CIRJE-F-632

Corrected Empirical Bayes Confidence Intervals in Nested Error Regression Models

Tatsuya Kubokawa
University of Tokyo

August 2009; Revised in June 2009

Corrected Empirical Bayes Confidence Intervals in Nested Error Regression Models

Tatsuya Kubokawa*
University of Tokyo

June 24, 2009

Abstract

In the small area estimation, the empirical best linear unbiased predictor (EBLUP) or the empirical Bayes estimator (EB) in the linear mixed model is recognized useful because it gives a stable and reliable estimate for a mean of a small area. In practical situations where EBLUP is applied to real data, it is important to evaluate how much EBLUP is reliable. One method for the purpose is to construct a confidence interval based on EBLUP. In this paper, we obtain an asymptotically corrected empirical Bayes confidence interval in a nested error regression model with unbalanced sample sizes and unknown components of variance. The coverage probability is shown to satisfy the confidence level in the second order asymptotics. It is numerically revealed that the corrected confidence interval is superior to the conventional confidence interval based on the sample mean in terms of the coverage probability and the expected width of the interval. Finally, it is applied to the posted land price data in Tokyo and the neighboring prefecture.

Key words and phrases: Best linear unbiased predictor, confidence interval, empirical Bayes procedure, finite population, linear mixed model, nested error regression model, second order correction, small area estimation.

1 Introduction

The empirical best linear unbiased predictor (EBLUP) or the empirical Bayes estimator (EB) in linear mixed models have been recognized as useful tools in small area estimation. In small area estimation, sample means may have unacceptable estimation errors because sample sizes of small areas are small. EBLUP is an alternative method to provide stable estimates with higher precisions by borrowing data in the surrounding areas. In practical situations where EBLUP is applied to real data, it is important to evaluate the estimation errors of EBLUP for each small area. One method is to provide estimates of the mean

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

squared errors of EBLUP, and it has been studied enough in the literature including Prasad and Rao (1990), Harville and Jeske (1992), Booth and Hobert (1998), Datta and Lahiri (2000), Rivest and Belmonte (2000), Das, Jiang and Rao (2004), Datta, Rao and Smith (2005) and others. Another method is to provide the confidence intervals based on EBLUP, and the two approaches to this issue have been studied. One is the method based on parametric bootstrap proposed by Hall and Maiti (2007) and Chatterjee, Lahiri and Li (2008), and the other is the method based on the Taylor series expansion. Although the confidence intervals based on the parametric bootstrap can be applied to the general linear mixed models, they are hard to implement. In contrast, the methods based on the Taylor series expansion are easy to implement, but the derivation depends on individual models. Since we specify a nested error regression model and the extended finite population model in this paper, we want to develop closed form confidence intervals based on the Taylor series expansion. This method has been used by Datta, Ghosh, Smith and Lahiri (2002) and Basu, Ghosh and Mukerjee (2003), who derived the asymptotically corrected empirical Bayes confidence interval in the Fay-Herriot model with a known error variance. The Fay-Herriot model is categorized into basic area level models where only aggregated data such as sample means are observed. When individual data are available, we can use basic unit level models to carry out more precise inference for small areas. A simple, but useful basic unit level model is a nested error regression model with unbalanced sample sizes and unknown components of variance. In fact, this model has been extensively used in the literature concerning the small area estimation as illustrated in Battese, Harter and Fuller (1988) and Rao (2003). In this paper, we shall construct corrected empirical Bayes confidence intervals based on EBLUP in the nested error regression model and the extended finite population model.

To explain more details, consider the nested error regression model (NERM) given by

$$y_{ij} = \boldsymbol{x}_{ij}'\boldsymbol{\beta} + v_i + e_{ij}, \quad i = 1, \ldots, k,\ j = 1, \ldots, n_i, \tag{1.1}$$

where k is the number of small areas, $\boldsymbol{x}_{ij}$ 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 e_{ij}'s are mutually independently distributed as $v_i \sim \mathcal{N}(0, \sigma_v^2)$ and $e_{ij} \sim \mathcal{N}(0, \sigma^2)$. Here, σ_v^2 and σ^2 are referred to as, respectively, 'between' and 'within' components of variance, and both are unknown. We want to construct a confidence interval of the mean $\mu_i = \overline{\boldsymbol{x}}_i'\boldsymbol{\beta} + v_i$ of the i-th small area for $\overline{\boldsymbol{x}}_i = \sum_{j=1}^{n_i} \boldsymbol{x}_{ij}/n_i$. Since the conditional distribution of the sample mean $\overline{y}_i$ given v_i is $\mathcal{N}(\mu_i, \sigma^2/n_i)$, a conventional confidence interval based on the sample mean $\overline{y}_i$ is

$$I_i^T \ :\ \overline{y}_i \pm t_{\alpha/2}\sqrt{\tilde{\sigma}_e^2/n_i}, \tag{1.2}$$

where $\tilde{\sigma}_e^2$ is an available unbiased estimator of σ^2 which is independent of $(\overline{y}_1, \ldots, \overline{y}_k)$, and $t_{\alpha/2}$ is the $\alpha/2$ upper quantile of a t-distribution with appropriate degrees of freedom. Although the coverage probability of I_i^T is exactly identical to the confidence coefficient $1 - \alpha$, the width of the confidence interval I_i^T is longer for smaller n_i since $\overline{y}_i$ has an unacceptable estimation error. Thus, we construct a confidence interval based on EBLUP or EB using the linear mixed model in (1.1).

Let ψ be the ratio of the variance components, namely, $\psi = \sigma_v^2/\sigma^2$, and let $\gamma_i =$

$\gamma_i(\psi) = 1/(1 + n_i\psi)$. Note that given $\overline{y}_i$, μ_i has conditionally

$$\mu_i|\overline{y}_i \sim \mathcal{N}\left(\widehat{\mu}_i^B(\boldsymbol{\beta}, \psi), (\sigma^2/n_i)(1-\gamma_i)\right) \tag{1.3}$$

where $\widehat{\mu}_i^B(\boldsymbol{\beta}, \psi)$ is the conditional mean $E[\mu_i|\overline{y}_i]$ given by

$$\widehat{\mu}_i^B(\boldsymbol{\beta}, \psi) = \overline{\boldsymbol{x}}_i'\boldsymbol{\beta} + (1-\gamma_i)(\overline{y}_i - \overline{\boldsymbol{x}}_i'\boldsymbol{\beta}). \tag{1.4}$$

This is also interpreted as the Bayes estimator, since the model (1.1) can be viewed as a Bayesian model. Then, the confidence interval of μ_i with respect to this conditional distribution is given by

$$I_i^B(\boldsymbol{\beta}, \psi, \sigma^2): \ \widehat{\mu}_i^B(\boldsymbol{\beta}, \psi) \pm z_{\alpha/2}\sqrt{(\sigma^2/n_i)(1-\gamma_i)}, \tag{1.5}$$

where $z_{\alpha/2}$ is the $\alpha/2$ upper quantile of a standard normal distribution. It is noted that the conditional coverage probability satisfies $P[\mu_i \in I_i^B(\boldsymbol{\beta}, \psi)|\overline{y}_i] = 1-\alpha$, which leads to $P[\mu_i \in I_i^B(\boldsymbol{\beta}, \psi)] = E[P[\mu_i \in I_i^B(\boldsymbol{\beta}, \psi)|\overline{y}_i]] = 1-\alpha$, namely, the unconditional coverage probability satisfies the confidence level. Since $\boldsymbol{\beta}$, ψ and σ^2 are unknown, we need to estimate these parameters. For known ψ, the generalized least squares estimator of $\boldsymbol{\beta}$ is given by Rao (2003) as

$$\widehat{\boldsymbol{\beta}}(\psi) = \big(\boldsymbol{A}(\psi) + \boldsymbol{B}\big)^{-1}\Big(\sum_{i=1}^k n_i\gamma_i(\psi)\overline{\boldsymbol{x}}_i\overline{y}_i + \sum_{i=1}^k\sum_{j=1}^{n_i}(\boldsymbol{x}_{ij} - \overline{\boldsymbol{x}}_i)(y_{ij} - \overline{y}_i)\Big), \tag{1.6}$$

where $\boldsymbol{A}(\psi) = \sum_{i=1}^k n_i\gamma_i(\psi)\overline{\boldsymbol{x}}_i\overline{\boldsymbol{x}}_i'$ and $\boldsymbol{B} = \sum_{i=1}^k\sum_{j=1}^{n_i}(\boldsymbol{x}_{ij} - \overline{\boldsymbol{x}}_i)(\boldsymbol{x}_{ij} - \overline{\boldsymbol{x}}_i)'$. When consistent estimators of σ^2 and ψ, denoted by $\hat{\sigma}^2$ and $\widehat{\psi}$, are available, the estimators $\widehat{\boldsymbol{\beta}}(\widehat{\psi})$, $\hat{\sigma}^2$ and $\widehat{\psi}$ are substituted into (1.5) to get the empirical Bayes confidence interval

$$I_i^{EB}(\widehat{\psi}, \hat{\sigma}^2): \ \widehat{\mu}_i^{EB}(\widehat{\psi}) \pm z_{\alpha/2}\sqrt{(\hat{\sigma}^2/n_i)(1-\hat{\gamma}_i)}, \tag{1.7}$$

where $\widehat{\mu}_i^{EB}(\widehat{\psi})$ is the empirical Bayes estimator (EB) given by

$$\widehat{\mu}_i^{EB}(\widehat{\psi}) = \widehat{\mu}_i^B(\widehat{\boldsymbol{\beta}}(\widehat{\psi}), \widehat{\psi}) = \overline{\boldsymbol{x}}_i'\widehat{\boldsymbol{\beta}}(\widehat{\psi}) + (1-\hat{\gamma}_i)(\overline{y}_i - \overline{\boldsymbol{x}}_i'\widehat{\boldsymbol{\beta}}(\widehat{\psi})),$$

for

$$\hat{\gamma}_i = \gamma_i(\widehat{\psi}) = 1/(1 + n_i\widehat{\psi}).$$

The empirical Bayes estimator $\widehat{\mu}_i^{EB}(\widehat{\psi})$ is known as the empirical best linear unbiased predictor (EBLUP) in the context of the linear mixed model. For smaller $n_i\widehat{\psi}$, the EB estimator $\widehat{\mu}_i^{EB}(\widehat{\psi})$ shrinks $\overline{y}_i$ more towards $\overline{\boldsymbol{x}}_i'\widehat{\boldsymbol{\beta}}(\widehat{\psi})$, which results in a stable estimate with a higher precision.

Although $I_i^{EB}(\widehat{\psi}, \hat{\sigma}^2)$ gives a stable confidence interval for small n_i, it has a drawback that the coverage probability $P[\mu_i \in I_i^{EB}(\widehat{\psi}, \hat{\sigma}^2)]$ cannot be guaranteed to be greater than or equal to the nominal confidence coefficient $1-\alpha$. As seen from the simulation experiment given in Section 3.1, it seems that $P[\mu_i \in I_i^{EB}(\widehat{\psi}, \hat{\sigma}^2)] < 1-\alpha$ for some ψ. A method for fixing this shortcoming is to adjust the significance quantile $z_{\alpha/2}$ as

$z_{\alpha/2}\{1+(2k)^{-1}h_i(\widehat{\psi})\}$ with an appropriate correction function $h_i(\widehat{\psi})$. That is, the corrected confidence interval is described as

$$I_i^{CEB}(\widehat{\psi},\hat{\sigma}^2):\ \widehat{\mu}_i^{EB}(\widehat{\psi})\pm z_{\alpha/2}\left[1+(2k)^{-1}h_i(\widehat{\psi})\right]\sqrt{(\hat{\sigma}^2/n_i)(1-\hat{\gamma}_i)}. \qquad (1.8)$$

This method was used in Datta *et al*. (2002) and Basu *et al*. (2003) for the Fay-Herriot model.

Another approch to constructing stable confidence intervals is to use the estimator of the mean squared error (MSE) of $\widehat{\mu}_i^{EB}(\widehat{\psi})$ instead of $(\hat{\sigma}^2/n_i)(1-\hat{\gamma}_i)$. Let $mse_i(\hat{\sigma}^2,\widehat{\psi})$ be the second-order unbiased estimator of the MSE given by $MSE_i(\sigma^2,\psi)=E[(\widehat{\mu}_i^{EB}(\widehat{\psi})-\mu_i)^2]$, namely, $E[mse_i(\hat{\sigma}^2,\widehat{\psi})]=MSE_i(\sigma^2,\psi)+O(k^{-3/2})$. Following Morris (1983), Prasad and Rao (1990) proposed another type of the empirical Bayes confidence interval

$$I_i^{EB*}(\widehat{\psi},\hat{\sigma}^2):\ \widehat{\mu}_i^{EB}(\widehat{\psi})\pm z_{\alpha/2}\sqrt{mse_i(\hat{\sigma}^2,\widehat{\psi})}. \qquad (1.9)$$

In this paper, we can modify $I_i^{EB*}(\widehat{\psi},\hat{\sigma}^2)$ to provide the corrected confidence interval

$$I_i^{CEB*}(\widehat{\psi},\hat{\sigma}^2):\ \widehat{\mu}_i^{EB}(\widehat{\psi})\pm z_{\alpha/2}\left[1+(2k)^{-1}h_i^*(\widehat{\psi})\right]\sqrt{mse_i(\hat{\sigma}^2,\widehat{\psi})}, \qquad (1.10)$$

for an appropriate correction function $h_i^*(\widehat{\psi})$.

In Section 2, we obtain the functions $h_i(\widehat{\psi})$ and $h_i^*(\widehat{\psi})$ such that the coverage probabilities satisfy the nominal confidence coefficient in the second order for large k, namely,

$$\begin{aligned} P[\mu_i\in I_i^{CEB}(\widehat{\psi},\hat{\sigma}^2)] &=1-\alpha+O(k^{-3/2}),\\ P[\mu_i\in I_i^{CEB*}(\widehat{\psi},\hat{\sigma}^2)] &=1-\alpha+O(k^{-3/2}), \end{aligned} \qquad (1.11)$$

as $k\to\infty$. Since the sample sizes n_i's are bounded in small area problems, it is common to consider the setup of k going to infinity. In the sense of (1.11), we call $I_i^{CEB}(\widehat{\psi},\hat{\sigma}^2)$ and $I_i^{CEB*}(\widehat{\psi},\hat{\sigma}^2)$ the corrected empirical Bayes confidence intervals. In Section 2, the corrected empirical Bayes confidence intervals are not only derived in the NERM, but also extended to a finite population model. The numerical performance of the corrected confidence intervals is investigated in Section 3 and it is revealed that $I_i^{CEB}(\widehat{\psi},\hat{\sigma}^2)$ and $I_i^{CEB*}(\widehat{\psi},\hat{\sigma}^2)$ are superior to the conventional confidence interval I_i^T in terms of both the coverage probability and the expected width of the interval. The proposed corrected confidence interval is applied to a real data set of the posted land price data in Tokyo and the neighboring prefecture, and it is shown to be useful. Concluding remarks are given in Section 4, and the proofs are given in the final section.

