1 | −0.00 | −3.98 | −4.07 | 63.09 | −5.06 | −4.97 | 60.83 | −2.59 | 61.98 |
| | | | | | **risk in pattern (b)** | | | | | |
| $G_1$ | 4.0 | 39.67 | 43.74 | 45.15 | 27.39 | 45.19 | 45.51 | 28.02 | 44.04 | 27.39 |
| $G_2$ | 0.6 | 12.61 | 14.69 | 15.52 | 25.76 | 15.42 | 15.68 | 25.60 | 14.09 | 25.97 |
| $G_3$ | 0.5 | 10.70 | 12.69 | 13.48 | 36.35 | 13.39 | 13.65 | 35.93 | 12.03 | 36.53 |
| $G_4$ | 0.4 | 8.66 | 10.49 | 11.23 | 52.54 | 11.17 | 11.41 | 51.67 | 9.75 | 52.62 |
| $G_5$ | 0.1 | 1.62 | 2.17 | 2.52 | 111.05 | 2.58 | 2.70 | 107.88 | 2.15 | 109.80 |

Table 3: Values of relative biases and risks of the MSE estimators where $\psi = 1$ and it is estimated by the Fay-Herriot estimator $\widehat{\psi}^{FH}$

| | | $M_{V0}$ | $M_{V0.5}$ | | | $M_{V1}$ | | | $M_M$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $d_i$ | $m^*_{CEB}$ | $m^*_{CEB}$ | $m^*_{V0.5}$ | $m_{V0.5}$ | $m^*_{CEB}$ | $m^*_{V1}$ | $m_{V1}$ | $m^*_{CEB}$ | $m_M$ |
| | | | | | **bias in pattern (a)** | | | | | |
| $G_1$ | 0.7 | −0.96 | −1.93 | −1.38 | 2.20 | −2.21 | −2.15 | 1.72 | −2.02 | 1.96 |
| $G_2$ | 0.6 | −0.55 | −1.61 | −0.97 | 2.59 | −1.93 | −1.84 | 1.98 | −1.72 | 2.27 |
| $G_3$ | 0.5 | −0.08 | −1.23 | −0.47 | 3.05 | −1.58 | −1.46 | 2.28 | −1.34 | 2.65 |
| $G_4$ | 0.4 | 0.50 | −0.75 | 0.14 | 3.60 | −1.15 | −0.99 | 2.64 | −0.87 | 3.11 |
| $G_5$ | 0.3 | 1.08 | −0.20 | 0.88 | 4.23 | −0.63 | −0.43 | 3.01 | −0.29 | 3.66 |
| | | | | | **risk in pattern (a)** | | | | | |
| $G_1$ | 0.7 | 6.17 | 6.95 | 7.42 | 3.83 | 7.26 | 7.36 | 3.70 | 6.58 | 3.24 |
| $G_2$ | 0.6 | 5.32 | 6.05 | 6.53 | 3.02 | 6.38 | 6.47 | 2.87 | 5.72 | 2.46 |
| $G_3$ | 0.5 | 4.51 | 5.09 | 5.58 | 2.21 | 5.41 | 5.51 | 2.05 | 4.80 | 1.69 |
| $G_4$ | 0.4 | 3.68 | 4.05 | 4.53 | 1.44 | 4.37 | 4.46 | 1.27 | 3.82 | 0.98 |
| $G_5$ | 0.3 | 2.73 | 2.92 | 3.40 | 0.79 | 3.23 | 3.32 | 0.62 | 2.81 | 0.43 |
| | | | | | **bias in pattern (b)** | | | | | |
| $G_1$ | 4.0 | −3.62 | −3.70 | −3.89 | −1.33 | −3.71 | −3.82 | −1.02 | −3.68 | −1.00 |
| $G_2$ | 0.6 | −1.25 | −2.02 | −1.91 | 1.55 | −2.23 | −2.24 | 1.40 | −2.08 | 1.65 |
| $G_3$ | 0.5 | −0.80 | −1.64 | −1.46 | 1.97 | −1.88 | −1.86 | 1.72 | −1.71 | 2.02 |
| $G_4$ | 0.4 | −0.20 | −1.16 | −0.90 | 2.46 | −1.43 | −1.39 | 2.09 | −1.23 | 2.47 |
| $G_5$ | 0.1 | 1.72 | 0.68 | 1.38 | 3.80 | 0.36 | 0.48 | 2.86 | 0.83 | 3.86 |
| | | | | | **risk in pattern (b)** | | | | | |
| $G_1$ | 4.0 | 22.51 | 23.21 | 23.45 | 21.19 | 23.43 | 23.46 | 21.29 | 22.97 | 20.89 |
| $G_2$ | 0.6 | 6.14 | 6.79 | 7.12 | 3.67 | 7.03 | 7.10 | 3.61 | 6.61 | 3.39 |
| $G_3$ | 0.5 | 5.16 | 5.72 | 6.04 | 2.70 | 5.95 | 6.03 | 2.63 | 5.54 | 2.43 |
| $G_4$ | 0.4 | 4.15 | 4.57 | 4.88 | 1.76 | 4.79 | 4.86 | 1.68 | 4.38 | 1.52 |
| $G_5$ | 0.1 | 0.96 | 0.73 | 0.93 | 0.16 | 0.85 | 0.89 | 0.09 | 0.89 | 0.20 |

Table 4: Values of the EBLUP and the variance constrained EBLUPs with their estimates of the MSE (the MSE estimates are given in parenthesis)

| Prefecture | $d_i$ | $y_i$ | $\boldsymbol{x}_i'\widehat{\boldsymbol{\beta}}$ | $EB$ | $cEB_{V0}$ | $cEB_{V0.5}$ | $cEB_{V1}$ |
|---|---|---|---|---|---|---|---|
| Ibaraki | 12.49 | 8.10 | 9.45 | 8.77 | 8.72 | 8.76 | 8.77 |
| | | | | (5.34) | (5.90) | (5.45) | (5.38) |
| Tochigi | 62.56 | 10.03 | 9.48 | 9.57 | 9.71 | 9.59 | 9.58 |
| | | | | (7.05) | (7.76) | (7.18) | (7.09) |
| Gunma | 5.38 | 5.21 | 9.99 | 6.63 | 6.08 | 6.54 | 6.61 |
| | | | | (3.48) | (4.20) | (3.57) | (3.51) |
| Saitama | 9.01 | 12.33 | 14.30 | 13.14 | 14.12 | 13.30 | 13.17 |
| | | | | (5.31) | (6.81) | (5.43) | (5.34) |
| Chiba | 91.77 | 30.71 | 12.17 | 14.43 | 15.71 | 14.63 | 14.46 |
| | | | | (7.73) | (9.96) | (7.87) | (7.75) |
| Tokyo | 3.65 | 15.45 | 13.16 | 14.94 | 16.34 | 15.17 | 14.98 |
| | | | | (2.79) | (5.01) | (2.90) | (2.81) |
| Kanagawa | 27.48 | 23.25 | 12.54 | 15.93 | 17.56 | 16.19 | 15.97 |
| | | | | (7.21) | (10.62) | (7.42) | (7.25) |

based on data of the disbursement 'Education' at the same city every November in the past ten years. Although the average disbursements in SFIE are reported every month, the sample sizes are around 100 for most prefectures, and data of the item 'Education' have high variability. On the other hand, we have data in the National Survey of Family Income and Expenditure (NSFIE) for 47 prefectures. Since NSFIE is based on much larger sample than SFIE, the average disbursements in NSEDI are more reliable, but this survey has been implemented every five years. In this study, we use the data of the item 'Education' of NSFIE in 2009, which is denoted by $X_i$ for $i = 1, \ldots, 47$. Thus, we apply the Fay-Herriot model (4.1) or

$$y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + v_i + \varepsilon_i, \quad i = 1, \ldots, k,$$

where $\boldsymbol{x}_i' = (1, X_i)$, $\boldsymbol{\beta} = (\beta_1, \beta_2)'$ and $k = 47$.

