ON ESTIMATION OF MEAN SQUARED ERRORS OF BENCHMARKED EMPIRICAL BAYES ESTIMATORS

Rebecca C. Steorts and Malay Ghosh

Carnegie Mellon University and University of Florida

Abstract:

We consider benchmarked empirical Bayes (EB) estimators under the basic area-level model of Fay and Herriot while requiring the standard benchmarking constraint. In this paper we determine the excess mean squared error (MSE) from constraining the estimates through benchmarking. We show that the increase due to benchmarking is $O(m^{-1})$, where m is the number of small areas. Furthermore, we find an asymptotically unbiased estimator of this MSE and compare it to the second-order approximation of the MSE of the EB estimator or equivalently of the MSE of the empirical best linear unbiased predictor (EBLUP), which was derived by Prasad and Rao (1990). Morever, using methods similar to those of Butar and Lahiri (2003), we compute a parametric bootstrap estimator of the MSE of the benchmarked EB estimator under the Fay-Herriot model and compare it to the MSE of the benchmarked EB estimator found by a second-order approximation. Finally, we illustrate our methods using SAIPE data from the U.S. Census Bureau as well as a simulation study.

Key words and phrases: Small-area, Fay-Herriot, Mean Squared Error, Empirical Bayes, Benchmarking, Parametric Bootstrap

1. Introduction

Small area estimation has become increasingly popular recently due to a growing demand for such statistics. It is well known that direct small-area estimators usually have large standard errors and coefficients of variation. In order to produce estimates for these small areas, it is necessary to borrow strength from other related areas. Accordingly, model-based estimates often differ widely from the direct estimates, especially for areas with small sample sizes. One problem that arises in practice is that the model-based estimates do not aggregate to the more reliable direct survey estimates. Agreement with the direct estimates is often a political necessity to convince legislators regarding the utility of small area estimates. The process of adjusting model-based estimates to correct this problem is known as benchmarking. Another key benefit of benchmarking is protection against model misspecification as pointed out by You, Rao and Dick (2004) and Datta, Ghosh, Steorts and Maples (2011).

In recent years, the literature on benchmarking has grown in small area estimation. Among others, Pfeffermann and Barnard (1991); You and Rao (2003); You, Rao and Dick (2004); Pfeffermann and Tiller (2006); and Ugarte, Militino and Goicoa (2009) have made an impact on the continuing development of this field. Specifically, Wang, Fuller and Qu (2008) provided a frequentist method wherein an augmented model was used to construct a best linear unbiased predictor (BLUP) that automatically satisfies the benchmarking constraint. In addition, Datta, Ghosh, Steorts and Maples (2011) developed very general benchmarked Bayes estimators, which covered most of the earlier estimators that have been motivated from either a frequentist or Bayesian perspective. Specifically, they found benchmarked Bayes estimators under the Fay and Herriot (1979) model.

Due to the fact that they borrow strength, model-based estimates typically show a substantial improvement over direct estimates in terms of mean squared error (MSE). It is of particular interest to determine how much of this advantage is lost by constraining the estimates through benchmarking. The aforementioned work of Wang, Fuller and Qu (2008) and Ugarte, Militino and Goicoa (2009) examined this question through simulation studies but did not derive any probabilistic results. They showed that the MSE of the benchmarked EB estimator was slightly larger than the MSE of the EB estimator for their simulation studies. In Section 3, we derive a second-order approximation of the MSE of the benchmarked Bayes EB estimator to show that the increase due to benchmarking is $O(m^{-1})$, where m is the number of small areas.

In this paper, we are concerned with the basic area-level model of Fay and Herriot (1979). We propose benchmarked EB estimators in Section 2. In Section 3, we derive a second-order asymptotic expansion of the MSE of the benchmarked EB estimator. In Section 4, we find an estimator of this MSE and compare it to the second-order approximation of the MSE of the EB estimator or equivalently the MSE of the EBLUP, which was derived by Prasad and Rao (1990).