2 Asymptotically Corrected Confidence Intervals

2.1 Nested error regression model (NERM)

We here provide the asymptotically corrected empirical Bayes confidence interval under the NERM in (1.1). This model is expressed in vector notations as

$$\boldsymbol{y}_i=\boldsymbol{X}_i\boldsymbol{\beta}+\boldsymbol{j}_{n_i}v_i+\boldsymbol{e}_i, \quad \text{for} \quad i=1,\dots,k, \qquad (2.1)$$

where $\boldsymbol{X}_i = (\boldsymbol{x}_{i1}, \ldots, \boldsymbol{x}_{i,n_i})'$ and n_i-vectors $\boldsymbol{y}_i = (y_{i1}, \ldots, y_{i,n_i})'$, $\boldsymbol{e}_i = (e_{i1}, \ldots, e_{i,n_i})'$ and $\boldsymbol{j}_{n_i} = (1, \ldots, 1)'$. Then the covariance matrix of $\boldsymbol{y}_i$ is $\mathbf{Cov}\,(\boldsymbol{y}_i) = \sigma^2 \boldsymbol{V}_i(\psi)$ for $\boldsymbol{V}_i(\psi) = \boldsymbol{I}_{n_i} + \psi \boldsymbol{J}_{n_i}$, where $\boldsymbol{I}_{n_i}$ is the $n_i \times n_i$ identity matrix, and $\boldsymbol{J}_{n_i} = \boldsymbol{j}_{n_i}\boldsymbol{j}_{n_i}'$ is the $n_i \times n_i$ matrix with every elements being one. Letting $\boldsymbol{y} = (\boldsymbol{y}_1', \ldots, \boldsymbol{y}_k')'$, $\boldsymbol{X} = (\boldsymbol{X}_1', \ldots, \boldsymbol{X}_k')'$, $\boldsymbol{v} = (v_1, \ldots, v_k)'$ and $\boldsymbol{e} = (\boldsymbol{e}_1', \ldots, \boldsymbol{e}_k')'$, we can rewrite the model (2.1) as

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

where $\boldsymbol{Z} = \text{block diag}(\boldsymbol{j}_{n_1}, \ldots, \boldsymbol{j}_{n_k})$, the block diagonal matrix. The covariance matrix of $\boldsymbol{y}$ is $\mathbf{Cov}\,(\boldsymbol{y}) = \sigma^2 \boldsymbol{V}(\psi)$ for $\boldsymbol{V} = \text{block diag}(\boldsymbol{V}_1(\psi), \ldots, \boldsymbol{V}_k(\psi))$. We shall use those vector notations through the paper for the convenience.

As stated in Section 1, the corrected empirical Bayes confidence interval of the i-th small area mean $\mu_i = \overline{\boldsymbol{x}}_i'\boldsymbol{\beta} + v_i$ is given by

$$I_i^{CEB}(\widehat{\psi}, \hat{\sigma}^2):\ \widehat{\mu}_i^{EB}(\widehat{\psi}) \pm z_{\alpha/2}\left[1 + (2k)^{-1} h_i(\widehat{\psi})\right]\sqrt{(\hat{\sigma}^2/n_i)(1 - \hat{\gamma}_i)}, \tag{2.2}$$

where the correction function $h_i(\widehat{\psi})$ is adjusted so that the coverage probability can satisfy the nominal confidence level in the second order for large k. To this end, we assume the following conditions:

(A1) The elements of $\boldsymbol{X}$ are uniformly bounded, and $\boldsymbol{X}'\boldsymbol{V}(\psi)^{-1}\boldsymbol{X}/k$ converges to a positive definite matrix as $k \to \infty$;

(A2) n_i's are bounded for $i = 1, \ldots, k$;

(A3) $\hat{\sigma}^2$ is an estimator of σ^2 which satisfies that $\hat{\sigma}^2 - \sigma^2 = O_p(k^{-1/2})$ and $Bias_\psi(\hat{\sigma}^2) = O(k^{-1})$ as $k \to \infty$.

(A4) $\widehat{\psi}$ is an estimator of ψ which satisfies that $\widehat{\psi} - \psi = O_p(k^{-1/2})$, $Bias_\psi(\widehat{\psi}) = O(k^{-1})$ and $\partial\widehat{\psi}/\partial\overline{y}_i = O(k^{-1})$ as $k \to \infty$.

Then, we can get the main theorem which will be proved in Section 5:

Theorem 2.1 *Assume the conditions* (A1)-(A4). *Define the correction function* $h_i(\psi)$ *by*

$$\begin{aligned} h_i(\psi) =& k\frac{\gamma_i}{\psi}\overline{\boldsymbol{x}}_i'(\boldsymbol{A}(\psi) + \boldsymbol{B})^{-1}\overline{\boldsymbol{x}}_i + 2k\frac{\gamma_i(1-\gamma_i)}{\psi^2}Var(\widehat{\psi}) \\ & - kE[H_{0,i}] + \frac{k}{4}E[z_{\alpha/2}^2 H_{0,i}^2 + K_{0,i}^2], \end{aligned} \tag{2.3}$$

where $Var(\widehat{\psi}) = E[(\widehat{\psi} - \psi)^2]$, $H_{0,i} = \gamma_i(\widehat{\psi} - \psi)/\psi + (\hat{\sigma}^2 - \sigma^2)/\sigma^2$, $K_{0,i} = \gamma_i(\widehat{\psi} - \psi)/\psi - (\hat{\sigma}^2 - \sigma^2)/\sigma^2$. *Then, the corrected empirical Bayes confidence interval given in* (2.2) *satisfies that* $P[\mu_i \in I_i^{CEB}(\widehat{\psi}, \hat{\sigma}^2)] = 1 - \alpha + O(k^{-3/2})$ *as* $k \to \infty$.

The function $h_i(\psi)$ can be obtained for specific estimators $\hat{\sigma}^2$ and $\widehat{\psi}$ by evaluating the biases $Bias(\hat{\sigma}^2)$, $Bias(\widehat{\psi})$, the variances $Var(\hat{\sigma}^2)$, $Var(\widehat{\psi})$ and the covariance $Cov(\hat{\sigma}^2, \widehat{\psi})$ asymptotically. Since $\psi = \sigma_v^2/\sigma^2$ is the ratio of the variance components, it may be convenient in some cases to express the moments $E[(\widehat{\psi} - \psi)^2]$, $E[H_{0,i}]$, $E[H_{0,i}^2]$ and $E[K_{0,i}^2]$ based on the estimators of the variance components $\hat{\sigma}^2$ and $\hat{\sigma}_v^2$, where $\hat{\sigma}_v^2 = \hat{\sigma}^2\widehat{\psi}$.

Proposition 2.1 *Let* $\hat{\sigma}_v^2 = \hat{\sigma}^2\widehat{\psi}$ *and assume the conditions* (A3) *and* (A4)*. Then, the correction function* $h_i(\psi)$ *given in* (2.3) *can be expressed as*

$$\begin{aligned} h_i(\psi) =& k\frac{\gamma_i}{\psi}\overline{\boldsymbol{x}}_i'(\boldsymbol{A}(\psi)+\boldsymbol{B})^{-1}\overline{\boldsymbol{x}}_i + 2k\gamma_i(1-\gamma_i)\tau_1(\psi) - k\{(1-\gamma_i)b_e(\psi)+\gamma_i b_v(\psi)\} \\ &+ \frac{k}{4}(z_{\alpha/2}^2+1)\left\{\gamma_i^2\tau_1(\psi) - 2\gamma_i\tau_2(\psi) + \tau_3(\psi)\right\} + O(k^{-3/2}), \end{aligned} \quad (2.4)$$

where for $T_e = (\hat{\sigma}^2-\sigma^2)/\sigma^2$ *and* $T_v = (\hat{\sigma}_v^2-\sigma_v^2)/\sigma_v^2$*,*

$$\begin{aligned} b_e(\psi) =& E[T_e], \quad b_v(\psi) = E[T_v], \\ \tau_1(\psi) =& E[(T_e-T_v)^2],\ \tau_2(\psi) = E[T_e(T_e-T_v)],\ \tau_3(\psi) = E[T_e^2]. \end{aligned} \quad (2.5)$$

It is noted that the function $h_i(\psi)$ given in (2.4) includes the value of ψ in the denominator at the first term in the r.h.s. of equation (2.4). Since the confidence interval is constructed using the estimator $\widehat{\psi}$, this causes the instability when $\widehat{\psi}$ is close to zero. Thus, it is reasonable to use the truncated estimator of $\widehat{\psi}$. An example of the truncation is given by

$$[\widehat{\psi}]^{TR} = \max\{\widehat{\psi}, k^{-a}\}, \quad (2.6)$$

for a positive constant a. The estimators $[\widehat{\psi}]^{TR}$ and $\widehat{\psi}$ can be shown to be asymptotically equivalent.

Proposition 2.2 *Assume that* $\psi > 0$ *and* $\widehat{\psi} \to \psi$ *in probability as* $k \to \infty$*. Let* a *and* b *be positive constants. Then,* $[\widehat{\psi}]^{TR} = \max\{\widehat{\psi}, k^{-a}\} = \widehat{\psi} + o_p(k^{-b})$ *as* $k \to \infty$*.*

Proof. For any $\varepsilon > 0$, we shall show that $P[k^b|\max(\widehat{\psi}, k^{-a}) - \widehat{\psi}| > \varepsilon] \to 0$ as $k \to \infty$. It is observed that $P[k^b|\max(\widehat{\psi}, k^{-a}) - \widehat{\psi}| > \varepsilon] = P[\widehat{\psi} < k^{-a}, \widehat{\psi} < k^{-a} - \varepsilon k^{-b}] \le P[\widehat{\psi} < k^{-a}]$. It is noted that $\widehat{\psi} \to \psi > 0$ while $k^{-a} \to 0$, which implies that $P[\widehat{\psi} < k^{-a}] \to 0$. ■

We next consider the empirical Bayes confidence intervals given in (1.9) and (1.10) using an estimator of the mean squared error (MSE) of $\widehat{\mu}_i^{EB}(\widehat{\psi})$. To this end, we begin with deriving the unbiased estimator of the MSE, which can be shown by using the same arguments as in Prasad and Rao (1990) and Datta and Lahiri (2000) under the following condition:

(A5) The estimatr $\widehat{\psi} = \widehat{\psi}(\boldsymbol{y})$ satisfies (i) $\widehat{\psi}(-\boldsymbol{y}) = \widehat{\psi}(\boldsymbol{y})$ and (ii) $\widehat{\psi}(\boldsymbol{y}+\boldsymbol{X\alpha}) = \widehat{\psi}(\boldsymbol{y})$ for any p-dimensional vector $\boldsymbol{\alpha}$.

Proposition 2.3 *Assume the conditions* (A1)*-*(A5) *and use the notations* $\tau_1(\psi)$, $b_e(\psi)$ *and* $b_v(\psi)$ *defined in* (2.5)*. Then the second-order approximation of the MSE of the empirical Bayes estimator* $\widehat{\mu}_i^{EB}(\widehat{\psi})$ *is given by*

$$\begin{aligned} MSE_i(\sigma^2,\psi) =& E[(\widehat{\mu}_i^{EB}(\widehat{\psi}) - \mu_i)^2] \\ =& \frac{\sigma^2}{n_i}(1-\gamma_i) + \sigma^2\gamma_i^2\overline{\boldsymbol{x}}_i'(\boldsymbol{A}(\psi)+\boldsymbol{B})^{-1}\overline{\boldsymbol{x}}_i + \sigma^2 n_i\gamma_i^3\psi^2\tau_1(\psi) + O(k^{-3/2}), \end{aligned} \quad (2.7)$$

and the second-order unbiased estimator of the MSE is given by

$$
\begin{aligned}
mse_i(\hat\sigma^2,\widehat\psi) =& \frac{\hat\sigma^2}{n_i}(1-\hat\gamma_i) + \hat\sigma^2\gamma_i^2\overline{\boldsymbol{x}}_i'(\boldsymbol{A}(\widehat\psi)+\boldsymbol{B})^{-1}\overline{\boldsymbol{x}}_i \\
&- \hat\sigma^2\widehat\psi\hat\gamma_i\{(1-\hat\gamma_i)b_e(\widehat\psi)+\hat\gamma_i b_v(\widehat\psi)\} + 2\hat\sigma^2 n_i\hat\gamma_i^3\widehat\psi^2\tau_1(\widehat\psi),
\end{aligned} \tag{2.8}
$$

namely, $E[mse_i(\hat\sigma^2,\widehat\psi)] = MSE_i(\sigma^2,\psi)+O(k^{-3/2})$.

This proposition was derived by Prasad and Rao (1990) for the estimators given in Example 2.1 and by Datta and Lahiri (2000) for the maximum likelihood and restricted maximum likelihood estimators given in Example 2.2. These results can be unified by Proposition 2.3 by assuming $\partial\widehat\psi/\partial\overline{y}_i = O_p(k^{-1})$.