The Fay-Herriot estimate of $\psi$ is 12.752, and the GLS estimates of $\beta_1$ and $\beta_2$ are 2.209 and 0.580, respectively. The EBLUP and the constrained EBLUP are given around (4.2). We provide the values of the EBLUP $EB$ and the variance constrained EBLUPs $cEB_{V0}$, $cEB_{V0.5}$ and $cEB_{V1}$ for $r = 0$, 0.5 and 1. It is noted that the mean constrained EBLUP is identical to the EBLUP $EB$ in this case. These values in seven prefectures around Tokyo are reported in Table 4 with the estimates of their MSEs based on $m^*_{CEB}$ given in (4.5). As seen from the table, the EBLUP $EB$ shrinks $y_i$ more toward $\boldsymbol{x}_i'\widehat{\boldsymbol{\beta}}$ for larger $d_i$. It is seen that $cEB_{V0}$ is more variable than $cEB_{V0.5}$ and $cEB_{V1}$, and that the value of $cEB_{V0}$ in Tokyo is beyond the range between $y_i$ and $\boldsymbol{x}_i'\widehat{\boldsymbol{\beta}}$, while the values of the EBLUP and the other constrained EBLUPs are between those values. Also, the MSE estimate of $cEB_{V0}$

over that of $EB$ is about 1.8 for Tokyo. Compared with $cEB_{V0}$, the estimates of $cEB_{V0.5}$ and $cEB_{V1}$ and the estimates of their MSEs are close to those of $EB$.

# 5 Concluding Remarks

In this paper, we have obtained the unified constrained estimator for the general mean-variance constraints based on the constrained Bayes estimator, and have derived asymptotic approximations of MSE of the constrained EBLUP. Using this result, we have shown that when the variance of estimates is constrained to be equal to the variance of prior distribution, the resulting constrained EBLUP has a larger MSE than EBLUP in the first order. This may be against the aim of EBLUP, since EBLUP is suggested to increase the precision of the estimates. Thus, we have considered to modify the variance constraints so that MSE of the constrained EBLUP is equal to MSE of EBLUP in the first order, and then we have derived a second order approximation of the MSE. Also, we have suggested an estimator of MSE of the constrained EBLUP based on the parametric bootstrap method, and have shown that it is a second order unbiased estimator of MSE. The performances of MSEs of the constrained EBLUPs and their estimators have been investigated by simulation and empirical studies.

**Acknowledgments.**

I would like to thank Dr. T. Maiti, Michigan State University, for his valuable suggestions. Research of the author was supported in part by Grant-in-Aid for Scientific Research (21540114 and 23243039) from Japan Society for the Promotion of Science.

# References

[1] Battese, G.E., Harter, R.M. and Fuller, W.A. (1988). An error-components model for prediction of county crop areas using survey and satellite data. *J. Amer. Statist. Assoc.*, **83**, 28-36.

[2] Butar, F.B. and Lahiri, P. (2003). On measures of uncertainty of empirical Bayes small-area estimators. *J. Statist. Plan. Inf.*, **112**, 63-76.

[3] Chatterjee, S., Lahiri, P., and Li, H. (2008). Parametric bootstrap approximation to the distribution of EBLUP and related prediction intervals in linear mixed models. *Ann. Statist.*, **36**, 1221-1245.

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

[5] Datta, G.S. and Lahiri, P. (2000). A unified measure of uncertainty of estimated best linear unbiased predictors in small area estimation problems. *Statist. Sinica*, **10**, 613-627.

[6] Datta, G.S., Rao, J.N.K. and Smith, D.D. (2005). On measuring the variability of small area estimators under a basic area level model. *Biometrika*, **92**, 183-196.

[7] Fay, R.E. and Herriot, R. (1979). Estimates of income for small places: An application of James-Stein procedures to census data. *J. Amer. Statist. Assoc.*, **74**, 269-277.

[8] Ghosh, M. (1992). Constrained Bayes estimation with applications. *J. American Statist. Assoc.*, **87**, 533-540.

[9] Ghosh, M. and Rao, J.N.K. (1994). Small area estimation: An appraisal. *Statist. Science*, **9**, 55-93.

[10] Hall, P. and Maiti, T. (2006). Nonparametric estimation of mean-squared prediction error in nested-error regression models. *Ann. Statist.*, **34**, 1733-1750.

[11] Kubokawa, T. (2011) On Measuring Uncertainty of Small Area Estimators with Higher Order Accuracy Discussion Paper Series, CIRJE-F-754. *J. Japan Statist. Soc.*, to appear.

[12] Kubokawa, T. and Nagashima, B. (2012). Parametric bootstrap methods for bias correction in linear mixed models. Discussion Paper Series, CIRJE-F-801. *J. Multivariate Anal.*, to appear.

[13] Louis, T.A. (1984). Estimating a population of parameter values using Bayes and empirical Bayes methods. *J. Amer. Statist. Assoc.*, **79**, 393-398.

[14] Pfeffermann, D. (2002). Small area estimation - new developments and directions. *Int. Statist. Rev.*, **70**, 125-143.

[15] Prasad, N.G.N. and Rao, J.N.K. (1990). The estimation of the mean squared error of small-area estimators. *J. Amer. Statist. Assoc.*, **85**, 163-171.

[16] Rao, J.N.K. (2003). *Small Area Estimation.* Wiley, New Jersey.

[17] Stein, C. (1981). Estimation of the mean of a multivariate normal distribution. *Ann. Statist.*, **9**, 1135-1151.

# A Appendix

## A.1 Assumptions and lemmas

We here give analytical results and their proofs for asymptotic approximation of MSE of $\widehat{\mu}^{CEB}$.

We begin by introducing the notations used here. Let $\mathcal{C}_{\boldsymbol{\psi}}^{[k]}$ denote a set of $k$ times continuously differentiable functions with respect to $\boldsymbol{\psi}$. For partial derivatives with respect to $\boldsymbol{\psi}$, we utilize the notations

$$\boldsymbol{A}_{(i)}(\boldsymbol{\psi}) = \frac{\partial \boldsymbol{A}(\boldsymbol{\psi})}{\partial \psi_i}, \quad \boldsymbol{A}_{(ij)}(\boldsymbol{\psi}) = \partial_{ij} \boldsymbol{A}(\boldsymbol{\psi}) = \frac{\partial^2 \boldsymbol{A}(\boldsymbol{\psi})}{\partial \psi_i \partial \psi_j},$$

and

$$\boldsymbol{A}_{(ijk)}(\boldsymbol{\psi}) = \partial_{ijk}\boldsymbol{A}(\boldsymbol{\psi}) = \frac{\partial^3 \boldsymbol{A}(\boldsymbol{\psi})}{\partial\psi_i\partial\psi_j\partial\psi_k},$$

where $\boldsymbol{A}(\boldsymbol{\psi})$ is a scalar, vector or matrix.