Finally, in Section 5, using methods similar to those of Butar and Lahiri (2003), we compute a parametric bootstrap estimator of the mean squared error of the benchmarked EB estimator under the Fay-Herriot (1979) model and compare it to our estimators from Section 2. Section 6 contains an application based on Small Area Income and Poverty Estimation Data (SAIPE) from the U.S. Census Bureau as well as a simulation study. Some concluding remarks are made in Section 7.

2. Benchmarked Empirical Bayes Estimators

Consider the area-level random effects model

$$\hat{\theta}_i = \theta_i + e_i, \quad \theta_i = \boldsymbol{x}_i^T \boldsymbol{\beta} + u_i; \quad i = 1, \ldots, m; \tag{2.1}$$

where e_i and u_i are mutually independent with $e_i \overset{ind}{\sim} N(0, D_i)$ and $u_i \overset{iid}{\sim} N(0, \sigma_u^2)$. This model was first considered in the context of estimating income for small areas (population less than 1000) by Fay and Herriot (1979). In model 2.1, D_i are known as are the $p \times 1$ design vectors $\boldsymbol{x}_i$. However, the vector of regression coefficients $\boldsymbol{\beta}_{p\times 1}$ is unknown.

When the variance component σ_u^2 is known and $\boldsymbol{\beta}$ has a uniform prior on $\mathbb{R}^p$, then the Bayes estimator of θ_i is given by $\hat{\theta}_i^B = (1 - B_i)\hat{\theta}_i + B_i \boldsymbol{x}_i^T \tilde{\boldsymbol{\beta}}$ where $B_i = D_i(\sigma_u^2 + D_i)^{-1}$, $\tilde{\boldsymbol{\beta}} \equiv \tilde{\boldsymbol{\beta}}(\sigma_u^2) = (X'V^{-1}X)^{-1}X'V^{-1}\hat{\boldsymbol{\theta}}$, and $V = \text{Diag}(\sigma_u^2 + D_1, \ldots, \sigma_u^2 + D_m)$. Suppose now we want to match the weighted average of some estimates δ_i to the weighted average of the direct estimates, which we denote by t. We will assume for our calculations that $t = \sum_i w_i \hat{\theta}_i =: \bar{\hat{\theta}}_w$. We denote the normalized weights by w_i, so that $\sum_i w_i = 1$. Under the loss $L(\theta, \delta) = \sum_i w_i(\theta_i - \delta_i)^2$ and subject to $\sum_i w_i \delta_i = \sum_i w_i \hat{\theta}_i$, the benchmarked Bayes estimator as derived in Datta, Ghosh, Steorts and Maples (2011) is given by

$$\hat{\theta}_i^{BM1} = \hat{\theta}_i^B + (\bar{\hat{\theta}}_w - \bar{\hat{\theta}}_w^B); \quad i = 1, \ldots, m; \tag{2.2}$$

where $\bar{\hat{\theta}}_w^B = \sum_i w_i \hat{\theta}_i^B$. In more realistic settings, σ_u^2 is unknown. Define $P_X = X(X^TX)^{-1}X^T$, $h_{ij} = \boldsymbol{x}_i^T(X^TX)^{-1}\boldsymbol{x_j}$, $\hat{u}_i = \hat{\theta}_i - \boldsymbol{x}_i^T \hat{\boldsymbol{\beta}}$, and $\hat{\boldsymbol{\beta}} = (X^TX)^{-1}X^T\hat{\boldsymbol{\theta}}$. In this paper, we consider the simple moment estimator given by $\hat{\sigma}_u^2 = \max\{0, \tilde{\sigma}_u^2\}$ where $\tilde{\sigma}_u^2 = (m-p)^{-1}\left[\sum_{i=1}^m \hat{u}_i^2 - \sum_{i=1}^m D_i(1 - h_{ii})\right]$, which is given in Prasad and Rao (1990). Then the benchmarked EB estimator of θ_i is given by

$$\hat{\theta}_i^{EBM1} = \hat{\theta}_i^{EB} + (\bar{\hat{\theta}}_w - \bar{\hat{\theta}}_w^{EB}), \tag{2.3}$$

where $\hat{\theta}_i^{EB} = (1-\hat{B}_i)\hat{\theta}_i + \hat{B}_i \boldsymbol{x}_i^T \tilde{\boldsymbol{\beta}}(\hat{\sigma}_u^2);\ \hat{B}_i = D_i(\hat{\sigma}_u^2 + D_i)^{-1};\ i = 1,\ldots,m$. The objective of the next two sections will be to obtain the MSE of the benchmarked EB estimator correct up to $O(m^{-1})$ and also to find an estimator of the MSE correct to the same order.