The empirical Bayes confidence interval suggested by Prasad and Rao (1990) is given by

$$
I_i^{EB*}(\widehat\psi,\hat\sigma^2):\ \widehat\mu_i^{EB}(\widehat\psi) \pm z_{\alpha/2}\sqrt{mse_i(\hat\sigma^2,\widehat\psi)}, \tag{2.9}
$$

and the corrected confidence interval is given by

$$
I_i^{CEB*}(\widehat\psi,\hat\sigma^2):\ \widehat\mu_i^{EB}(\widehat\psi) \pm z_{\alpha/2}\left[1+(2k)^{-1}h_i^*(\widehat\psi)\right]\sqrt{mse_i(\hat\sigma^2,\widehat\psi)}, \tag{2.10}
$$

where for $\ell_i(\widehat\psi) = k\{mse_i(\hat\sigma^2,\widehat\psi)n_i/\{\hat\sigma^2(1-\hat\gamma_i)\}-1\}$ and $h_i(\psi)$ given in (2.3), $h_i^*(\widehat\psi)$ is defined by

$$
\begin{aligned}
h_i^*(\widehat\psi) =& h_i(\widehat\psi) - \ell_i(\widehat\psi) \\
=& \frac{k}{4}(z_{\alpha/2}^2+1)\left\{\hat\gamma_i^2\tau_1(\widehat\psi) - 2\hat\gamma_i\tau_2(\widehat\psi)+\tau_3(\widehat\psi)\right\}.
\end{aligned} \tag{2.11}
$$

In fact, noting that $mse_i(\hat\sigma^2,\widehat\psi) = (\hat\sigma^2/n_i)(1-\hat\gamma_i)\{1+\ell_i(\widehat\psi)/k\}$ and $\ell_i(\widehat\psi) = O_p(1)$, we can see that

$$
\begin{aligned}
[1+\frac{1}{2k}h_i^*(\widehat\psi)]\sqrt{mse_i(\hat\sigma^2,\widehat\psi)} =& [1+\frac{1}{2k}h_i^*(\widehat\psi)]\sqrt{\hat\sigma^2(1-\hat\gamma_i)/n_i}\sqrt{1+\ell_i(\widehat\psi)/k} \\
=& [1+\frac{1}{2k}\{h_i^*(\widehat\psi)+\ell_i(\widehat\psi)\}]\sqrt{\hat\sigma^2(1-\hat\gamma_i)/n_i}+O_p(k^{-3/2}),
\end{aligned}
$$

and $h_i^*(\widehat\psi)+\ell_i(\widehat\psi) = h_i(\widehat\psi)$. This implies that the two corrected empirical Bayes confidence intervals $I_i^{CEB}(\widehat\psi,\hat\sigma^2)$ and $I_i^{CEB*}(\widehat\psi,\hat\sigma^2)$ are equivalent in the second-order asymptotics. Hence from Theorem 2.1, we get the following proposition.

Proposition 2.4 *Assume the conditions* (A1)-(A4)*. Then, the corrected empirical Bayes confidence interval* $I_i^{CEB*}(\widehat\psi,\hat\sigma^2)$ *given in* (2.10) *satisfies that* $P[\mu_i \in I_i^{CEB*}(\widehat\psi,\hat\sigma^2)] = 1-\alpha+O(k^{-3/2})$ *as* $k\to\infty$.

We now provide a couple of examples for some specific estimators of σ^2 and σ_v^2.

Example 2.1 (Prasad-Rao estimators) Prasad and Rao (1990) suggested estimators based on unbiased estimators of σ^2 and σ_v^2, which are useful because they have simple and explicit forms. Let $\widehat{\boldsymbol{\beta}}_1 = \boldsymbol{B}^- \sum_{i=1}^k \sum_{j=1}^{n_i} (\boldsymbol{x}_{ij} - \overline{\boldsymbol{x}}_i)(y_{ij} - \overline{y}_i)$ and $S_1 = \sum_{i=1}^k \sum_{j=1}^{n_i} \big\{(y_{ij} - \overline{y}_i) - (\boldsymbol{x}_{ij} - \overline{\boldsymbol{x}}_i)'\widehat{\boldsymbol{\beta}}_1\big\}^2$ where $\boldsymbol{B}$ is given below (1.6) and $\boldsymbol{B}^-$ denotes the generalized inverse of the matrix $\boldsymbol{B}$. Then, an unbiased estimator of σ^2 is given by

$$\hat{\sigma}^{2U} = S_1/(N - k - r_1), \tag{2.12}$$

where $N = \sum_{i=1}^k n_i$, and r_1 is the rank of the matrix $\boldsymbol{B}$. It can be seen that S_1 is independent of $\overline{y}_1, \dots, \overline{y}_k$ and that $S_1/\sigma^2 \sim \chi^2_{N-k-r_1}$ and $\overline{y}_i \sim \mathcal{N}(\overline{\boldsymbol{x}}_i'\boldsymbol{\beta}, \sigma^2/(n_i\gamma_i))$ for $i = 1, \dots, k$. To estimate σ_v^2, we can use the Henderson method III. For $\widehat{\boldsymbol{\beta}}_0 = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$, we consider the sum of squares $S = (\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}_0)'(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}_0)$, which can be rewritten as $S = (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) - (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})$. The expectation of S is $E[S] = \mathrm{tr}\,[\boldsymbol{I}_N - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}']\mathbf{Cov}\,(\boldsymbol{y}) = \sigma^2(N - p) + N_*\sigma_v^2$, where I_N is the $N \times N$ identity matrix, and $N_* = N - \mathrm{tr}\,\{(\boldsymbol{X}'\boldsymbol{X})^{-1}\sum_{i=1}^k n_i^2 \overline{\boldsymbol{x}}_i\overline{\boldsymbol{x}}_i'\}$. Hence, an unbiased estimator of σ_v^2 is given by

$$\hat{\sigma}_v^{2U} = N_*^{-1}\left\{S - (N - p)\hat{\sigma}^{2U}\right\}.$$

Based on $\hat{\sigma}^{2U}$ and $\hat{\sigma}_v^{2U}$, the ratio ψ can be estimated by

$$\widehat{\psi}^U = \hat{\sigma}_v^{2U}/\hat{\sigma}^{2U} = N_*^{-1}\left\{S/\hat{\sigma}^{2U} - (N - p)\right\}.$$

However, this estimator has a drawback of taking negative values with a positive probability, and it may be serious when ψ is small. Instead of $\widehat{\psi}^U$, we here use the truncated estimator

$$\widehat{\psi}^{TR} = \max\{\widehat{\psi}^U, k^{-2/3}\}, \tag{2.13}$$

which is positive, consistent and $\widehat{\psi}^{TR} = \widehat{\psi}^U + o_p(k^{-a})$ for any $a > 0$ as shown in Proposition 2.2.

We here verify that $\partial\widehat{\psi}^{TR}/\partial\overline{y}_i = O_p(k^{-1})$ as $k \to \infty$. From Proposition 2.2 and the definition of $\widehat{\psi}^U$, it is sufficient to show that $\partial S/\partial\overline{y}_i = O_p(1)$. It is noted that S is expressed as $S = S_{(1)} + S_{(2)}$ where $S_{(1)} = \sum_{i=1}^k \sum_{j=1}^{n_i} \{(y_{ij} - \overline{y}_i) - (\boldsymbol{x}_{ij} - \overline{\boldsymbol{x}}_i)'\widehat{\boldsymbol{\beta}}_0\}^2$ and $S_{(2)} = \sum_{i=1}^k n_i(\overline{y}_i - \overline{\boldsymbol{x}}_i'\widehat{\boldsymbol{\beta}}_0)^2$. Since $\widehat{\boldsymbol{\beta}}_0 = \widehat{\boldsymbol{\beta}}_1 - (\boldsymbol{B} + \boldsymbol{A}_0)^{-1}\boldsymbol{A}_0(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_{20})$ for $\boldsymbol{A}_0 = \boldsymbol{A}(0) = \sum_{i=1}^k n_i\overline{\boldsymbol{x}}_i\overline{\boldsymbol{x}}_i'$ and $\widehat{\boldsymbol{\beta}}_{20} = \boldsymbol{A}_0^{-1}\sum_{i=1}^k n_i\overline{\boldsymbol{x}}_i\overline{y}_i$, $S_{(1)}$ can be expressed as

$$S_{(1)} = S_1 + (\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_{20})'\boldsymbol{A}_0(\boldsymbol{B} + \boldsymbol{A}_0)^{-1}\boldsymbol{B}(\boldsymbol{B} + \boldsymbol{A}_0)^{-1}\boldsymbol{A}_0(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_{20}).$$

Since $\widehat{\boldsymbol{\beta}}_0 = \widehat{\boldsymbol{\beta}}_{20} + (\boldsymbol{B} + \boldsymbol{A}_0)^{-1}\boldsymbol{B}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_{20})$, $S_{(2)}$ is rewritten as

$$S_{(2)} = S_2 + (\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_{20})'\boldsymbol{B}(\boldsymbol{B} + \boldsymbol{A}_0)^{-1}\boldsymbol{A}_0(\boldsymbol{B} + \boldsymbol{A}_0)^{-1}\boldsymbol{B}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_{20}).$$

Combining these expressions of $S_{(1)}$ and $S_{(2)}$ gives that

$$S = S_1 + S_2 + (\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_{20})'\boldsymbol{C}(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_{20}),$$

where

$$\boldsymbol{C} = \boldsymbol{A}_0(\boldsymbol{B} + \boldsymbol{A}_0)^{-1}\boldsymbol{B}(\boldsymbol{B} + \boldsymbol{A}_0)^{-1}\boldsymbol{A}_0 + \boldsymbol{B}(\boldsymbol{B} + \boldsymbol{A}_0)^{-1}\boldsymbol{A}_0(\boldsymbol{B} + \boldsymbol{A}_0)^{-1}\boldsymbol{B},$$

which is equal to $\boldsymbol{C} = \boldsymbol{A}_0 - \boldsymbol{A}_0(\boldsymbol{B} + \boldsymbol{A}_0)^{-1}\boldsymbol{A}_0$. Noting that S_1 and $\widehat{\boldsymbol{\beta}}_1$ are independent of $(\overline{y}_1, \ldots, \overline{y}_k)$, it follows that

$$\frac{\partial S}{\partial \overline{y}_i} = \frac{\partial S_2}{\partial \overline{y}_i} - 2(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_{20})'\boldsymbol{C}\frac{\partial \widehat{\boldsymbol{\beta}}_{20}}{\partial \overline{y}_i}.$$

Since $\partial S_2/\partial \overline{y}_i = 2n_i\{\overline{y}_i - \overline{\boldsymbol{x}}_i'\widehat{\boldsymbol{\beta}}_{20}\}(1 + n_i\overline{\boldsymbol{x}}_i'\boldsymbol{A}_0^{-1}\overline{\boldsymbol{x}}_i) + 2n_i\sum_{j\neq i} n_j\{\overline{y}_j - \overline{\boldsymbol{x}}_j'\widehat{\boldsymbol{\beta}}_{20}\}\overline{\boldsymbol{x}}_j'\boldsymbol{A}_0^{-1}\overline{\boldsymbol{x}}_i$, it is seen that $\partial S_2/\partial \overline{y}_i = O_p(1)$. Also, it is observed that $(\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_{20})'\boldsymbol{C}\{\partial\widehat{\boldsymbol{\beta}}_{20}/\partial \overline{y}_i\} = (\widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_{20})'\boldsymbol{C}\boldsymbol{A}_0^{-1}n_i\overline{\boldsymbol{x}}_i = O_p(1)$. Thus, it is shown that $\partial\widehat{\psi}^{TR}/\partial \overline{y}_i = O_p(k^{-1})$.

To get the corrected empirical Bayes confidence, we need to evaluate the moments of T_e and T_v for $T_v = (\hat{\sigma}_v^{2U} - \sigma_v^2)/\sigma_v^2$ and $T_e = (\hat{\sigma}^{2U} - \sigma^2)/\sigma^2$. Clearly, $E[T_e] = 0$ and $E[T_v] = 0$. From (5.4)-(5.6) in Prasad and Rao (1891), it follows that $E[T_e^2] = 2/(N-k) + O(k^{-3/2})$, $E[T_vT_e] = -2k/\{N(N-k)\psi\} + O(k^{-3/2})$ and $E[T_v^2] = 2(N^2\psi^2)^{-1}\big[k^2/(N-k) + \sum_{i=1}^k \gamma_i^{-2}\big] + O(k^{-3/2})$. Then the values of $\tau_1(\psi)$, $\tau_2(\psi)$ and $\tau_3(\psi)$ given in Proposition 2.1 for the unbiased estimators $\hat{\sigma}^{2U}$ and $\hat{\sigma}_v^{2U}$ can be approximated as

$$\begin{aligned}
\tau_1(\psi) =& \frac{2}{N^2\psi^2}\Big\{\sum_{i=1}^k \gamma_i^{-2} + \frac{1}{N-k}\big(\sum_{i=1}^k \gamma_i^{-1}\big)^2\Big\} + O(k^{-3/2}),\\
\tau_2(\psi) =& \frac{2}{N(N-k)\psi}\sum_{i=1}^k \gamma_i^{-1} + O(k^{-3/2}),\\
\tau_3(\psi) =& 2/(N-k) + O(k^{-3/2}).
\end{aligned}$$

These approximations with the estimator $\widehat{\psi}^{TR}$ are substituted into (2.4) and (2.11) to get $h_i(\widehat{\psi}^{TR})$ and $h_i^*(\widehat{\psi}^{TR})$, and the corresponding corrected confidence intervals are obtained. ■

Example 2.2 (ML and REML estimators) The general method for estimating variance components is the maximum likelihood (ML) or the restricted maximum likelihood (REML) estimators, though the iteration methods such as the Newton method are necessary for solving the likelihood equations numerically.

The moments of the ML and REML estimators can be derived by using the arguments as in Datta and Lahiri (2000). The ML estimator $\widehat{\psi}^*$ of ψ is given as the solution of the equation

$$\sum_{i=1}^k \{n_i\gamma_i(\widehat{\psi}^*)\}^2\{\overline{y}_i - \overline{\boldsymbol{x}}_i'\widehat{\boldsymbol{\beta}}(\widehat{\psi}^*)\}^2 = \hat{\sigma}^2(\widehat{\psi}^*)\sum_{i=1}^k n_i\gamma_i(\widehat{\psi}^*),$$

where $\hat{\sigma}^2(\psi)$ is defined by

$$\hat{\sigma}^2(\psi) = \frac{1}{N}\Big(\sum_{i=1}^k\sum_{j=1}^{n_i}\big\{(y_{ij} - \overline{y}_i) - (\boldsymbol{x}_{ij} - \overline{\boldsymbol{x}}_i)'\widehat{\boldsymbol{\beta}}(\psi)\big\}^2 + \sum_{i=1}^k n_i\gamma_i(\psi)\big\{\overline{y}_i - \overline{\boldsymbol{x}}_i'\widehat{\boldsymbol{\beta}}(\psi)\big\}^2\Big).$$

From Proposition 2.2, we here use the truncated estimator of $\widehat{\psi}^*$ given by $\widehat{\psi}^M = \max\big\{\widehat{\psi}^*, k^{-2/3}\big\}$. Then, the ML estimators of σ^2 and σ_v^2 are written by $\hat{\sigma}^{2M} = \hat{\sigma}^2(\widehat{\psi}^M)$ and $\hat{\sigma}_v^{2M} = \hat{\sigma}^{2M}\widehat{\psi}^M$.