For $0 \le i, j, k \le q$, let $\lambda_1(\boldsymbol{\Sigma}) \le \cdots \le \lambda_N(\boldsymbol{\Sigma})$ be the eigenvalues of $\boldsymbol{\Sigma}$ and let those of $\boldsymbol{\Sigma}_{(i)}$, $\boldsymbol{\Sigma}_{(ij)}$ and $\boldsymbol{\Sigma}_{(ijk)}$ be $\lambda_a^i(\boldsymbol{\Sigma})$, $\lambda_a^{ij}(\boldsymbol{\Sigma})$ and $\lambda_a^{ijk}(\boldsymbol{\Sigma})$ for $a = 1, \ldots, N$ respectively, where $|\lambda_1^i(\boldsymbol{\Sigma})| \le \cdots \le |\lambda_N^i(\boldsymbol{\Sigma})|$, $|\lambda_1^{ij}(\boldsymbol{\Sigma})| \le \cdots \le |\lambda_N^{ij}(\boldsymbol{\Sigma})|$ and $|\lambda_1^{ijk}(\boldsymbol{\Sigma})| \le \cdots \le |\lambda_N^{ijk}(\boldsymbol{\Sigma})|$.

Throughout the paper, we assume the following conditions for large $N$ and $1 \le i, j, k \le q$:

(A1) The elements of $\boldsymbol{X}$, $\boldsymbol{Z}$, $\boldsymbol{Q}$, $\boldsymbol{R}$ and $\boldsymbol{c}$ are uniformly bounded, $p$ and $q$ are bounded, and $\boldsymbol{Z}'\boldsymbol{Z} = [O(1)]_{M\times M}$, $\boldsymbol{Z}\boldsymbol{Z}' = [O(1)]_{N\times N}$, where $[O(1)]_{M\times M}$ means that every element of the matrix is of $O(1)$. The matrices $\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X}$ and $\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W}$ are positive definite and $\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X}/N$ and $\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{W}/N$ converge to positive definite matrices;

(A2) (i) $\boldsymbol{\Sigma}(\boldsymbol{\psi}) \in \mathcal{C}_{\boldsymbol{\psi}}^{[2]}$, and $\lim_{N\to\infty}\lambda_1 > 0$, $\lim_{N\to\infty}\lambda_N < \infty$, $\lim_{N\to\infty}|\lambda_N^i| < \infty$ and $\lim_{N\to\infty}|\lambda_N^{ij}| < \infty$. (ii) $\boldsymbol{s} = \boldsymbol{s}(\boldsymbol{\psi}) \in \mathcal{C}_{\boldsymbol{\psi}}^{[2]}$, and $(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'\boldsymbol{s}(\boldsymbol{\psi}) = O_p(1)$, $(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'\boldsymbol{s}_{(i)} = O_p(1)$, $(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'\boldsymbol{s}_{(ij)} = O_p(1)$ and $\boldsymbol{s}_{(i)}'\boldsymbol{s}_{(j)} = O(1)$.

(A3) $\widehat{\boldsymbol{\psi}} = \widehat{\boldsymbol{\psi}}(\boldsymbol{y}) = (\widehat{\psi}_1, \ldots, \widehat{\psi}_q)'$ is an estimator of $\boldsymbol{\psi}$ which satisfies that $\widehat{\boldsymbol{\psi}}(-\boldsymbol{y}) = \widehat{\boldsymbol{\psi}}(\boldsymbol{y})$ and $\widehat{\boldsymbol{\psi}}(\boldsymbol{y} + \boldsymbol{X}\boldsymbol{\alpha}) = \widehat{\boldsymbol{\psi}}(\boldsymbol{y})$ for any $p$-dimensional vector $\boldsymbol{\alpha}$.

(A4) $\widehat{\boldsymbol{\psi}} - \boldsymbol{\psi}$ is expanded as

$$\widehat{\boldsymbol{\psi}} - \boldsymbol{\psi} = \widehat{\boldsymbol{\psi}}^{\dagger} + \widehat{\boldsymbol{\psi}}^{\dagger\dagger} + O_p(N^{-3/2}), \tag{A.1}$$

where $\widehat{\boldsymbol{\psi}}^{\dagger} = O_p(N^{-1/2})$ and $\widehat{\boldsymbol{\psi}}^{\dagger\dagger} = O_p(N^{-1})$. For $\widehat{\boldsymbol{\psi}}^{\dagger} = (\widehat{\psi}_1^{\dagger}, \ldots, \widehat{\psi}_q^{\dagger})'$, it is assumed that $\widehat{\psi}_i^{\dagger}$ satisfies that (i) $E[\widehat{\psi}_i^{\dagger}] = O(N^{-1})$, (ii) $\boldsymbol{s}_{(j)}'\boldsymbol{\Sigma}\boldsymbol{\nabla}_y\widehat{\psi}_i^{\dagger} = O_p(N^{-1})$ and (iii) $\boldsymbol{c}'\boldsymbol{\nabla}_y\boldsymbol{\nabla}_y'\widehat{\psi}_a^{\dagger}\boldsymbol{s}_{(a)} = O_p(N^{-1})$.

(A5) The functions $\boldsymbol{t}_1(\boldsymbol{y})$ and $t_2(\boldsymbol{y})$ in the constraints (C1) and (C2) are expressed as

$$\begin{aligned} \boldsymbol{t}_1(\boldsymbol{y}) =& \boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{\theta}}^{EB} + \boldsymbol{m}(\boldsymbol{y}), \quad \boldsymbol{m}(\boldsymbol{y}) = O_p(N^{1/2}), \\ t_2(\boldsymbol{y}) =& (\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB} + h(\widehat{\boldsymbol{\psi}}), \quad h(\boldsymbol{\psi}) = O(N^{1-r}), \end{aligned} \tag{A.2}$$

for $0 \le r \le 1$.