3. Second-Order Approximation to MSE

Wang et al. (2008) construct a simulation study to compare the MSE of the benchmarked EB estimator to the MSE of the EB estimator. In this section, we derive a second order expansion for the MSE of the benchmarked Bayes estimator under the same regularity conditions and assuming the standard benchmarking constraint. That is, assuming the model proposed in Section 2, we obtain a second-order approximation to the MSE of the empirical benchmarked Bayes estimator derived in Section 2. Define $h_{ij}^V = \boldsymbol{x}_i^T(X^TV^{-1}X)^{-1}\boldsymbol{x_j}$ and assume that $\sigma_u^2 > 0$. The following regularity conditions are necessary for establishing Theorem 1:

(i) $0 < D_L \leq \inf_{1\leq i\leq m} D_i \leq \sup_{1\leq i\leq m} D_i \leq D_U < \infty$;
(ii) $\max_{1\leq i\leq m} h_{ii} = O(m^{-1})$; and
(iii) $\max_{1\leq i\leq m} w_i = O(m^{-1})$.

Condition (iii) requires a kind of homogeneity of the small areas, in particular, that there do not exist a few large areas which dominate the remaining small areas in terms of the w_i. Conditions (i) and (ii) are similar to those of Prasad and Rao (1990) and are very often assumed in the small area estimation literature.

Before stating Theorem 1, we first present some lemmas whose proofs are provided in the supplementary material and are used in the proof of Theorem 1. The proof of Theorem 1 can be found in Appendix B.

Lemma 1: *Let $r > 0$ be arbitrary. Then*

(i) $E\left[\left\{\dfrac{\partial \hat{\theta}_i^B}{\partial \sigma_u^2}\right\}^{2r}\right] = O(1)$, *and*

(ii) $E\left[\sup_{\sigma_u^2\geq 0}\left|\dfrac{\partial^2 \hat{\theta}_i^B}{\partial (\sigma_u^2)^2}\right|^{2r}\right] = O(1)$.

Recall that $\boldsymbol{u} = \hat{\boldsymbol{\theta}} - X\boldsymbol{\beta} \sim N(0, V)$. We have the following collection of results:

Lemma 2: *Let $r > 0$ and assume $\max_{1 \le i \le m} \boldsymbol{x}_i^T \boldsymbol{\beta} = O(1)$. Then*

$$||\hat{\boldsymbol{\theta}} - X\tilde{\boldsymbol{\beta}}||^{2r} = O_p(m^r) \quad \textit{and} \quad E\left[||\hat{\boldsymbol{\theta}} - X\tilde{\boldsymbol{\beta}}||^{2r}\right] = O(m^r).$$

Lemma 3: *Let $\boldsymbol{z} \sim N_p(\boldsymbol{0}, \Sigma)$ with matrices $A_{p\times p}$ and $B_{p\times p}$, where B is symmetric. Then*

(i) $Cov(\boldsymbol{z}^T A\boldsymbol{z}, \boldsymbol{z}^T B\boldsymbol{z}) = 2tr(A\Sigma B\Sigma)$

(ii) $Cov(\boldsymbol{z}^T A\boldsymbol{z}, (\boldsymbol{z}^T B\boldsymbol{z})^2) = 4tr(A\Sigma B\Sigma)tr(B\Sigma) + 8tr(A\Sigma B\Sigma B\Sigma)$.

Lemma 4: $E[(\tilde{\sigma}_u^2 - \sigma_u^2)^2] = 2(m-p)^{-2}\sum_{i=1}^m (\sigma_u^2 + D_i)^2 + O(m^{-2})$.