Let $\boldsymbol{\xi} = (\xi_1, \xi_2)' = (\sigma^2, \sigma_v^2)'$ and let $\ell(\boldsymbol{\beta}, \boldsymbol{\xi})$ be the log likelihood function of $(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k)$. The Fisher information matrices are given by $\boldsymbol{I}_{\boldsymbol{\beta},\boldsymbol{\beta}} = -E[\partial^2 \ell(\boldsymbol{\beta}, \boldsymbol{\xi})/\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}']$ and $\boldsymbol{I}_{\boldsymbol{\xi},\boldsymbol{\xi}} = -E[\partial^2 \ell(\boldsymbol{\beta}, \boldsymbol{\xi})/\partial \boldsymbol{\xi} \partial \boldsymbol{\xi}']$, where $\boldsymbol{I}_{\boldsymbol{\beta},\boldsymbol{\xi}} = -E[\partial^2 \ell(\boldsymbol{\beta}, \boldsymbol{\xi})/\partial \boldsymbol{\beta} \partial \boldsymbol{\xi}'] = \mathbf{0}$ because of the orthogonality of the parameters $(\boldsymbol{\beta}, \boldsymbol{\xi})$. From Datta and Lahiri (2000) and Rao (2003), it follows that $\boldsymbol{I}_{\boldsymbol{\beta},\boldsymbol{\beta}} = (\boldsymbol{A}(\psi) + \boldsymbol{B})/\sigma^2$,

$$\boldsymbol{I}_{\boldsymbol{\xi},\boldsymbol{\xi}} = \frac{1}{2\sigma^4} \begin{pmatrix} N - k + \sum_{i=1}^k \gamma_i^2 & \sum_{i=1}^k n_i \gamma_i^2 \\ \sum_{i=1}^k n_i \gamma_i^2 & \sum_{i=1}^k n_i^2 \gamma_i^2 \end{pmatrix},$$

and

$$\widehat{\boldsymbol{\xi}}^M - \boldsymbol{\xi} = \frac{1}{2\sigma^2} \boldsymbol{I}_{\boldsymbol{\xi},\boldsymbol{\xi}}^{-1} \begin{pmatrix} g_1(\boldsymbol{y}) \\ g_2(\boldsymbol{y}) \end{pmatrix} + O_p(k^{-1}),$$

where $\widehat{\boldsymbol{\xi}}^M = (\hat{\sigma}^{2M}, \hat{\sigma}_v^{2M})'$, $g_1(\boldsymbol{y}) = \sum_{i=1}^k \sum_{j=1}^{n_i} \{(y_{ij} - \overline{y}_i) - (\boldsymbol{x}_{ij} - \overline{\boldsymbol{x}}_i)'\boldsymbol{\beta}\}^2/\sigma^2 + \sum_{i=1}^k n_i \gamma_i (2 - \gamma_i)(\overline{y}_i - \overline{\boldsymbol{x}}_i'\boldsymbol{\beta})^2/\sigma^2 - N - \sum_{i=1}^k (1 - \gamma_i)$ and $g_2(\boldsymbol{y}) = \sum_{i=1}^k n_i^2 \gamma_i^2 (\overline{y}_i - \overline{\boldsymbol{x}}_i'\boldsymbol{\beta})^2/\sigma^2 - \sum_{i=1}^k n_i \gamma_i$. Noting that $\boldsymbol{I}_{\boldsymbol{\xi},\boldsymbol{\xi}} = O(k)$, $\partial g_1(\boldsymbol{y})/\partial \overline{y}_i = 2 n_i \gamma_i (2 - \gamma_i)(\overline{y}_i - \overline{\boldsymbol{x}}_i'\boldsymbol{\beta})/\sigma^2 = O_p(1)$ and $\partial g_2(\boldsymbol{y})/\partial \overline{y}_i = 2 n_i^2 \gamma_i^2 (\overline{y}_i - \overline{\boldsymbol{x}}_i'\boldsymbol{\beta})/\sigma^2 = O_p(1)$, we can see that $\partial \widehat{\boldsymbol{\xi}}^M/\partial \overline{y}_i = O_p(k^{-1})$. Combining this fact and the equation

$$\frac{\partial \widehat{\psi}^M}{\partial \overline{y}_i} = \frac{1}{\hat{\sigma}^{2M}} \frac{\partial \hat{\sigma}_v^{2M}}{\partial \overline{y}_i} - \frac{\widehat{\psi}^M}{\hat{\sigma}^{2M}} \frac{\partial \hat{\sigma}^{2M}}{\partial \overline{y}_i}$$

can show that $\partial \widehat{\psi}^M/\partial \overline{y}_i = O_p(k^{-1})$. From Datta and Lahiri (2000), it follows that $\mathbf{Var}(\widehat{\boldsymbol{\xi}}^M) = \boldsymbol{I}_{\boldsymbol{\xi},\boldsymbol{\xi}}^{-1} + O(k^{-3/2})$ and

$$\mathbf{Bias}\left(\widehat{\boldsymbol{\xi}}^M\right) = \frac{1}{2} \boldsymbol{I}_{\boldsymbol{\xi},\boldsymbol{\xi}}^{-1} \begin{pmatrix} \operatorname{tr}[\boldsymbol{I}_{\boldsymbol{\beta},\boldsymbol{\beta}}^{-1}(\partial \boldsymbol{I}_{\boldsymbol{\beta},\boldsymbol{\beta}}/\partial \xi_1)] \\ \operatorname{tr}[\boldsymbol{I}_{\boldsymbol{\beta},\boldsymbol{\beta}}^{-1}(\partial \boldsymbol{I}_{\boldsymbol{\beta},\boldsymbol{\beta}}/\partial \xi_2)] \end{pmatrix} + O(k^{-3/2}).$$

Hence,

$$\begin{aligned} Var(\hat{\sigma}^{2M}) =& 2\sigma^4 \sum_{i=1}^k n_i^2 \gamma_i^2 / d(\psi) + O(k^{-3/2}) = \sigma^4 E[T_e^2], \\ Var(\hat{\sigma}_v^{2M}) =& 2\sigma^4 (N - k + \sum_{i=1}^k \gamma_i^2)/d(\psi) + O(k^{-3/2}) = \sigma_v^4 E[T_v^2], \\ Cov(\hat{\sigma}^{2M}, \hat{\sigma}_v^{2M}) =& -2\sigma^4 \sum_{i=1}^k n_i \gamma_i^2 / d(\psi) + O(k^{-3/2}) = \sigma^2 \sigma_v^2 E[T_e T_v], \end{aligned}$$

where $d(\psi) = (N - k + \sum_{i=1}^k \gamma_i^2) \sum_{i=1}^k n_i^2 \gamma_i^2 - (\sum_{i=1}^k n_i \gamma_i^2)^2$. Then it is seen that $\tau_1(\psi)$, $\tau_2(\psi)$ and $\tau_3(\psi)$ in Proposition 2.1 are given by $\tau_1(\psi) = 2N/\{\psi^2 d(\psi)\} + O(k^{-3/2})$, $\tau_2(\psi) = 2\sum_{i=1}^k n_i \gamma_i / \{\psi d(\psi)\} + O(k^{-3/2})$ and $\tau_3(\psi) = 2\sum_{i=1}^k n_i^2 \gamma_i^2 / d(\psi) + O(k^{-3/2})$. Also, the biases are given by

$$\begin{aligned} Bias(\hat{\sigma}^{2M}) =& \frac{\sigma^2}{d(\psi)} \Big\{ -p \sum_{i=1}^k n_i^2 \gamma_i^2 + (\sum_{i=1}^k n_i \gamma_i) c(\psi) \Big\} + O(k^{-3/2}) = \sigma^2 E[T_e], \\ Bias(\hat{\sigma}_v^{2M}) =& \frac{\sigma^2}{d(\psi)} \Big\{ p \sum_{i=1}^k n_i \gamma_i^2 - (N - k + \sum_{i=1}^k \gamma_i) c(\psi) \Big\} + O(k^{-3/2}) = \sigma_v^2 E[T_v], \end{aligned}$$

where $c(\psi) = \mathrm{tr}\,[(\boldsymbol{A}(\psi)+\boldsymbol{B})^{-1}\sum_{i=1}^k n_i^2\gamma_i^2\overline{\boldsymbol{x}}_i\overline{\boldsymbol{x}}_i']$. Substituting these quantities into (2.4) and (2.11), we get the correction terms $h_i(\psi)$ and $h_i^*(\psi)$ for the ML estimators $\hat{\sigma}^{2M}$ and $\widehat{\psi}^M$. Thus, Theorem 2.1 implies that the corresponding corrected confidence intervals based on the ML estimators has the coverage accuracy $O(k^{-3/2})$.

For the REML estimator of ψ, it is given as the solution ψ_* of the equation

$$\sum_{i=1}^k \{n_i\gamma_i(\widehat{\psi}_*)\}^2\{\overline{y}_i - \overline{\boldsymbol{x}}_i'\widehat{\boldsymbol{\beta}}(\widehat{\psi}_*)\}^2$$
$$= \frac{N}{N-p}\hat{\sigma}^2(\widehat{\psi}_*)\sum_{i=1}^k n_i\gamma_i(\widehat{\psi}_*) - \sum_{i=1}^k \{n_i\gamma_i(\widehat{\psi}_*)\}^2\overline{\boldsymbol{x}}_i'(\boldsymbol{A}(\widehat{\psi}_*)+\boldsymbol{B})^{-1}\overline{\boldsymbol{x}}_i.$$

From Proposition 2.2, we use the truncated estimator of $\widehat{\psi}_*$ given by $\widehat{\psi}^R = \max\{\widehat{\psi}_*, k^{-2/3}\}$. Then, the REML estimators of σ^2 and σ_v^2 are written by $\hat{\sigma}^{2R} = \{N/(N-p)\}\hat{\sigma}^2(\widehat{\psi}^R)$ and $\hat{\sigma}_v^{2R} = \hat{\sigma}^{2R}\widehat{\psi}^R$. From Datta and Lahiri (2000), it follows that $Bias(\hat{\sigma}^{2R}) = Bias(\hat{\sigma}_v^{2R}) = O(k^{-3/2})$, and that the asymptotic variance and covariance of $\hat{\sigma}^{2R}$ and $\hat{\sigma}_v^{2R}$ are equal to those of the ML estimators $\hat{\sigma}^{2M}$ and $\hat{\sigma}_v^{2M}$. Thus, substituting these quantities into (2.4) and (2.11), we get the correction terms for the REML estimators $\hat{\sigma}^{2R}$ and $\widehat{\psi}^R$. ■

2.2 Extension to the estimation of finite population means

The results given in Section 2.1 can be extended to the estimation of means of k finite populations. Let Y_{ij} denote the value of a characteristic of interest for the jth unit of the ith finite population where $i = 1,\ldots,k;\ j = 1,\ldots,N_i$. We assume that there exist the auxiliary variables $\boldsymbol{x}_{ij}$ which can be associated with Y_{ij} as

$$Y_{ij} = \boldsymbol{x}_{ij}'\boldsymbol{\beta} + v_i + e_{ij}, \quad i = 1,\ldots,k,\ j = 1,\ldots,N_i,$$

where $\boldsymbol{x}_{ij}$, $\boldsymbol{\beta}$, v_i and e_{ij} are defined similarly as in (1.1). We also assume that the population sizes N_i's and all the auxiliary variables $\boldsymbol{x}_{ij}$'s are known and bounded, but only some of the Y_{ij}'s are observed through the following sampling procedure. For each i, a subset of $\{1,\ldots,N_i\}$ is called a sample from the ith population. Let S_i denote the set of all possible samples of fixed known size n_i taken from $\{1,\ldots,N_i\}$. A sampling design $p(s_i)$ is the probability of selecting the sample s_i from S_i. Then $p[s_i] \geq 0$ for all $s_i \in S_i$ with $\sum_{s_i\in S_i} p[s_i] = 1$. Let $\boldsymbol{s} = (s_1,\ldots,s_k)$ and $\boldsymbol{S} = S_1\times\cdots\times S_k$. Since the sampling is carried out independetly for $i = 1,\ldots,k$, it is seen that $P[\boldsymbol{s}] = P[s_1]\times\cdots\times P[s_k]$ and $\sum_{\boldsymbol{s}\in\boldsymbol{S}} P[\boldsymbol{s}] = 1$. Given $\boldsymbol{s}$, the subset of $\{Y_{i1},\ldots,Y_{iN_i}\}$ is observed for $i = 1,\ldots,k$. Suppose, without of loss of generality $s_i = \{1,\ldots,n_i\}$. Thus, the sampled values of Y_{ij} are denoted by $y_{i1},\ldots,y_{in_i}$, and unobserved varaibles are denoted by $Y_{i,n_i+1}^*,\ldots,Y_{i,N_i}^*$.

The objective is to estimate the population mean $\overline{Y}_i = N_i^{-1}\sum_{j=1}^{N_i} Y_{ij}$ based on the samples $s_1,\ldots,s_k$, namely, $\{y_{ij}|j = 1,\ldots,n_i; i = 1,\ldots,k\}$ and the auxilliary variables $\{\boldsymbol{x}_{ij}|j = 1,\ldots,N_i; i = 1,\ldots,k\}$. Let $\boldsymbol{y}_i = (y_{i1},\ldots,y_{in_i})'$ and $\boldsymbol{Y}_i^* = (Y_{i,n_i+1}^*,\ldots,Y_{i,N_i}^*)'$. Correspondingly, let $\boldsymbol{X}_i = (\boldsymbol{x}_{i1},\ldots,\boldsymbol{x}_{i,n_i})'$ and $\boldsymbol{X}_i^* = (\boldsymbol{x}_{i,n_i+1},\ldots,\boldsymbol{x}_{i,N_i})'$. It is noted that given s_i, $(\boldsymbol{y}_i, \boldsymbol{Y}_i^*)$ is jointly distributed as

$$\begin{pmatrix} \boldsymbol{y}_i \\ \boldsymbol{Y}_i^* \end{pmatrix} | s_i \sim \mathcal{N}_{N_i}\left(\begin{pmatrix} \boldsymbol{X}_i \\ \boldsymbol{X}_i^* \end{pmatrix}\boldsymbol{\beta}, \sigma^2 \begin{pmatrix} \boldsymbol{\Lambda}_{11} & \boldsymbol{\Lambda}_{12} \\ \boldsymbol{\Lambda}_{21} & \boldsymbol{\Lambda}_{22} \end{pmatrix}\right),$$

where $\boldsymbol{\Lambda}_{11} = \boldsymbol{I}_{n_i} + \psi \boldsymbol{j}_{n_i}\boldsymbol{j}_{n_i}'$, $\boldsymbol{\Lambda}_{12} = \boldsymbol{\Lambda}_{21}' = \psi \boldsymbol{j}_{n_i}\boldsymbol{j}_{N_i-n_i}'$ and $\boldsymbol{\Lambda}_{22} = \boldsymbol{I}_{N_i-n_i} + \psi \boldsymbol{j}_{N_i-n_i}\boldsymbol{j}_{N_i-n_i}'$. The conditional distribution of $\boldsymbol{Y}_i^*$ given $\boldsymbol{y}_i$ and s_i is