We shall derive approximations of MSE of the constrained EBLUP $\widehat{\mu}^{CEB}$ given in (3.2) under the general constraints given in (A5). It is noted that $MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB}) = I_1 + I_2 + 3I_3$, where

$$\begin{aligned} I_1 =& E[(\widehat{\mu}^{EB} - \mu)^2], \\ I_2 =& E\big[\big\{\{a(\boldsymbol{y}) - 1\}\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\Omega}\widehat{\boldsymbol{\theta}}^{EB} + \boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})\big\}^2\big], \\ I_3 =& E\big[(\widehat{\mu}^{EB} - \mu)\big\{\{a(\boldsymbol{y}) - 1\}\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\Omega}\widehat{\boldsymbol{\theta}}^{EB} + \boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})\big\}\big], \end{aligned} \tag{A.3}$$

for $\widehat{\boldsymbol{\theta}}^{EB}=\boldsymbol{X}\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\psi}})+\boldsymbol{Z}\widehat{\boldsymbol{Q}}\boldsymbol{Z}'\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{y}-\boldsymbol{X}\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\psi}}))$ and $\widehat{\mu}^{EB}=\boldsymbol{c}'\widehat{\boldsymbol{\theta}}^{EB}$. Since $\widehat{\mu}^{B}=E[\mu|\boldsymbol{y}]=\boldsymbol{c}'\widehat{\boldsymbol{\theta}}^{B}$ for $\widehat{\boldsymbol{\theta}}^{B}=\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})$, $I_3$ is rewritten as

$$I_3=E\big[(\widehat{\mu}^{EB}-\widehat{\mu}^{B})\big\{\{a(\boldsymbol{y})-1\}\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB}+\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})\big\}\big]. \tag{A.4}$$

The following lemmas are useful for the purpose.

**Lemma A.1 (Stein identity)** *Let $\boldsymbol{y}\sim\mathcal{N}_N(\boldsymbol{X}\boldsymbol{\beta},\boldsymbol{\Sigma})$. Then,*

$$E[(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})'\boldsymbol{g}(\boldsymbol{y})]=E[\boldsymbol{\nabla}_{\boldsymbol{y}}'\{\boldsymbol{\Sigma}\boldsymbol{g}(\boldsymbol{y})\}], \tag{A.5}$$

*where $\boldsymbol{g}(\boldsymbol{y})=(g_1(\boldsymbol{y}),\ldots,g_N(\boldsymbol{y}))'$ is an absolutely continuous function and $\boldsymbol{\nabla}_{\boldsymbol{y}}$ is the differential operator defined by $\boldsymbol{\nabla}_{\boldsymbol{y}}=\partial/\partial\boldsymbol{y}$. Let $\boldsymbol{A}$ be an $N\times N$ matrix independent of $\boldsymbol{y}$, and let $f(\boldsymbol{y})$ be a scalar function which is twice-differentiable with respect to $\boldsymbol{y}$. Then,*

$$E[\boldsymbol{u}'\boldsymbol{A}\boldsymbol{u}f(\boldsymbol{y})]=\mathrm{tr}\,[\boldsymbol{\Sigma}\boldsymbol{A}]E[f(\boldsymbol{y})]+\mathrm{tr}\,[\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma}E[\boldsymbol{\nabla}_{\boldsymbol{y}}\boldsymbol{\nabla}_{\boldsymbol{y}}'f(\boldsymbol{y})]]. \tag{A.6}$$

The identity (A.5) is from Stein (1973), and the equation (A.6) can be derived by using the Stein identity.

**Lemma A.2** *Assume the conditions* (A1)-(A6)*. Then,*

$$(\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB}/N=\nu_0+\nu_1, \tag{A.7}$$

*where*

$$\nu_0=\{\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{P}_{\boldsymbol{\Omega}}\boldsymbol{X}\boldsymbol{\beta}+\mathrm{tr}\,[\boldsymbol{Z}'\boldsymbol{P}_{\boldsymbol{\Omega}}\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z}]\}/N,$$

*and $\nu_1$ is a function with $O_p(N^{-1/2})$. Also,*

$$a(\boldsymbol{y})-1=A(\boldsymbol{\omega})-1+O_p(N^{-1/2-r}), \tag{A.8}$$

*where $A(\boldsymbol{\omega})-1=O(N^{-r})$ for $A(\boldsymbol{\omega})$ defined in* (3.6)*.*

**Proof.** To verify (A.7), note that $\widehat{\boldsymbol{\theta}}^{EB}$ is rewritten as $\widehat{\boldsymbol{\theta}}^{EB}=\widehat{\boldsymbol{\theta}}^{B}+(\widehat{\boldsymbol{\theta}}^{EB}-\widehat{\boldsymbol{\theta}}^{B})$, where

$$\widehat{\boldsymbol{\theta}}^{EB}-\widehat{\boldsymbol{\theta}}^{B}=\widehat{\boldsymbol{\theta}}^{B}+\widehat{\boldsymbol{R}}\widehat{\boldsymbol{\Sigma}}^{-1}\{\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\psi}})-\boldsymbol{\beta}\}+(\boldsymbol{Z}\widehat{\boldsymbol{Q}}\boldsymbol{Z}'\widehat{\boldsymbol{\Sigma}}^{-1}-\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1})(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta}),$$

and it is seen that $\widehat{\boldsymbol{\theta}}^{EB}-\widehat{\boldsymbol{\theta}}^{B}=[O_p(N^{-1/2})]_{N\times N}$. Also note that $E[(\widehat{\boldsymbol{\theta}}^{B})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{B}]=\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{P}_{\boldsymbol{\Omega}}\boldsymbol{X}\boldsymbol{\beta}+\mathrm{tr}\,[\boldsymbol{Z}'\boldsymbol{P}_{\boldsymbol{\Omega}}\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z}]$. Then,

$$\begin{aligned}(\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB}=&(\widehat{\boldsymbol{\theta}}^{B})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{B}+2(\widehat{\boldsymbol{\theta}}^{EB}-\widehat{\boldsymbol{\theta}}^{B})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{B}+(\widehat{\boldsymbol{\theta}}^{EB}-\widehat{\boldsymbol{\theta}}^{B})'\boldsymbol{P}_{\boldsymbol{\Omega}}(\widehat{\boldsymbol{\theta}}^{EB}-\widehat{\boldsymbol{\theta}}^{B})\\=&(\widehat{\boldsymbol{\theta}}^{B})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{B}+O_p(N^{1/2}),\end{aligned}$$

since $(\widehat{\boldsymbol{\theta}}^{EB}-\widehat{\boldsymbol{\theta}}^{B})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{B}=O_p(N^{1/2})$ and $(\widehat{\boldsymbol{\theta}}^{EB}-\widehat{\boldsymbol{\theta}}^{B})'\boldsymbol{P}_{\boldsymbol{\Omega}}(\widehat{\boldsymbol{\theta}}^{EB}-\widehat{\boldsymbol{\theta}}^{B})=O_p(1)$. Thus, it is sufficient to verify that $(\widehat{\boldsymbol{\theta}}^{B})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{B}=E[(\widehat{\boldsymbol{\theta}}^{B})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{B}]+O_p(N^{1/2})$. Since

$$\frac{1}{N}E\Big[\Big\{(\widehat{\boldsymbol{\theta}}^{B})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{B}-E[(\widehat{\boldsymbol{\theta}}^{B})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{B}]\Big\}^2\Big]=\frac{2}{N}\mathrm{tr}\,\Big[\Big\{\boldsymbol{Z}'\boldsymbol{P}_{\boldsymbol{\Omega}}\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z}\Big\}^2\Big],$$

which is of $O(1)$ from (A1). This proves the approximation (A.7).