$$\boldsymbol{Y}_i^*|\boldsymbol{y}_i, s_i \sim \mathcal{N}_{N_i-n_i}(\boldsymbol{X}_i^*\boldsymbol{\beta} + \boldsymbol{\Lambda}_{21}\boldsymbol{\Lambda}_{11}^{-1}(\boldsymbol{y}_i - \boldsymbol{X}_i\boldsymbol{\beta}),\ \sigma^2\boldsymbol{\Lambda}_{22.1}),$$

where $\boldsymbol{\Lambda}_{22.1} = \boldsymbol{\Lambda}_{22} - \boldsymbol{\Lambda}_{21}\boldsymbol{\Lambda}_{11}^{-1}\boldsymbol{\Lambda}_{12}$. Since $\boldsymbol{\Lambda}_{21}\boldsymbol{\Lambda}_{11}^{-1} = \psi(1+n_i\psi)^{-1}\boldsymbol{j}_{N_i-n_i}\boldsymbol{j}_{n_i}'$ and $\boldsymbol{\Lambda}_{22.1} = \boldsymbol{I}_{N_i-n_i} + \psi(1+n_i\psi)^{-1}\boldsymbol{j}_{N_i-n_i}\boldsymbol{j}_{N_i-n_i}'$, it is seen that

$$\boldsymbol{Y}_i^*|\boldsymbol{y}_i, s_i \sim \mathcal{N}_{N_i-n_i}\Big(\boldsymbol{X}_i^*\boldsymbol{\beta} + (1-\gamma_i)(\overline{y}_i - \overline{\boldsymbol{x}}_i'\boldsymbol{\beta})\boldsymbol{j}_{N_i-n_i},\ \sigma^2(\boldsymbol{I}_{N_i-n_i} + \psi\gamma_i\boldsymbol{j}_{N_i-n_i}\boldsymbol{j}_{N_i-n_i}')\Big),$$

for $\gamma_i = 1/(1+n_i\psi)$. It is here noted that $\overline{Y}_i$ is expressed as

$$\begin{aligned}\overline{Y}_i =&(y_{i1} + \cdots + y_{i,n_i})/N_i + (Y_{i,n_i+1}^* + \cdots + Y_{i,N_i}^*)/N_i \\ =&(1-f_i)\overline{y}_i + \boldsymbol{j}_{N_i-n_i}'\boldsymbol{Y}_i^*/N_i,\end{aligned}$$

for $f_i = (N_i - n_i)/N_i$. Then from the above conditinal distribution, it follows that

$$\overline{Y}_i|\overline{y}_i, s_i \sim \mathcal{N}(\widehat{\mu}_i^B(\boldsymbol{\beta}, \psi, s_i),\ \sigma^2 f_i(1-f_i\gamma_i)/n_i), \tag{2.14}$$

where $\mu_i^B(\boldsymbol{\beta}, \psi, s_i)$ is the conditional mean $E[\overline{Y}_i|\overline{y}_i, s_i]$ given by

$$\begin{aligned}\widehat{\mu}_i^B(\boldsymbol{\beta}, \psi, s_i) =&(1-f_i)\overline{y}_i + f_i\{(\overline{\boldsymbol{x}}_i^*)'\boldsymbol{\beta} + (1-\gamma_i)(\overline{y}_i - \overline{\boldsymbol{x}}_i'\boldsymbol{\beta})\} \\ =&\overline{\boldsymbol{x}}_{i(p)}'\boldsymbol{\beta} + (1-f_i\gamma_i)(\overline{y}_i - \overline{\boldsymbol{x}}_i'\boldsymbol{\beta}),\end{aligned} \tag{2.15}$$

for $\overline{\boldsymbol{x}}_i^* = (N_i - n_i)^{-1}\sum_{j=n_i+1}^{N_i}\boldsymbol{x}_{ij}$ and $\overline{\boldsymbol{x}}_{i(p)} = N_i^{-1}\sum_{j=1}^{N_i}\boldsymbol{x}_{ij}$, since

$$\overline{\boldsymbol{x}}_{i(p)} = (1-f_i)\overline{\boldsymbol{x}}_i + f_i\overline{\boldsymbol{x}}_i^*.$$

The estimator $\widehat{\mu}_i^B(\boldsymbol{\beta}, \psi, s_i)$ is viewed as the Bayes estimator in the finite population as stated in Ghosh and Meeden (1986). Since conditional distribution (2.14) and conditional mean (2.15) correspond to posterior distribution (1.3) and Bayes estimator (1.4), the same arguments as in the previous sections can be used to construct the corrected confidence interval of $\overline{Y}_i$.

Using the estimators $\widehat{\psi}$, $\widehat{\boldsymbol{\beta}}(\widehat{\psi})$ and $\hat{\sigma}^2$ given in Section 2.1, we get the empirical Bayes estimator $\widehat{\mu}_i^{EB}(\widehat{\psi}, \boldsymbol{s}) = \widehat{\mu}_i^B(\widehat{\boldsymbol{\beta}}(\widehat{\psi}), \widehat{\psi}, s_i)$ and provide the corrected empirical Bayes confidence interval

$$I_i^{CEB}(\widehat{\psi}, \hat{\sigma}^2, \boldsymbol{s}):\ \widehat{\mu}_i^{EB}(\widehat{\psi}, \boldsymbol{s}) \pm z_{\alpha/2}\left[1 + (2k)^{-1}h_i(\widehat{\psi}, \boldsymbol{s})\right]\sqrt{(\hat{\sigma}^2/n_i)f_i(1-f_i\hat{\gamma}_i)}, \tag{2.16}$$

where $h_i(\psi, \boldsymbol{s})$ is given by

$$\begin{aligned}h_i(\psi, \boldsymbol{s}) =&\frac{kn_i}{f_i(1-f_i\gamma_i)}\Big\{\boldsymbol{d}_i'(\boldsymbol{A}(\psi) + \boldsymbol{B})^{-1}\boldsymbol{d}_i + 2f_i^2 n_i\gamma_i^3 E[(\widehat{\psi} - \psi)^2]\Big\} \\ &- kE[H_i] + \frac{k}{4}E[z_{\alpha/2}^2 H_i^2 + K_i^2],\end{aligned} \tag{2.17}$$

where

$$\boldsymbol{d}_i = f_i\{\overline{\boldsymbol{x}}_i^* - (1-\gamma_i)\overline{\boldsymbol{x}}_i\} = \overline{\boldsymbol{x}}_{i(p)} - (1-f_1\gamma_i)\overline{\boldsymbol{x}}_i,$$

and

$$\begin{aligned} H_i =& a_i(\widehat{\psi}-\psi)/\psi + (\hat\sigma^2-\sigma^2)/\sigma^2, \\ K_i =& a_i(\widehat{\psi}-\psi)/\psi - (\hat\sigma^2-\sigma^2)/\sigma^2, \end{aligned} \tag{2.18}$$

for $a_i = f_i n_i \psi\gamma_i^2/(1-f_i\gamma_i)$. Then, we can get the following theorem which will be proved in Section 5.

Theorem 2.2 *Assume the conditions (A1)-(A4). Then, the corrected empirical Bayes confidence interval given in (2.16) satisfies that* $P[\overline{Y}_i \in I_i^{CEB}(\widehat{\psi},\hat\sigma^2,\boldsymbol{s})] = 1-\alpha+O(k^{-3/2})$ *as* $k\to\infty$.

Proposition 2.5 *Let* $\hat\sigma_v^2 = \hat\sigma^2\widehat{\psi}$ *and assume the conditions (A3) and (A4). Then, the correction function* $h_i(\psi)$ *given in (2.17) can be expressed as*

$$\begin{aligned} h_i(\psi,\boldsymbol{s}) =& \frac{kn_i}{f_i(1-f_i\gamma_i)}\Big\{\boldsymbol{d}_i'(\boldsymbol{A}(\psi)+\boldsymbol{B})^{-1}\boldsymbol{d}_i + 2f_i^2n_i\gamma_i^3\psi^2\tau_1(\psi)\Big\} - k\{(1-a_i)b_e(\psi)+a_ib_v(\psi)\} \\ &+ \frac{k}{4}(z_{\alpha/2}^2+1)\left\{a_i^2\tau_1(\psi) - 2a_i\tau_2(\psi)+\tau_3(\psi)\right\} + O(k^{-3/2}), \end{aligned} \tag{2.19}$$

where $b_e(\psi)$, $b_v(\psi)$, $\tau_1(\psi)$, $\tau_2(\psi)$ *and* $\tau_3(\psi)$ *are defined in Proposition 2.1.*

Proposition 2.6 *Assume the conditions (A3) and (A4). Given* $\boldsymbol{s}$*, the second-order approximation of the MSE of the empirical Bayes estimator* $\widehat{\mu}_i^{EB}(\widehat{\psi},\boldsymbol{s})$ *is given by*

$$\begin{aligned} MSE_i(\sigma^2,\psi,\boldsymbol{s}) =& E[(\widehat{\mu}_i^{EB}(\widehat{\psi},\boldsymbol{s})-\overline{Y}_i)^2] \\ =& \frac{\sigma^2}{n_i}f_i(1-f_i\gamma_i) + \sigma^2\boldsymbol{d}_i'(\boldsymbol{A}(\psi)+\boldsymbol{B})^{-1}\boldsymbol{d}_i + \sigma^2 f_i^2 n_i\gamma_i^3\psi^2\tau_1(\psi) + O(k^{-3/2}), \end{aligned} \tag{2.20}$$

and the second-order unbiased estimator of the MSE is given by

$$\begin{aligned} mse_i(\hat\sigma^2,\widehat{\psi},\boldsymbol{s}) =& \frac{\hat\sigma^2}{n_i}f_i(1-f_i\hat\gamma_i)\big\{1-(1-\hat a_i)b_e(\widehat{\psi}) - \hat a_i b_v(\widehat{\psi})\big\} \\ &+ \hat\sigma^2\boldsymbol{d}_i'(\boldsymbol{A}(\widehat{\psi})+\boldsymbol{B})^{-1}\boldsymbol{d}_i + 2\hat\sigma^2 f_i 62 n_i\hat\gamma_i^3\widehat{\psi}^2\tau_1(\widehat{\psi}), \end{aligned} \tag{2.21}$$

for $\hat a_i = f_i n_i\widehat{\psi}\hat\gamma_i^2/(1-f_i\hat\gamma_i)$*, namely,* $E[mse_i(\hat\sigma^2,\widehat{\psi},\boldsymbol{s})] = MSE_i(\sigma^2,\psi,\boldsymbol{s})+O(k^{-3/2})$.

The empirical Bayes confidence interval and its corrected interval are given by

$$I_i^{EB*}(\widehat{\psi},\hat\sigma^2,\boldsymbol{s}) : \widehat{\mu}_i^{EB}(\widehat{\psi},\boldsymbol{s}) \pm z_{\alpha/2}\sqrt{mse_i(\hat\sigma^2,\widehat{\psi},\boldsymbol{s})}, \tag{2.22}$$

$$I_i^{CEB*}(\widehat{\psi},\hat\sigma^2,\boldsymbol{s}) : \widehat{\mu}_i^{EB}(\widehat{\psi},\boldsymbol{s}) \pm z_{\alpha/2}\left[1+(2k)^{-1}h_i^*(\widehat{\psi},\boldsymbol{s})\right]\sqrt{mse_i(\hat\sigma^2,\widehat{\psi},\boldsymbol{s})}, \tag{2.23}$$

where

$$h_i^*(\widehat{\psi}, \boldsymbol{s}) = h_i(\widehat{\psi}, \boldsymbol{s}) - k\Big\{mse_i(\hat{\sigma}^2, \widehat{\psi}, \boldsymbol{s})n_i/\{\hat{\sigma}^2 f_i(1-f_i\hat{\gamma}_i)\} - 1\Big\}$$
$$= \frac{k}{4}(z_{\alpha/2}^2 + 1)\left\{\hat{a}_i^2\tau_1(\widehat{\psi}) - 2\hat{a}_i\tau_2(\widehat{\psi}) + \tau_3(\widehat{\psi})\right\}.$$

The same argument as in the proof of Proposition 2.4 can be used to verify the following proposition.

Proposition 2.7 *Assume the conditions* (A1)-(A4). *Then, the corrected empirical Bayes confidence interval* $I_i^{CEB*}(\widehat{\psi}, \hat{\sigma}^2, \boldsymbol{s})$ *given in* (2.23) *satisfies that* $P[\mu_i \in I_i^{CEB*}(\widehat{\psi}, \hat{\sigma}^2, \boldsymbol{s})] = 1 - \alpha + O(k^{-3/2})$ *as* $k \to \infty$.

3 Numerical studies

3.1 Comparison of the confidence intervals

We now investigate the numerical performances of the confidence intervals given in the previous sections under nested error regression model (1.1) or (2.1) without covariates through simulations experiments. The confidence intervals we want to compare are the conventional confidence interval based on a t-distribution

$$I_i^T \ : \ \overline{y}_i \pm t_{\alpha/2}\sqrt{\hat{\sigma}^{2U}/n_i}, \tag{3.1}$$

the two kinds of the empirical Bayes confidence intervals

$$I_i^{EB} : \ \widehat{\mu}_i^{EB}(\widehat{\psi}) \pm z_{\alpha/2}\sqrt{(\hat{\sigma}^2/n_i)(1-\hat{\gamma}_i)},$$
$$I_i^{EB*} : \ \widehat{\mu}_i^{EB}(\widehat{\psi}) \pm z_{\alpha/2}\sqrt{mse_i(\hat{\sigma}^2, \widehat{\psi})},$$

and the corrected empirical Bayes confidence interval

$$I_i^{CEB*} : \ \widehat{\mu}_i^{EB}(\widehat{\psi}) \pm z_{\alpha/2}\left[1 + (2k)^{-1}h_i^*(\widehat{\psi})\right]\sqrt{mse_i(\hat{\sigma}^2, \widehat{\psi})}, \tag{3.2}$$

where the truncated Prasad-Rao estimator $(\widehat{\psi}^{TR}, \hat{\sigma}^{2U})$ is used for (ψ, σ^2). Here $t_{\alpha/2}$ is the $\alpha/2$ upper quantile of a t-distribution with $(N-k-r_1)$-degrees of freedom, and $\hat{\sigma}^{2U}$ and $\widehat{\psi}^{TR}$ are given in (2.12) and (2.13).