To verify (A.8), note that $a(\boldsymbol{y})-1$ is rewritten as

$$a(\boldsymbol{y})-1=\frac{\{a(\boldsymbol{y})\}^2-1}{a(\boldsymbol{y})+1}=\frac{h(\widehat{\boldsymbol{\psi}})}{(\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB}\{a(\boldsymbol{y})+1\}}.$$

It is here noted that

$$\begin{aligned}\frac{h(\widehat{\boldsymbol{\psi}})}{N}&=\frac{h(\boldsymbol{\psi})}{N}+\sum_a\frac{h_{(a)}(\boldsymbol{\psi})}{N}(\widehat{\psi}_a-\psi_a)+O_p(N^{-1-r}),\\ \frac{N}{(\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB}}&=\frac{1}{\nu_0}-\frac{\nu_1}{\nu_0}+O_p(N^{-1}).\end{aligned}$$

Since $1/\{1+a(\boldsymbol{y})\}$ can be similarly approximated as

$$\{1+a(\boldsymbol{y})\}^{-1}=\{1+\sqrt{1+h(\boldsymbol{\psi})/(N\nu_0)}\}^{-1}+O_p(N^{-1/2-r}),$$

we can see that

$$\begin{aligned}\frac{h(\widehat{\boldsymbol{\psi}})}{(\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB}\{a(\boldsymbol{y})+1\}}=&\frac{h(\boldsymbol{\psi})}{N\nu_0\{1+\sqrt{1+h(\boldsymbol{\psi})/(N\nu_0)}\}}+O_p(N^{-1/2-r}),\\ =&\{A(\boldsymbol{\omega})-1\}+O_p(N^{-1/2-r}).\end{aligned}$$

Clearly, $A(\boldsymbol{\omega})-1=O(N^{-r})$ and we get Lemma A.2. ■

**Lemma A.3** *Assume the conditions* (A1)-(A6). *Then,* $I_3=O(N^{-1})$ *for* $r\geq 0$.

**Proof.** Note that $I_3=I_{31}+I_{32}$, where

$$\begin{aligned}I_{31}=&E\big[(\widehat{\mu}^{EB}-\widehat{\mu}^B)\{a(\boldsymbol{y})-1\}\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB}\big],\\ I_{32}=&E\big[(\widehat{\mu}^{EB}-\widehat{\mu}^B)\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})\big].\end{aligned}$$

Since $\widehat{\mu}^B=\boldsymbol{c}'\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{s}'(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})$ for $\boldsymbol{s}=\boldsymbol{s}(\boldsymbol{\psi})=\boldsymbol{\Sigma}^{-1}\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{c}$, it is seen that

$$\widehat{\mu}^{EB}-\widehat{\mu}^B=(\boldsymbol{c}-\widehat{\boldsymbol{s}})'\boldsymbol{X}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})+(\widehat{\boldsymbol{s}}-\boldsymbol{s})'(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta}),\tag{A.9}$$

for $\widehat{\boldsymbol{s}}=\boldsymbol{s}(\widehat{\boldsymbol{\psi}})=\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{Z}\widehat{\boldsymbol{Q}}\boldsymbol{Z}'\boldsymbol{c}$. Clearly, $\widehat{\mu}^{EB}-\widehat{\mu}^B=O_p(N^{-1/2})$. Since $\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})=O_p(N^{-1/2})$, it is observed that $I_{32}=O(N^{-1})$.

To evaluate $I_{31}$, we approximate $\widehat{\mu}^{EB}-\widehat{\mu}^B$ as

$$\begin{aligned}\widehat{\mu}^{EB}-\widehat{\mu}^B=&(\boldsymbol{c}-\boldsymbol{s})'\boldsymbol{X}(\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\psi}})-\boldsymbol{\beta})+(\widehat{\boldsymbol{s}}-\boldsymbol{s})'(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})+O_p(N^{-1})\\ =&(\boldsymbol{c}-\boldsymbol{s})'\boldsymbol{X}(\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\psi}})-\boldsymbol{\beta})+\sum_{a=1}^q\boldsymbol{s}_{(a)}'(\widehat{\psi}_a-\psi_a)(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})+O_p(N^{-1}),\end{aligned}$$

where $\boldsymbol{s}_{(a)} = \partial \boldsymbol{s}(\boldsymbol{\psi})/\partial \psi_a$. Since $\widehat{\boldsymbol{\beta}}_{(a)}(\boldsymbol{\psi}) = \partial\widehat{\boldsymbol{\beta}}(\boldsymbol{\psi})/\partial\psi_a = O_p(N^{-1/2})$, it is noted that $\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\psi}}) = \widehat{\boldsymbol{\beta}}(\boldsymbol{\psi}) + \sum_{a=1}^q \widehat{\boldsymbol{\beta}}_{(a)}(\boldsymbol{\psi})(\widehat{\psi}_a - \psi_a) + O_p(N^{-3/2})$, namely, $\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\psi}}) - \widehat{\boldsymbol{\beta}}(\boldsymbol{\psi}) = O_p(N^{-1})$. Then, $\widehat{\mu}^{EB} - \widehat{\mu}^B$ can be further approximated as

$$\begin{aligned}\widehat{\mu}^{EB} - \widehat{\mu}^B =&(\boldsymbol{c}-\boldsymbol{s})'\boldsymbol{X}(\widehat{\boldsymbol{\beta}}(\boldsymbol{\psi}) - \boldsymbol{\beta}) + \sum_{a=1}^q \boldsymbol{s}'_{(a)}(\widehat{\psi}_a - \psi_a)(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) + O_p(N^{-1})\\ =&(\boldsymbol{c}-\boldsymbol{s})'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})\\ &+\sum_{a=1}^q (\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})'\boldsymbol{s}_{(a)}\widehat{\psi}_a^\dagger + O_p(N^{-1}),\end{aligned} \tag{A.10}$$

From Lemma A.2, it follows that $a(\boldsymbol{y})-1 = A(\boldsymbol{\omega})-1+O_p(N^{-1/2-r})$. Since $\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB} = \boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{B} + \boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}(\widehat{\boldsymbol{\theta}}^{EB} - \widehat{\boldsymbol{\theta}}^{B}) = \boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{B} + O_p(N^{-1/2})$, $I_{31}$ can be approximated as

$$\begin{aligned}I_{31} =&E\Big[(\widehat{\mu}^{EB} - \widehat{\mu}^B)\big\{\{A(\boldsymbol{\omega})-1\} + O_p(N^{-1/2-r})\big\}\big\{\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{B} + O_p(N^{-1/2})\big\}\Big]\\ =&E\Big[(\widehat{\mu}^{EB} - \widehat{\mu}^B)\{A(\boldsymbol{\omega})-1\}\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{B}\Big] + O(N^{-1}) + O(N^{-1-r}).\end{aligned}$$

Using the approximation (A.10), we can evaluate $I_{31}$ as

$$\begin{aligned}I_{31} =&E\Big[\big\{(\boldsymbol{c}-\boldsymbol{s})'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta}) + \sum_{a=1}^q (\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})'\boldsymbol{s}_{(a)}\widehat{\psi}_a^\dagger\big\}\\ &\times \{A(\boldsymbol{\omega})-1\}\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\{\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})\}\Big] + O(N^{-1}).\end{aligned}$$

It is easy to see that

$$\begin{aligned}&E\big[(\boldsymbol{c}-\boldsymbol{s})'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\{\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})\}\big]\\ &\quad= (\boldsymbol{c}-\boldsymbol{s})'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{P}_{\boldsymbol{\Omega}}\boldsymbol{c},\end{aligned}$$

which is of $O(N^{-1})$. The Stein identity (A.5) and the equation (A.6) are applied to get that

$$\begin{aligned}&E\Big[(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})'\boldsymbol{s}_{(a)}\widehat{\psi}_a^\dagger\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\{\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})\}\Big]\\ &\quad= \boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\boldsymbol{X}\boldsymbol{\beta}E[\boldsymbol{s}'_{(a)}\boldsymbol{\nabla}_{\boldsymbol{y}}\widehat{\psi}_a^\dagger] + \boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\big\{\boldsymbol{s}_{(a)}E[\widehat{\psi}_a^\dagger] + \boldsymbol{\Sigma}E[\boldsymbol{\nabla}_y\boldsymbol{\nabla}'_{\boldsymbol{y}}\widehat{\psi}_a^\dagger]\boldsymbol{s}_{(a)}\big\},\end{aligned}$$

which can be verified to be of $O(N^{-1})$ from the condition (A4). These show that $I_{31} = O(N^{-1})$ and the proof is complete. ■

## A.2 Approximation of MSE

We now asymptotic approximations of MSE of the constrained EBLUP in the cases of $r = 0$, $r = 1/2$ and $r = 1$.

**[Scenario 1] Case of** $r = 0$**.** Note that $h(\boldsymbol{\psi}) = O(N)$ for $t_2(\boldsymbol{y})$ in (3.1). From Lemma A.3, it follows that

$$I_3 = O(N^{-1/2}). \tag{A.11}$$

We can obtain the limiting value of $I_2$ and get the following theorem.

**Theorem A.1** *Assume the conditions* (A1)-(A6) *and* $r = 0$. *Then, MSE of* $\widehat{\mu}^{CEB}$ *is approximated as*

$$MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB}) = MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB}) + \{A(\boldsymbol{\omega}) - 1\}^2 B(\boldsymbol{\omega}) + O(N^{-1/2}), \tag{A.12}$$

*where* $A(\boldsymbol{\omega})$ *and* $B(\boldsymbol{\omega})$ *are defined in* (3.6) *and* (3.7).

**Proof.** Note that $I_2$ is expressed as

$$I_2 = E\Big[\Big\{\frac{h(\widehat{\boldsymbol{\psi}})}{(\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\Omega}\widehat{\boldsymbol{\theta}}^{EB}\{a(\boldsymbol{y})+1\}}\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\Omega}\widehat{\boldsymbol{\theta}}^{EB} + \boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})\Big\}^2\Big]. \tag{A.13}$$

From (A.7), it is noted that

$$\{a(\boldsymbol{y})\}^2 = \frac{h(\widehat{\boldsymbol{\psi}})}{(\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\Omega}\widehat{\boldsymbol{\theta}}^{EB}} + 1 = \{A(\boldsymbol{\omega})\}^2 + O_p(N^{-1/2}),$$

where Since $\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y}) = O_p(N^{-1/2})$, it is seen that

$$\begin{aligned} I_2 =& E\Big[\{a(\boldsymbol{y}) - 1\}^2\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\Omega}\widehat{\boldsymbol{\theta}}^{EB}(\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\Omega}\boldsymbol{\Omega}^{-1}\boldsymbol{c}\Big] + O(N^{-1/2}) \\ =& \{A(\boldsymbol{\omega}) - 1\}^2 E\big[\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\Omega}\widehat{\boldsymbol{\theta}}^{B}(\widehat{\boldsymbol{\theta}}^{B})'\boldsymbol{P}_{\Omega}\boldsymbol{\Omega}^{-1}\boldsymbol{c}\big] + O(N^{-1/2}) \\ =& \{A(\boldsymbol{\omega}) - 1\}^2 B(\boldsymbol{\omega}) + O(N^{-1/2}), \end{aligned} \tag{A.14}$$

which proves the theorem. ∎

**[Scenario 2] Case of $r = 1/2$.** In this case, it is noted that $h(\boldsymbol{\psi}) = O(N^{1/2})$, and we get the following theorem.

**Theorem A.2** *Assume the conditions* (A1)-(A6) *and* $r = 1/2$. *Then, MSE of* $\widehat{\mu}^{CEB}$ *is approximated as*

$$MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB}) = MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB}) + \overline{I}_2(\boldsymbol{\omega}) + 2\overline{I}_3(\boldsymbol{\omega}) + O(N^{-3/2}). \tag{A.15}$$

*for* $\overline{I}_2(\boldsymbol{\omega})$ *and* $\overline{I}_3(\boldsymbol{\omega})$ *given in* (3.11) *and* (3.12).

**Proof.** Since $h(\boldsymbol{\psi}) = O(N^{1/2})$, it is seen that $\{a(\boldsymbol{y})\}^2 - 1$ is equal to

$$\begin{aligned} \frac{h(\widehat{\boldsymbol{\psi}})}{(\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\Omega}\widehat{\boldsymbol{\theta}}^{EB}} =& \frac{h(\boldsymbol{\psi})}{\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{P}_{\Omega}\boldsymbol{X}\boldsymbol{\beta} + \mathrm{tr}\,[\boldsymbol{Z}'\boldsymbol{P}_{\Omega}\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z}]} + O_p(N^{-1}) \\ =& \{A(\boldsymbol{\omega})\}^2 - 1 + O_p(N^{-1}), \end{aligned}$$

for $A(\boldsymbol{\omega})$ given in (3.6). Note that $\{a(\boldsymbol{y})\}^2 - 1 = O_p(N^{-1/2})$ and $\{A(\boldsymbol{\omega})\}^2 - 1 = O(N^{-1/2})$. Since $\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y}) = O_p(N^{-1/2})$, it is observed from (A.13) that $I_2 = O(N^{-1})$ and

$$\begin{aligned} I_2 =& E\Big[\Big\{\frac{h(\widehat{\boldsymbol{\psi}})}{(\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\Omega}\widehat{\boldsymbol{\theta}}^{EB}\{a(\boldsymbol{y})+1\}}\Big\}^2\big\{\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\Omega}\widehat{\boldsymbol{\theta}}^{EB}\big\}^2\Big] + E\big[\{\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})\}^2\big] \\ &+ 2E\Big[\frac{h(\widehat{\boldsymbol{\psi}})}{(\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\Omega}\widehat{\boldsymbol{\theta}}^{EB}\{a(\boldsymbol{y})+1\}}\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\Omega}\widehat{\boldsymbol{\theta}}^{EB}\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})\Big] \\ =& I_{21} + I_{22} + I_{23}. \quad \text{(say)} \end{aligned}$$