The simulation experiments are carried out under model (1.1) or (2.1) without covariates for $k = 20$. In this case, $\boldsymbol{x}_{ij}'\boldsymbol{\beta} = \mu$ (i.e., $\boldsymbol{x}_{ij} = 1$), and we can put $\mu = 0$ without any loss of generality since the confidence intervals are translation-invariant. The sample sizes n_i's are given as $n_1 = \cdots = n_5 = 2$, $n_6 = \cdots = n_{10} = 4$, $n_{11} = \cdots = n_{15} = 6$ and $n_{16} = \cdots = n_{20} = 8$. The total number of the sample sizes is $N = \sum_{i=1}^{20} n_i = 100$. We handle the cases that $\sigma^2 = 1$ and $\psi = \sigma_v^2$ takes the values from 0 to 2. We generate 10,000 random sets of the response variables $\boldsymbol{y} = (y_{11}, \ldots, y_{1,n_1}; \ldots; y_{k1}, \ldots, y_{k,n_k})'$ based on model (1.1), and the frequency of the confidence interval which includes the mean μ_i

is counted for $i = 1, \ldots, k$. The coverage probability is estimated by dividing the total number of the frequency by $10,000$. The expected width of the confidence interval can be also estimated by a similar method.

The coverage probabilities and the expected widths of the confidence intervals I_i^T, I_i^{EB}, I_i^{EB*} and I_i^{CEB*} are obtained through the above simulation for each area $i = 1, \ldots, k$. The average values of the coverage probabilities over the total k areas for $k = 20$ are illustrated in Figure 1, where the confidence coefficient is $1 - \alpha = 0.95$, and the x-axis denotes the value of ψ. The average values of the expected widths of the confidence intervals over the total k areas are illustrated in Figure 2. From Figures 1 and 2, we can observe the following:

(1) The corrected empirical Bayes confidence interval I_i^{CEB*} satisfies the nominal confidence level, while the empirical Bayes confidence intervals I_i^{EB} and I_i^{EB*} violate the confidence level for $\psi > 0.2$ as seen from Figure 1. Since I_i^{CEB*} has a larger expected width than I_i^{EB} and I_i^{EB*} as illustrated in Figure 2, it is seen that I_i^{CEB*} extends the widths of I_i^{EB} and I_i^{EB*} so as to satisfy the nominal confidence level.

(2) The expected width of I_i^{CEB*} is much smaller than that of I_i^T, while both the confidene intervals satisfy the nominal confidence level. This means that I_i^{CEB*} improves on I_i^T.

(3) As seen from the forms of the confidence intervals given in (1.7), (1.8), (2.9) and (2.10), the confidence intervals become instable when $\widehat{\psi}$ is close to zero. Thus we need to use the truncated estimator like (2.13), but in the case of small k, such a truncation affects the performances of I_i^{CEB*}, I_i^{EB} and I_i^{EB*} for small values of ψ. This is the reason that the coverage probablities of I_i^{CEB*}, I_i^{EB} and I_i^{EB*} are over 0.95 for $\psi < 0.2$ in Figure 1.

(4) We have investigated the performances of the corrected empirical Bayes confidence interval I_i^{CEB} and have found that the performances of I_i^{CEB} are quite similar to those of I_i^{CEB*}, though the numerical results are omitted here.

We thus conclude that the corrected empirical Bayes confidence intervals I_i^{CEB} and I_i^{CEB*} not only satisfy the nominal confidence level by extending the widths of I_i^{EB} and I_i^{EB*}, but also are superior to the conventional interval I_i^T in terms of the coverage probability and the expected width.

3.2 Example: *Posted Land Price Data*

We here apply the proposed confidence interval to the posted land price data along the Keikyu train line. This train line connects the suburbs in the Kanagawa prefecture to the Tokyo metropolitan area. Those who live in the suburbs in the Kanagawa prefecture take this line to work or study in Tokyo everyday. Thus, it is expected that the land price depends on the distance from Tokyo.

The posted land price data are available for 48 stations on the Keikyu train line, and we consider each station as a small area, namely, $k = 48$. For the i-th station, data of n_i land spots are available in 2001. For $j = 1, \ldots, n_i$, y_{ij} denotes the value which is transformed by logarithm from the posted land price (Yen) for the unit meter squares of the j-th spot, T_i is the time to take from the nearby station i to the Tokyo station around

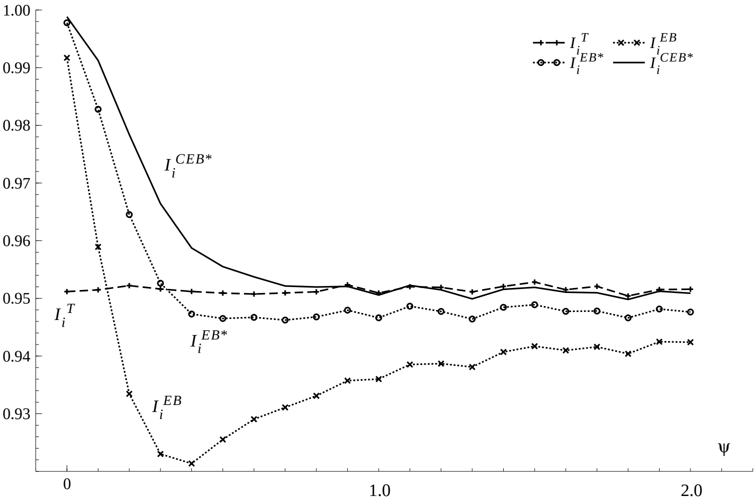

Figure 1: Coverage probabilities of I_i^T, I_i^{EB}, I_i^{EB*} and I_i^{CEB*} for $k = 20$ (The x-axix denotes the value of ψ from 0 to 2)

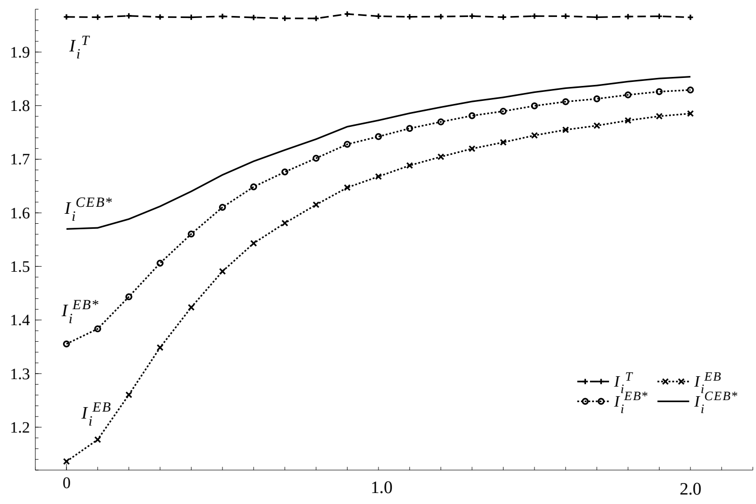

Figure 2: Expected widths of I_i^T, I_i^{EB}, I_i^{EB*} and I_i^{CEB*} for $k = 20$ (The x-axix denotes the value of ψ from 0 to 2)

8:30 in the morning, D_{ij} is the geographical distance from the spot j to the station i and FAR_{ij} denotes the floor-area ratio, or ratio of building volume to lot area of the spot j. Using the Akaike information criterion, Kubokawa and Srivastava (2007) selected the regressor variables of and proposed the nested error regression model

$$y_{ij} = \beta_0 + FAR_{ij}\beta_1 + T_i\beta_2 + (T_i^2)\beta_3 + D_{ij}\beta_4 + v_i + e_{ij}. \tag{3.3}$$

Then, the estimates of the parameters are given by $\hat{\sigma}^2 = 0.020803$, $\widehat{\psi}^{TR} = 0.406572$, $\hat{\sigma}_v^2 = \widehat{\psi}^{TR} \times \hat{\sigma}^2 = 0.008458$ and

$$\widehat{\boldsymbol{\beta}}(\widehat{\psi}^{TR}) = (\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_2, \widehat{\beta}_3, \widehat{\beta}_4) = (13.448, 0.0010105, -0.032302, 0.00018892, -6.0411\times 10^{-5}).$$

It is interesting to note that the land price decreases through the quadratic function $f(T_1)$ of the time T_1, namely, $f(T_1) = \beta_0 + \beta_2 T_1 + \beta_3 T_1^2 = 13.448 - 0.032302T_1 + 0.00018892T_1^2$.

Now we give the confidence intervals of the average land price for the unit meter squares at the i-th station, namely, $\mu_i = \beta_0 + \overline{FAR}_i\beta_1 + T_i\beta_2 + (T_i^2)\beta_3 + \overline{D}_i\beta_4 + v_i$ for $i = 1, \ldots, 48$, where $\overline{FAR}_i = \sum_{j=1}^{n_i} FAR_{ij}/n_i$ and $\overline{D}_i = \sum_{j=1}^{n_i} D_{ij}/n_i$. The upper and lower bounds of the confidence interval I_i^T and the corrected empirical Bayes confidence interval I_i^{CEB} based on the Prasad-Rao estimators are computed by (3.1) and (3.2), and those values transformed by exponential are plotted in Figure 3 for $i = 1, \ldots, 48$, where the X-axis denotes the stations from No.1 to No.48, namely No.1 is the station closest to the Tokyo station and No.48 is the station farthest from Tokyo. The widths of the confidence intervals are also plotted in Figure 4. The values of n_i are indicated in Figures 3 and 4 with a different scale. It is revealed that I_i^{CEB} is more stabilized and shorter than I_i^T for smaller n_i's and that the confidence intervals have the general pattern of decrease in i on the X-axis.

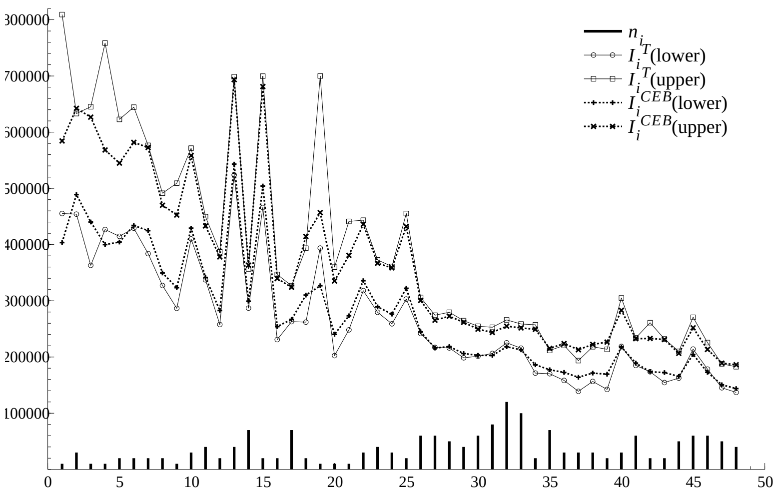

Figure 3: Confidence intervals of the means based on I_i^T and I_i^{CEB} for $i = 1, \ldots, 48$

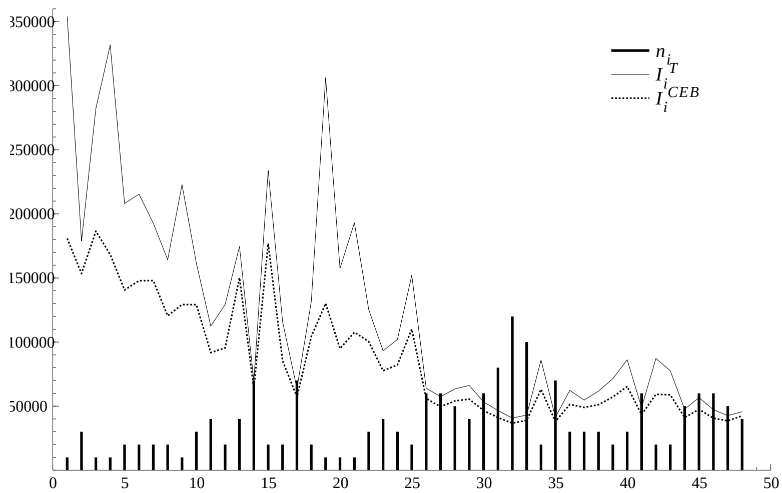

Figure 4: Widths of the confidence intervals I_i^T and I_i^{CEB} for $i = 1, \dots, 48$

4 Concluding Remarks

In this paper, we have obtained the asymptotically corrected empirical Bayes confidence intervals whose coverage probabilities can satisfy the confidence level in the second order asymptotics. This is an extension of the results of Datta *et al.* (2002) and Basu *et al.* (2003) to the nested error regression model and to the finite population model with unbalanced sample sizes and unknown components of variance. The corrected confidence intervals have been numerically shown to be superior to the conventional confidence interval based on the sample mean in terms of the coverage probability and the expected width of the interval. The usefulness has also been shown through the application to the posted land price data in Tokyo and the neighboring prefecture.

5 Appendix

We shall prove the theorems and propositions given in Section 2.