Noting that $a(\boldsymbol{y}) = \{1 + [\{a(\boldsymbol{y})\}^2 - 1]\}^{1/2} = 1 + O_p(N^{-1/2})$, we can evaluate $I_{21}$ as

$$
\begin{aligned}
I_{21} =& E\Big[\Big\{\frac{h(\widehat{\boldsymbol{\psi}})}{2(\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB}}\Big\}^2\{\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB}\}^2\Big] + O(N^{-3/2})\\
=& \frac{1}{4}\big[\{A(\boldsymbol{\omega})\}^2 - 1\big]^2 B(\boldsymbol{\omega}) + O(N^{-3/2}).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
I_{23} =& E\Big[\frac{h(\widehat{\boldsymbol{\psi}})}{2(\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB}}\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB}\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})\Big] + O(N^{-3/2})\\
=& \frac{1}{2}\big[\{A(\boldsymbol{\omega})\}^2 - 1\big]E\big[\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB}\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})\big] + O(N^{-3/2}).
\end{aligned}
$$

Since $\widehat{\boldsymbol{\theta}}^{EB} = \widehat{\boldsymbol{\theta}}^{B} + [O_p(N^{-1/2})]_{N\times 1}$, it is seen that

$$
\begin{aligned}
I_{23} =& 2^{-1}\big[\{A(\boldsymbol{\omega})\}^2 - 1\big]E\big[\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\{\boldsymbol{X\beta} + \boldsymbol{ZQZ}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{X\beta})\}\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})\big]\\
&+ O(N^{-3/2}).
\end{aligned}
\tag{A.16}
$$

The Stein identity given in (A.5) gives the expression

$$
\begin{aligned}
I_{23} =& 2^{-1}\big[\{A(\boldsymbol{\omega})\}^2 - 1\big]\big\{\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\boldsymbol{X\beta}E[\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})] + \boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\boldsymbol{ZQZ}'E[\boldsymbol{\nabla}_{\boldsymbol{y}}\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})]\big\}\\
&+ O(N^{-3/2}).
\end{aligned}
$$

Hence, $I_2$ is approximated as $I_2 = \overline{I}_2(\boldsymbol{\omega}) + O(N^{-3/2})$ for $\overline{I}_2(\boldsymbol{\omega})$ given in (3.11).

For $I_3$, it is noted that $I_3 = O(N^{-1})$ and $I_3$ is expressed as

$$
\begin{aligned}
I_3 =& E\Big[(\widehat{\mu}^{EB} - \widehat{\mu}^{B})\Big\{\frac{h(\widehat{\boldsymbol{\psi}})}{(\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB}\{a(\boldsymbol{y}) + 1\}}\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB} + \boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})\Big\}\Big]\\
=& E\Big[(\widehat{\mu}^{EB} - \widehat{\mu}^{B})\Big\{\frac{h(\widehat{\boldsymbol{\psi}})}{2(\widehat{\boldsymbol{\theta}}^{EB})'\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB}}\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{EB} + \boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})\Big\}\Big] + O(N^{-3/2}),
\end{aligned}
$$

Using the same argument as in the evaluation of $I_{23}$, we can see that $I_3 = I_{31} + I_{32} + O(N^{-3/2})$, where

$$
\begin{aligned}
I_{31} =& \frac{1}{2}\big[\{A(\boldsymbol{\omega})\}^2 - 1\big]E\Big[(\widehat{\mu}^{EB} - \widehat{\mu}^{B})\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\boldsymbol{\Omega}}\widehat{\boldsymbol{\theta}}^{B}\Big],\\
I_{32} =& E\Big[(\widehat{\mu}^{EB} - \widehat{\mu}^{B})\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})\Big].
\end{aligned}
\tag{A.17}
$$

Since $\widehat{\mu}^{EB}-\widehat{\mu}^{B}=(\boldsymbol{c}-\boldsymbol{s})'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})+\sum_{a=1}^{q}(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})'\boldsymbol{s}_{(a)}\widehat{\psi}_a^{\dagger}$ from (A.9), it is seen that

$$\begin{aligned}\overline{I}_{31}=&\frac{1}{2}\big[\{A(\boldsymbol{\omega})\}^2-1\big]\Big\{\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\Omega}\boldsymbol{X}\boldsymbol{\beta}\sum_{a=1}^{q}E[\boldsymbol{s}_{(a)}'\boldsymbol{\Sigma}\boldsymbol{\nabla}_y\widehat{\psi}_a^{\dagger}]\\&+(\boldsymbol{c}-\boldsymbol{s})'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{P}_{\Omega}\boldsymbol{\Omega}^{-1}\boldsymbol{c}\\&+\sum_{a=1}^{q}\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_{\Omega}\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\big\{\boldsymbol{s}_{(a)}E[\widehat{\psi}_a^{\dagger}]+E[\boldsymbol{\nabla}_y\boldsymbol{\nabla}_y'\widehat{\psi}_a^{\dagger}]\boldsymbol{\Sigma}\boldsymbol{s}_{(a)}\big\}\Big\},\end{aligned}\tag{A.18}$$

$$\begin{aligned}\overline{I}_{32}=&(\boldsymbol{c}-\boldsymbol{s})'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'E[\boldsymbol{\nabla}_y\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})]\\&+\sum_{a=1}^{q}E[\boldsymbol{s}_{(a)}'\boldsymbol{\Sigma}\{\boldsymbol{\nabla}_y\widehat{\psi}_a^{\dagger}\}\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})]+\sum_{a=1}^{q}E[\widehat{\psi}_a^{\dagger}\boldsymbol{s}_{(a)}'\boldsymbol{\nabla}_y\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})].\end{aligned}\tag{A.19}$$

From the condition (A4), it follows that $E[\boldsymbol{s}_{(a)}'\boldsymbol{\Sigma}\boldsymbol{\nabla}_y\widehat{\psi}_a^{\dagger}]=O(N^{-1})$, $E[\widehat{\psi}_a^{\dagger}]=O(N^{-1})$, $\boldsymbol{c}'E[\boldsymbol{\nabla}_y\boldsymbol{\nabla}_y'\widehat{\psi}_a^{\dagger}]\boldsymbol{\Sigma}\boldsymbol{s}_{(a)}=O(N^{-1})$ and $\boldsymbol{s}_{(a)}'\boldsymbol{\Sigma}\{\boldsymbol{\nabla}_y\widehat{\psi}_a^{\dagger}\}=O_p(N^{-1})$. Thus, $I_3$ is approximated as $I_3=\overline{I}_3(\boldsymbol{\omega})+O(N^{-3/2})$ for $\overline{I}_3(\boldsymbol{\omega})$ given in (3.12). Therefore, the proof is complete. ■

**[Scenario 3] Case of** $r=1$**.** In this case, from Theorem A.2, it follows that $I_2=E[\{\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})\}^2]+O(N^{-3/2})$ and $I_3=E[(\widehat{\mu}^{EB}-\widehat{\mu}^{B})\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})]=\overline{I}_{32}+O(N^{-3/2})$. Thus, we get the following theorem.

**Theorem A.3** *Assume the conditions* (A1)-(A6) *and* $r=1$. *Then, MSE of* $\widehat{\mu}^{CEB}$ *is approximated as*

$$MSE(\boldsymbol{\omega},\widehat{\mu}^{CEB})=MSE(\boldsymbol{\psi},\widehat{\mu}^{EB})+E[\{\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})\}^2]+2\overline{I}_{32}(\boldsymbol{\omega})+O(N^{-3/2}),\tag{A.20}$$

*for* $\overline{I}_{32}$ *given in* (3.17).