Proof of Theorems 2.1 and 2.2. We begin with proving Theorem 2.2 for $i = 1$, namely, $P[\overline{Y}_1 \in I_1^{CEB}] = 1 - \alpha + O(k^{-3/2})$ as $k \to \infty$. Since

$$P[\overline{Y}_1 \in I_1^{CEB}] = \sum_{\boldsymbol{s} \in \boldsymbol{S}} P[\overline{Y}_1 \in I_1^{CEB}(\widehat{\psi}, \hat{\sigma}^2, \boldsymbol{s})|s_i]P[\boldsymbol{s}],$$

it is sufficient to show that $P[\overline{Y}_1 \in I_1^{CEB}(\widehat{\psi}, \hat{\sigma}^2, \boldsymbol{s})|\boldsymbol{s}] = 1 - \alpha + O(k^{-3/2})$. Since the conditional distribution of $\overline{Y}_1$ given $\overline{y}_1$ is

$$\overline{Y}_1|\overline{y}_1 \sim \mathcal{N}(\widehat{\mu}_1^B(\boldsymbol{\beta}, \psi, s_1), \sigma^2(f_1/n_1)(1 - f_1\gamma_1))$$

for $\gamma_1 = 1/(1 + n_1\psi)$ and $f_1 = (N_1 - n_1)/N_1$. For the notational convenience, we omit the condition $\boldsymbol{s}$ as $P[\cdot|\boldsymbol{s}] = P[\cdot]$. Then, the coverage probability of $I_1^{CEB}(\widehat{\psi}, \hat{\sigma}^2, \boldsymbol{s})$ can be written as

$$P[\overline{Y}_1 \in I_1^{CEB}(\widehat{\psi}, \hat{\sigma}^2, \boldsymbol{s})] = P\left[-z + G(-z) < \frac{\overline{Y}_1 - \widehat{\mu}_1^B(\boldsymbol{\beta}, \psi, s_1)}{\sqrt{\sigma^2 (f_1/n_1)(1 - f_1\gamma_1)}} < z + G(z)\right]$$
$$= E[\Phi(z + G(z)) - \Phi(-z + G(-z))], \tag{5.1}$$

where

$$\begin{aligned} G(z) =& \frac{\widehat{\mu}_1^{EB}(\widehat{\psi}, \boldsymbol{s}) - \widehat{\mu}_1^B(\boldsymbol{\beta}, \psi, s_1) + z[1 + h_1(\widehat{\psi}, \boldsymbol{s})/(2k)]\sqrt{\hat{\sigma}^2 (f_1/n_1)(1 - f_1\hat{\gamma}_1)}}{\sqrt{\sigma^2 (f_1/n_1)(1 - f_1\gamma_1)}} - z \\ =& \frac{\boldsymbol{d}_1'(\widehat{\boldsymbol{\beta}}(\widehat{\psi}) - \boldsymbol{\beta}) - f_1(\hat{\gamma}_1 - \gamma_1)(\overline{y}_1 - \overline{\boldsymbol{x}}_1'\boldsymbol{\beta}) + f_1(\hat{\gamma}_1 - \gamma_1)\overline{\boldsymbol{x}}_1'(\widehat{\boldsymbol{\beta}}(\widehat{\psi}) - \boldsymbol{\beta})}{\sqrt{\sigma^2 (f_1/n_1)(1 - f_1\gamma_1)}} \\ &+ z\left[1 + h_1(\widehat{\psi}, \boldsymbol{s})/(2k)\right] \frac{\sqrt{\hat{\sigma}^2 (f_1/n_1)(1 - f_1\hat{\gamma}_1)}}{\sqrt{\sigma^2 (f_1/n_1)(1 - f_1\gamma_1)}} - z, \end{aligned}$$

for $\boldsymbol{d}_1 = \overline{\boldsymbol{x}}_{1(p)} - (1 - f_1\gamma_1)\overline{\boldsymbol{x}}_1$. We begin with approximating $G(z)$ up to order $O_p(k^{-1})$. Note that

$$\hat{\gamma}_1 = \gamma_1 - n_1\gamma_1^2(\widehat{\psi} - \psi) + n_1^2\gamma_1^3(\widehat{\psi} - \psi)^2 + O_p(k^{-3/2}),$$

which yields

$$\begin{aligned} &\frac{\boldsymbol{d}_1'(\widehat{\boldsymbol{\beta}}(\widehat{\psi}) - \boldsymbol{\beta}) - f_1(\hat{\gamma}_1 - \gamma_1)(\overline{y}_1 - \overline{\boldsymbol{x}}_1'\boldsymbol{\beta}) + f_1(\hat{\gamma}_1 - \gamma_1)\overline{\boldsymbol{x}}_1'(\widehat{\boldsymbol{\beta}}(\widehat{\psi}) - \boldsymbol{\beta})}{\sqrt{\sigma^2 (f_1/n_1)(1 - f_1\gamma_1)}} \\ &\quad = \frac{\hat{u}_1}{\sqrt{k}} + \frac{\hat{u}_2}{k} + O_p(k^{-3/2}), \end{aligned} \tag{5.2}$$

where

$$\begin{aligned} \hat{u}_1 =& \frac{\sqrt{k}}{\sqrt{\sigma^2 (f_1/n_1)(1 - f_1\gamma_1)}} \left[\boldsymbol{d}_1'(\widehat{\boldsymbol{\beta}}(\widehat{\psi}) - \boldsymbol{\beta}) + f_1 n_1 \gamma_1^2 (\widehat{\psi} - \psi)(\overline{y}_1 - \overline{\boldsymbol{x}}_1'\boldsymbol{\beta})\right], \\ \hat{u}_2 =& -\frac{k f_1 n_1 \gamma_1^2}{\sqrt{\sigma^2 (f_1/n_1)(1 - f_1\gamma_1)}} \left[(\widehat{\psi} - \psi)\overline{\boldsymbol{x}}_1'(\widehat{\boldsymbol{\beta}}(\widehat{\psi}) - \boldsymbol{\beta}) + n_1\gamma_1(\widehat{\psi} - \psi)^2(\overline{y}_1 - \overline{\boldsymbol{x}}_1'\boldsymbol{\beta})\right]. \end{aligned}$$

It is also noted that

$$\begin{aligned} \frac{\sqrt{1 - f_1\hat{\gamma}_1}}{\sqrt{1 - f_1\gamma_1}} =& 1 + \frac{f_1 n_1 \gamma_1^2}{2(1 - f_1\gamma_1)}(\widehat{\psi} - \psi) \\ &- \frac{f_1^2 n_1^2 \gamma_1^4 + 4 f_1 n_1^2 \gamma_1^3 (1 - f_1\gamma_1)}{8(1 - f_1\gamma_1)^2}(\widehat{\psi} - \psi)^2 + O_p(k^{-3/2}), \\ \frac{\sqrt{\hat{\sigma}^2}}{\sqrt{\sigma^2}} =& 1 + \frac{1}{2\sigma^2}(\hat{\sigma}^2 - \sigma^2) - \frac{1}{8\sigma^4}(\hat{\sigma}^2 - \sigma^2)^2 + O_p(k^{-3/2}), \end{aligned}$$

which yield that

$$\left[1 + h_1(\widehat{\psi}, \boldsymbol{s})/(2k)\right] \frac{\sqrt{\hat{\sigma}^2 (1 - f_1\hat{\gamma}_1)}}{\sqrt{\sigma^2 (1 - f_1\gamma_1)}} - 1 = \frac{\hat{v}_1}{\sqrt{k}} + \frac{\hat{v}_2}{k} + \frac{h_1(\widehat{\psi})}{2k} + O_p(k^{-3/2}), \tag{5.3}$$

where

$$\hat{v}_1 = \frac{\sqrt{k}}{2} H_1,$$
$$\hat{v}_2 = -\frac{k f_1 n_1^2 \gamma_1^3}{2(1 - f_1\gamma_1)} (\widehat{\psi} - \psi)^2 - \frac{k}{8} K_1^2,$$

for $H_1 = a_1(\widehat{\psi} - \psi)/\psi + (\hat{\sigma}^2 - \sigma^2)/\sigma^2$ and $K_1 = a_1(\widehat{\psi} - \psi)/\psi - (\hat{\sigma}^2 - \sigma^2)/\sigma^2$ for $a_1 = f_1 n_1 \gamma_1^2 \psi/(1 - f_1\gamma_1)$. Combining (5.2) and (5.3) gives the approximation

$$G(z) = \frac{\hat{u}_1}{\sqrt{k}} + \frac{\hat{u}_2}{k} + z\left\{\frac{\hat{v}_1}{\sqrt{k}} + \frac{\hat{v}_2}{k} + \frac{h_1(\widehat{\psi}, \boldsymbol{s})}{2k}\right\} + O_p(k^{-3/2}). \tag{5.4}$$

Since $G(z) = O_p(k^{-1/2})$, $\Phi(z + G(z))$ is evaluated as

$$\begin{aligned}\Phi(z + G(z)) =& \Phi(z) + G(z)\phi(z) + \frac{G^2(z)}{2}\phi'(z) + \frac{1}{2}\int_z^{z+G(z)} (z + G(z) - x)^2 \phi''(x)\mathrm{d}x \\ =& \Phi(z) + \left\{G(z) - zG^2(z)/2\right\}\phi(z) + O_p(k^{-3/2}),\end{aligned}$$

so that from (5.1),

$$\begin{aligned}&P[\overline{Y}_1 \in I_1^{CEB}(\widehat{\psi}, \hat{\sigma}^2, s_1)] \\ &\quad= \Phi(z) - \Phi(-z) + \phi(z)E\left[G(z) - G(-z) - (z/2)\{G(z)^2 + G(-z)^2\}\right] + O(k^{-3/2}) \\ &\quad= 1 - \alpha + \phi(z)E\left[G(z) - G(-z) - (z/2)\{G(z)^2 + G(-z)^2\}\right] + O(k^{-3/2}).\end{aligned} \tag{5.5}$$

From (5.4), it is seen that

$$\begin{aligned}&E\big[G(z) - G(-z) - (z/2)\{G(z)^2 + G(-z)^2\}\big] \\ &= E\left[2z\frac{\hat{v}_1}{\sqrt{k}} + \frac{z}{k}[2\hat{v}_2 + h_1(\widehat{\psi}, \boldsymbol{s})] - \frac{z}{k}\hat{u}_1^2 - \frac{z^3}{k}\hat{v}_1^2\right] \\ &= \frac{z}{k}\left\{E[h_1(\widehat{\psi}, \boldsymbol{s})] - \left[-kE[H_1] + E[-2\hat{v}_2 + \hat{u}_1^2 + z^2\hat{v}_1^2]\right]\right\}. \\ &= \frac{z}{k}\left\{E[h_1(\widehat{\psi}, \boldsymbol{s})] - E\left[\hat{u}_1^2 + \frac{k f_1 n_1^2 \gamma_1^3}{1 - f_1\gamma_1}(\widehat{\psi} - \psi)^2 - kH_1 + kK_1^2/4 + kz^2H_1^2/4\right]\right\},\end{aligned} \tag{5.6}$$

from the definitions of $\hat{v}_1$ and $\hat{v}_2$. Hence, we need to evaluate the term $E[\hat{u}_1^2]$, which is written as

$$E[\hat{u}_1^2] = \frac{n_1 k}{\sigma^2 f_1(1 - f_1\gamma_1)} E\left[\left\{\boldsymbol{d}_1'(\widehat{\boldsymbol{\beta}}(\widehat{\psi}) - \boldsymbol{\beta}) + f_1 n_1 \gamma_1^2(\widehat{\psi} - \psi)U_1\right\}^2\right], \tag{5.7}$$

where $U_1 = \overline{y}_1 - \overline{\boldsymbol{x}}_1'\boldsymbol{\beta}$. The Taylor expansion with respect to $\widehat{\psi}$ at ψ gives the expression

$$\widehat{\boldsymbol{\beta}}(\widehat{\psi}) = \widehat{\boldsymbol{\beta}}(\psi) + \widehat{\boldsymbol{\beta}}^{(1)}(\psi)(\widehat{\psi} - \psi) + \frac{\widehat{\boldsymbol{\beta}}^{(2)}(\psi)}{2}(\widehat{\psi} - \psi)^2 + O_p(k^{-3/2}),$$

where $\widehat{\boldsymbol{\beta}}^{(1)}(\psi) = (\partial/\partial\psi)\widehat{\boldsymbol{\beta}}(\psi)$ and $\widehat{\boldsymbol{\beta}}^{(2)}(\psi) = (\partial^2/\partial\psi^2)\widehat{\boldsymbol{\beta}}(\psi)$. It can be seen that $\widehat{\boldsymbol{\beta}}(\psi) - \boldsymbol{\beta} = O_p(k^{-1/2})$, $\widehat{\boldsymbol{\beta}}^{(1)}(\psi) = O_p(k^{-1/2})$ and $\widehat{\boldsymbol{\beta}}^{(2)}(\psi) = O_p(k^{-1/2})$. Thus, it is observed that $\widehat{\boldsymbol{\beta}}(\widehat{\psi}) - \widehat{\boldsymbol{\beta}}(\psi) = \widehat{\boldsymbol{\beta}}^{(1)}(\psi)(\widehat{\psi} - \psi) + O_p(k^{-3/2})$, and

$$
\begin{aligned}
&E\big[\{\boldsymbol{d}_1'(\widehat{\boldsymbol{\beta}}(\widehat{\psi}) - \boldsymbol{\beta}) + f_1 n_1 \gamma_1^2(\widehat{\psi} - \psi)U_1\}^2\big] \\
&\quad= E\big[\{\boldsymbol{d}_1'(\widehat{\boldsymbol{\beta}}(\psi) - \boldsymbol{\beta}) + \boldsymbol{d}_1'\widehat{\boldsymbol{\beta}}^{(1)}(\psi)(\widehat{\psi} - \psi) + f_1 n_1 \gamma_1^2(\widehat{\psi} - \psi)U_1\}^2\big] + O(k^{-3/2}) \\
&\quad= E\big[\boldsymbol{d}_1'(\widehat{\boldsymbol{\beta}}(\psi) - \boldsymbol{\beta})(\widehat{\boldsymbol{\beta}}(\psi) - \boldsymbol{\beta})'\boldsymbol{d}_1\big] + f_1^2 n_1^2 \gamma_1^4 E\big[(\widehat{\psi} - \psi)^2 U_1^2\big] \\
&\qquad + 2 f_1 n_1 \gamma_1^2 E\big[\boldsymbol{d}_1'(\widehat{\boldsymbol{\beta}}(\psi) - \boldsymbol{\beta})(\widehat{\psi} - \psi)U_1\big] + O(k^{-3/2}) \\
&\quad= I_1 + I_2 + I_3 + O(k^{-3/2}). \quad \text{(say)}
\end{aligned}
$$

It is easy to see that $I_1 = \boldsymbol{d}_1'(\boldsymbol{A}(\psi) + \boldsymbol{B})^{-1}\boldsymbol{d}_1\sigma^2$. To evaluate I_2 and I_3, the following Stein identity is useful. Note that $\overline{y}_1 \sim \mathcal{N}(\overline{\boldsymbol{x}}_1'\boldsymbol{\beta}, \sigma^2/(n_1\gamma_1))$. For an absolutely continuous function $g(\overline{y}_1)$, Stein (1981) showed that

$$
E[g(\overline{y}_1)(\overline{y}_1 - \overline{\boldsymbol{x}}_1'\boldsymbol{\beta})] = \frac{\sigma^2}{n_1\gamma_1} E\left[\frac{\partial}{\partial \overline{y}_1} g(\overline{y}_1)\right],
$$

which is called the Stein identity. Using the Stein identity, we observe that

$$
\begin{aligned}
E\big[\boldsymbol{d}_1'(\widehat{\boldsymbol{\beta}}(\psi) - \boldsymbol{\beta})(\widehat{\psi} - \psi)U_1\big] &= \frac{\sigma^2}{n_1\gamma_1} E\left[\frac{\partial}{\partial \overline{y}_1}\left\{\boldsymbol{d}_1'(\widehat{\boldsymbol{\beta}}(\psi) - \boldsymbol{\beta})(\widehat{\psi} - \psi)\right\}\right] \\
&= \frac{\sigma^2}{n_1\gamma_1} E\left[\boldsymbol{d}_1'\left\{\frac{\partial}{\partial \overline{y}_1}\widehat{\boldsymbol{\beta}}(\psi)\right\}(\widehat{\psi} - \psi) + \boldsymbol{d}_1'(\widehat{\boldsymbol{\beta}}(\psi) - \boldsymbol{\beta})\frac{\partial}{\partial \overline{y}_1}\widehat{\psi}\right].
\end{aligned}
$$

Since $\{y_{ij} - \overline{y}_i, i = 1, \ldots, k,\ j = 1, \ldots, n_i\}$ are independent of $\{\overline{y}_1, \ldots, \overline{y}_k\}$, we can see that $\partial\widehat{\boldsymbol{\beta}}(\psi)/\partial\overline{y}_1 = (\boldsymbol{A}(\psi) + \boldsymbol{B})^{-1} n_1 \gamma_1(\psi)\overline{\boldsymbol{x}}_1$, which is $O_p(k^{-1})$. From the assumption of the theorem, $\partial\widehat{\psi}/\partial\overline{y}_1 = O_p(k^{-1})$. These imply that $I_3 = 0 + O(k^{-3/2})$. For I_2, the Stein identity is used to get that

$$
\begin{aligned}
E\big[(\widehat{\psi} - \psi)^2 U_1^2\big] &= \frac{\sigma^2}{n_1\gamma_1} E\left[\frac{\partial}{\partial \overline{y}_1}\left\{(\widehat{\psi} - \psi)^2 U_1\right\}\right] \\
&= \frac{\sigma^2}{n_1\gamma_1} E\left[2(\widehat{\psi} - \psi)U_1\left\{\frac{\partial\widehat{\psi}}{\partial \overline{y}_1}\right\} + (\widehat{\psi} - \psi)^2\right] \\
&= \frac{\sigma^2}{n_1\gamma_1} E[(\widehat{\psi} - \psi)^2] + O(k^{-3/2}),
\end{aligned}
$$

so that $I_2 = \sigma^2 f_1^2 n_1 \gamma_1^3 E[(\widehat{\psi} - \psi)^2] + O(k^{-3/2})$. Hence from (5.7), we obtain that

$$
E[\hat{u}_1^2] = \frac{n_1 k}{f_1(1 - f_1\gamma_1)}\left\{\boldsymbol{d}_1'(\boldsymbol{A}(\psi) + \boldsymbol{B})^{-1}\boldsymbol{d}_1 + f_1^2 n_1 \gamma_1^3 E[(\widehat{\psi} - \psi)^2]\right\} + O(k^{-3/2}).
$$

From (5.6), we get that

$$
\begin{aligned}
&E\left[\hat{u}_1^2+\frac{kf_1n_1^2\gamma_1^3}{1-f_1\gamma_1}(\widehat{\psi}-\psi)^2-kH_1+kK_1^2/4+kz^2H_1^2/4\right]\\
=&\frac{n_1k}{f_1(1-f_1\gamma_1)}\left\{\boldsymbol{d}_1'(\boldsymbol{A}(\psi)+\boldsymbol{B})^{-1}\boldsymbol{d}_1+2f_1^2n_1\gamma_1^3E[(\widehat{\psi}-\psi)^2]\right\}\\
&-kE[H_1]+\frac{k}{4}E[z^2H_1^2+K_1^2]+O(k^{-1/2}),
\end{aligned}
$$

which is equal to $h_1(\psi)$ given by (2.3). Hence from (5.6), it is seen that

$$
\begin{aligned}
&E\big[G(z)-G(-z)-(z/2)\{G(z)^2+G(-z)^2\}\big]\\
=&\frac{z}{k}\left\{E[h_1(\widehat{\psi},\boldsymbol{s})]-E\left[\hat{u}_1^2+\frac{kf_1n_1^2\gamma_1^3}{1-f_1\gamma_1}(\widehat{\psi}-\psi)^2-kH_1+kK_1^2/4+kz^2H_1^2/4\right]\right\}\\
=&O(k^{-3/2}).
\end{aligned}
$$

From (5.5), we thus conclude that $P[\mu_1\in I_1^{CEB}]=1-\alpha+O(k^{-3/2})$, and the proof of Theorem 2.2 is complete.

Since Theorem 2.1 corresponds to Theorem 2.2 with $f_i=1$, $a_i=\gamma_i$ and $\boldsymbol{d}_i=\overline{\boldsymbol{x}}_i$, the same arguments can prove Theorem 2.1. ■

Proof of Propositions 2.1 and 2.5. Since Proposition 2.1 corresponds to Proposition 2.5 with $f_i=1$ and $a_i=\gamma_i$, it is enough to show Proposition 2.5. Note that $\widehat{\psi}$ is approximated as

$$
\widehat{\psi}=\frac{\hat{\sigma}_v^2}{\hat{\sigma}^2}=\hat{\sigma}_v^2\left\{\frac{1}{\sigma^2}-\frac{\hat{\sigma}^2-\sigma^2}{\sigma^4}+\frac{(\hat{\sigma}^2-\sigma^2)^2}{\sigma^6}\right\}+O_p(k^{-3/2}),
$$

or

$$
(\widehat{\psi}-\psi)/\psi=T_v-T_e-T_vT_e+T_e^2+O_p(k^{-3/2}).
$$

Thus, the moments are approximated as

$$
\begin{aligned}
E[(\widehat{\psi}-\psi)^2]=&\psi^2E\left[T_v^2-2T_vT_e+T_e^2\right]+O(k^{-3/2}),\\
E\left[H_i\right]=&E\left[a_iT_v+(1-a_i)T_e\right]+a_iE\left[T_e^2-T_vT_e\right]+O(k^{-3/2}),\\
E\left[H_i^2\right]=&E\left[a_i^2T_v^2+2a_i(1-a_i)T_vT_e+(1-a_i)^2T_e^2\right]+O(k^{-3/2}),\\
E\left[K_i^2\right]=&E\left[a_i^2T_v^2-2a_i(1+a_i)T_vT_e+(1+a_i)^2T_e^2\right]+O(k^{-3/2}).
\end{aligned}
\tag{5.8}
$$

Using the notations $\tau_1(\psi)$, $\tau_2(\psi)$ and $\tau_3(\psi)$ defined in Proposition 2.1, we can see that $E[(\widehat{\psi}-\psi)^2]=\psi^2\tau_1(\psi)+O(k^{-3/2})$, $E[H_i^2]=a_i^2\tau_1(\psi)-2a_i\tau_2(\psi)+\tau_3(\psi)+O(k^{-3/2})$, $E[K_i^2]=a_i^2\tau_1(\psi)+2a_i\tau_2(\psi)+\tau_3(\psi)+O(k^{-3/2})$. Substituting these terms into (2.17), we can get expression (2.19). ■

Proof of Propositions 2.3 and 2.6. We begin with proving Proposition 2.6, where the given sample $\boldsymbol{s}$ in $\widehat{\mu}_i^{EB}(\widehat{\psi},\boldsymbol{s})$ is omitted here for the notational convenience. Following

Pradad and Rao (1990) and Datta and Lahiri (2000), the MSE of $\widehat{\mu}_i^{EB}(\widehat{\psi})$ can be written as

$$
\begin{aligned}
MSE_i(\sigma^2, \psi) = & E[\{\widehat{\mu}_i^{EB}(\widehat{\psi}) - \overline{Y}_i\}^2] \\
= & E[\{\widehat{\mu}_i^{B}(\boldsymbol{\beta}, \psi) - \overline{Y}_i\}^2] + E[\{\widehat{\mu}_i^{EB}(\psi) - \widehat{\mu}_i^{B}(\boldsymbol{\beta}, \psi)\}^2] + E[\{\widehat{\mu}_i^{EB}(\widehat{\psi}) - \widehat{\mu}_i^{EB}(\psi)\}^2] \\
= & \frac{\sigma^2}{n_i} f_i(1 - f_i\gamma_i) + \sigma^2 \boldsymbol{d}_i'(\boldsymbol{A}(\psi) + \boldsymbol{B})^{-1}\boldsymbol{d}_i + g_{3i}(\sigma^2, \psi),
\end{aligned}
$$

where $\boldsymbol{d}_i = f_i\{\overline{\boldsymbol{x}}_i^* - (1 - \gamma_i)\overline{\boldsymbol{x}}_i\} = \overline{\boldsymbol{x}}_{i(p)} - (1 - f_1\gamma_i)\overline{\boldsymbol{x}}_i$ and $g_{3i}(\sigma^2, \psi) = E[\{\widehat{\mu}_i^{EB}(\widehat{\psi}) - \widehat{\mu}_i^{EB}(\psi)\}^2]$. From the Taylor expansion, it follows that

$$
g_{3i}(\sigma^2, \psi) = E[\{\partial\widehat{\mu}_i^{EB}(\psi)/\partial\psi\}^2(\widehat{\psi} - \psi)^2] + O(k^{-3/2}),
$$

where $\partial\widehat{\mu}_i^{EB}(\psi)/\partial\psi = f_i n_i \gamma_i^2(\overline{y}_i - \overline{\boldsymbol{x}}_i'\widehat{\boldsymbol{\beta}}(\psi)) + \boldsymbol{d}_i'\{\partial\widehat{\boldsymbol{\beta}}(\psi)/\partial\psi\}$. It can be observed that $\{\partial\widehat{\boldsymbol{\beta}}(\psi)/\partial\psi\} = O_p(k^{-1/2})$ and $\widehat{\boldsymbol{\beta}}(\psi) - \boldsymbol{\beta} = O_p(k^{-1/2})$, so that g_{3i} is approximated as

$$
g_{3i}(\sigma^2, \psi) = f_i^2 n_i^2 \gamma_i^4 E[(\overline{y}_i - \overline{\boldsymbol{x}}_i'\boldsymbol{\beta})^2(\widehat{\psi} - \psi)^2] + O(k^{-3/2}).
$$

From the Stein identity used in the above proof, g_{3i} can be evaluated as

$$
\begin{aligned}
g_{3i}(\sigma^2, \psi) = & \sigma^2 f_i^2 n_i \gamma_i^3 E[\frac{\partial}{\partial \overline{y}_i}\{(\overline{y}_i - \overline{\boldsymbol{x}}_i'\boldsymbol{\beta})(\widehat{\psi} - \psi)^2\}] + O(k^{-3/2}) \\
= & \sigma^2 f_i^2 n_i \gamma_i^3 E[(\widehat{\psi} - \psi)^2 + 2(\overline{y}_i - \overline{\boldsymbol{x}}_i'\boldsymbol{\beta})(\widehat{\psi} - \psi)\frac{\partial\widehat{\psi}}{\partial \overline{y}_i}] + O(k^{-3/2}) \\
= & \sigma^2 f_i^2 n_i \gamma_i^3 E[(\widehat{\psi} - \psi)^2] + O(k^{-3/2}),
\end{aligned}
$$

since $\partial\widehat{\psi}/\partial\overline{y}_i = O_p(k^{-1})$. From (5.8), it is also expressed as $g_{3i}(\sigma^2, \psi) = \sigma^2 f_i^2 n_i \gamma_i^3 \psi^2 E[(T_e - T_v)^2] + O(k^{-3/2}) = \sigma^2 f_i^2 n_i \gamma_i^3 \psi^2 \tau_1(\psi) + O(k^{-3/2})$, and we get the expression given in (2.7).

For the asymptotically unbiased estimator of the MSE, it is noted that

$$
\begin{aligned}
E[\frac{\hat{\sigma}^2}{n_i} f_i(1 - f_i\hat{\gamma}_i)] = & \frac{\sigma^2}{n_i} f_i(1 - f_i\gamma_i) + f_i\frac{1 - f_i\gamma_i}{n_i} E[\hat{\sigma}^2 - \sigma^2] + \sigma^2 f_i^2 \gamma_i^2 E[\widehat{\psi} - \psi] \\
& + f_i^2\gamma_i^2 E[(\hat{\sigma}^2 - \sigma^2)(\widehat{\psi} - \psi)] - \sigma^2 f_i^2 n_i \gamma_i^3 E[(\widehat{\psi} - \psi)^2] + O(k^{-3/2}) \\
= & \frac{\sigma^2}{n_i} f_i(1 - f_i\gamma_i)\big\{1 + (1 - a_i)b_e(\psi) + a_i b_v(\psi)\big\} - g_{3i}(\psi) + O(k^{-3/2}).
\end{aligned}
$$

Combining this fact and (2.20), we can get the second-order unbiased estimator given in (2.21).

Since Proposition 2.3 corresponds to Proposition 2.6 with $f_i = 1$ and $\boldsymbol{d}_i = \overline{\boldsymbol{x}}_i$, the same arguments can prove Proposition 2.3. ■

Acknowledgments. The author is grateful to two referees for their valuable comments and helpful suggestions. The research of the author was supported in part by a grant from the Ministry of Education, Japan, Nos. 16500172 and 21540114 and in part by a grant from the 21st Century COE Program at Faculty of Economics, University of Tokyo.

References

[1] Basu, R., Ghosh, J.K., and Mukerjee, R. (2003). Empirical Bayes prediction intervals in a normal regression model: higher order asymptotics. *Statist. Prob. Letters*, **63**, 197-203.

[2] 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.

[3] Booth, J.G. and Hobert, J.P. (1998). Standard errors of prediction in generalized linear mixed models. *J. Amer. Statist. Assoc.*, **93**, 262-272.

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

[5] Das, K., Jiang, J. and Rao, J.N.K. (2004). Mean squared error of empirical predictor. *Ann. Statist.*, **32**, 818-840.

[6] Datta, G.S., Ghosh, M., Smith, D.D. and Lahiri, P. (2002). On an asymptotic theory of conditional and unconditional coverage probabilities of empirical Bayes confidence Intervals. *Scandinavian J. Statist.*, **29**, 139-152.

[7] 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.

[8] 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.

[9] Ghosh, M. and Meeden, G. (1986). Empirical Bayes estimation in finite population sampling. *J. Amer. Statist. Assoc.*, **81**, 1058-1062.

[10] Hall, P. and Maiti, T. (2006). On parametric bootstrap methods for small area prediction. *J. Royal Statist. Soc.*, **68**, 221-238.

[11] Harville, D.A. and Jeske, D.R. (1992). Mean squared error of estimation or prediction under a general linear model. *J. Amer. Statist. Assoc.*, **87**, 724-731.

[12] Kubokawa, T. and Srivastava, M.S. (2007). Akaike information criterion for selecting variables in a nested error regression model. Discussion paper series, CIRJE-F-525, Faculty of Economics, University of Tokyo,

[13] 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.

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

[15] Rivest, L.-P. and Belmonte, E. (2000). A conditional mean squared error of small area estimators. *Survey Methodology*, **26**, 67-78.

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