Consider the case that $\boldsymbol{m}(\boldsymbol{y})=\boldsymbol{W}'\boldsymbol{\Omega}(\boldsymbol{y}-\widehat{\boldsymbol{\theta}}^{EB})=\boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{R}}\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{y}-\boldsymbol{X}\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\psi}}))$. Then, $\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})$ is approximated as

$$\begin{aligned}\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})=&\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{R}\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})+\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{W}'\boldsymbol{\Omega}(\widehat{\boldsymbol{R}}\widehat{\boldsymbol{\Sigma}}^{-1}-\boldsymbol{R}\boldsymbol{\Sigma}^{-1})(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})\\&-\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{W}'\boldsymbol{\Omega}\widehat{\boldsymbol{R}}\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}(\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\psi}})-\boldsymbol{\beta})\\=&\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{R}\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})+O_p(N^{-1}).\end{aligned}$$

This implies that $E[\{\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})\}^2]=\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{R}\boldsymbol{\Sigma}^{-1}\boldsymbol{R}\boldsymbol{\Omega}\boldsymbol{W}\boldsymbol{c}_{\boldsymbol{w}}+O(N^{-3/2})$. Also,

$$\begin{aligned}(\boldsymbol{c}-\boldsymbol{s})'&\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'E[\boldsymbol{\nabla}_{\boldsymbol{y}}\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})]\\&=(\boldsymbol{c}-\boldsymbol{s})'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{R}\boldsymbol{\Omega}\boldsymbol{W}\boldsymbol{c}_{\boldsymbol{w}}+O(N^{-3/2}),\\E[\widehat{\psi}_a^{\dagger}\boldsymbol{s}_{(a)}'\boldsymbol{\nabla}_{\boldsymbol{y}}\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})]&=E[\widehat{\psi}_a^{\dagger}]\boldsymbol{s}_{(a)}'\boldsymbol{\Sigma}^{-1}\boldsymbol{R}\boldsymbol{\Omega}\boldsymbol{W}\boldsymbol{c}_{\boldsymbol{w}}+O(N^{-3/2}).\end{aligned}$$

Since $E[\widehat{\psi}_a^{\dagger}]=O(N^{-1})$ and $\boldsymbol{s}_{(a)}'\boldsymbol{\Sigma}^{-1}\boldsymbol{R}\boldsymbol{\Omega}\boldsymbol{W}\boldsymbol{c}_{\boldsymbol{w}}=O(N^{-1})$, it is seen that

$$E[\widehat{\psi}_a^{\dagger}\boldsymbol{s}_{(a)}'\boldsymbol{\nabla}_{\boldsymbol{y}}\boldsymbol{c}_{\boldsymbol{w}}'\boldsymbol{m}(\boldsymbol{y})]=O(N^{-3/2}).$$

Hence, we get the following proposition.

**Proposition A.1** *Assume the conditions* (A1)-(A6) *and* $r = 1$*. In the case that* $\boldsymbol{m}(\boldsymbol{y}) = \boldsymbol{W}'\boldsymbol{\Omega}(\boldsymbol{y} - \widehat{\boldsymbol{\theta}}^{EB})$*, MSE of* $\widehat{\mu}^{CEB}$ *is approximated as*

$$\begin{aligned} MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB}) =& MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB}) + \boldsymbol{c}_w'\boldsymbol{W}'\boldsymbol{\Omega}\boldsymbol{R}\boldsymbol{\Sigma}^{-1}\boldsymbol{R}\boldsymbol{\Omega}\boldsymbol{W}\boldsymbol{c_w} \\ &+ 2(\boldsymbol{c} - \boldsymbol{s})'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{R}\boldsymbol{\Omega}\boldsymbol{W}\boldsymbol{c_w} + O(N^{-3/2}). \end{aligned} \quad \text{(A.21)}$$

## A.3 Estimation of MSE

We next provide a second-order unbiased estimator of the $MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB})$. It is noted that $I_3 = O(N^{-1})$ from Lemma A.2. This implies that $E[I_3^*] = I_3 + O(N^{-3/2})$ for $I_3^*$ given in (3.26). Also, note that $\widehat{I}_2$ given in (3.29) is an unbiased estimator of $I_2$. Hence, we get the following theorem.

**Theorem A.4** *Assume the conditions* (A1)-(A6)*. Let* $mse(\widehat{\mu}^{EB})$ *be a second-order unbiased estimator of* $MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB})$*, namely,* $E[mse(\widehat{\mu}^{EB})] = MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB}) + O(N^{-3/2})$*. Then, the estimator given by*

$$mse^*(\widehat{\mu}^{CEB}) = mse(\widehat{\mu}^{EB}) + \widehat{I}_2 + 2I_3^* \quad \text{(A.22)}$$

*satisfies that* $E[mse^*(\widehat{\mu}^{CEB})] = MSE(\boldsymbol{\psi}, \widehat{\mu}^{CEB}) + O(N^{-3/2})$*.*

This theorem holds for all $r \geq 0$. In the case of $r = 1/2$, we can obtain another second-order unbiased estimator based on the second-order approximation of MSE. It is noted that $I_{23}$ given in (A.16) is expressed as

$$\begin{aligned} I_{23} =& 2^{-1}\big[\{A(\boldsymbol{\omega})\}^2 - 1\big]\big\{\boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_\Omega\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{c}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}_\Omega\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}E\big[(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})\}\boldsymbol{c}_w'\boldsymbol{m}(\boldsymbol{y})\big]\big\} \\ &+ O(N^{-3/2}). \end{aligned}$$

Since $E\big[(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})\}\boldsymbol{c}_w'\boldsymbol{m}(\boldsymbol{y})\big]$ can be estimated based on the parametric bootstrap method as

$$E_*\big[(\boldsymbol{y}^* - \boldsymbol{X}\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\psi}}))\}\boldsymbol{c}_w'\boldsymbol{m}(\boldsymbol{y}^*)|\boldsymbol{y}\big].$$

Using the same argument, we can estimate $\overline{I}_2(\boldsymbol{\psi})$ and $\overline{I}_3(\boldsymbol{\psi})$ by $\overline{I}_2^*$ and $\overline{I}_3^*$, respectively, given in (3.32) and (3.33).

**Theorem A.5** *Assume the conditions* (A1)-(A6) *and* $r = 1/2$*. Let* $mse(\widehat{\mu}^{EB})$ *be a second-order unbiased estimator of* $MSE(\boldsymbol{\psi}, \widehat{\mu}^{EB})$*. Define* $mse^{**}(\widehat{\mu}^{CEB})$ *by*

$$mse^{**}(\widehat{\mu}^{CEB}) = mse(\widehat{\mu}^{EB}) + \overline{I}_2^* + 2\overline{I}_3^*.$$

*Then,* $mse^{**}(\widehat{\mu}^{CEB})$ *is a second-order unbiased estimator of* $MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB})$*, namely,*

$$E[mse^{**}(\widehat{\mu}^{CEB})] = MSE(\boldsymbol{\omega}, \widehat{\mu}^{CEB}) + O(N^{-3/2}).$$