Theorem 1. *Assume regularity conditions (i)–(iii) hold. Then $E[(\hat{\theta}_i^{EBM1} - \theta_i)^2] = g_{1i}(\sigma_u^2) + g_{2i}(\sigma_u^2) + g_{3i}(\sigma_u^2) + g_4(\sigma_u^2) + o(m^{-1})$, where*

$$\begin{aligned}
g_{1i}(\sigma_u^2) &= B_i\sigma_u^2 \\
g_{2i}(\sigma_u^2) &= B_i^2 h_{ii}^V \\
g_{3i}(\sigma_u^2) &= B_i^3 D_i^{-1} Var(\tilde{\sigma}_u^2) \\
g_4(\sigma_u^2) &= \sum_{i=1}^m w_i^2 B_i^2 V_i - \sum_{i=1}^m \sum_{j=1}^m w_i w_j B_i B_j h_{ij}^V,
\end{aligned}$$

and where $Var(\tilde{\sigma}_u^2) = 2(m-p)^{-2}\sum_{k=1}^m (\sigma_u^2 + D_k)^2 + o(m^{-1})$.

Remark 1: We note that the the MSE of the benchmarked EB estimator in Theorem 1 will always be non-negative. It is clear that $g_{1i}(\sigma_u^2)$, $g_{2i}(\sigma_u^2)$, and $g_{3i}(\sigma_u^2)$ are non-negative. Next, we establish the non-negativity of $g_4(\sigma_u^2)$. Let $\boldsymbol{q} = (\boldsymbol{q}_1, \ldots, \boldsymbol{q}_m)$, where $q_i = w_i B_i V_i^{1/2}$. We can write $g_4(\sigma_u^2) = \boldsymbol{q}^T(I - \tilde{P}_X^T)\boldsymbol{q}$, where $\tilde{P}_X^T = V^{-1/2}X(X^TV^{-1}X)^{-1}X^TV^{-1/2}$. Thus, $g_4(\sigma_u^2) \ge 0$, and hence, the MSE in Theorem 1 will always be non-negative.

4. Estimator of MSE Approximation

We now obtain an estimator of the MSE approximation for the Fay-Herriot model (assuming normality). Theorem 2 shows that the expectation of the MSE estimator is correct up to $O(m^{-1})$.

Lemma 5: *Suppose that*

$$\sup_{t\in T} |h'(t)| = O(m^{-1}) \tag{4.1}$$

for some interval $T \subseteq \mathbb{R}$. *If* $\hat{\sigma}_u^2, \sigma_u^2 \in T$ *w.p.* 1, *then* $E[h(\hat{\sigma}_u^2)] = h(\sigma_u^2) + o(m^{-1})$.

Proof. Consider the expansion $h(\hat{\sigma}_u^2) = h(\sigma_u^2) + h'(\sigma_u^{*2})(\hat{\sigma}_u^2 - \sigma_u^2)$ for some σ_u^{*2} between σ_u^2 and $\hat{\sigma}_u^2$. Then $\sigma_u^{*2} \in T$ a.s., and $h'(\sigma_u^{*2}) \leq \sup_{t\in T} |h'(t)|$ a.s. as well. This implies $E[h'(\sigma_u^{*2})(\hat{\sigma}_u^2 - \sigma_u^2)] \leq \sup_{t\in T} |h'(t)| E|\hat{\sigma}_u^2 - \sigma_u^2| = O(m^{-3/2})$ by equation (4.1) and since $E|\hat{\sigma}_u^2 - \sigma_u^2| \leq E^{\frac{1}{2}}[(\hat{\sigma}_u^2 - \sigma_u^2)^2]$. Hence, if assumption (4.1) holds, then $E[h(\hat{\sigma}_u^2)] = h(\sigma_u^2) + o(m^{-1})$. □

Theorem 2. $E[g_{1i}(\hat{\sigma}_u^2) + g_{2i}(\hat{\sigma}_u^2) + 2g_{3i}(\hat{\sigma}_u^2) + g_4(\hat{\sigma}_u^2)] = g_{1i}(\sigma_u^2) + g_{2i}(\sigma_u^2) + g_{3i}(\sigma_u^2) + g_4(\sigma_u^2) + o(m^{-1})$, *where* $g_{1i}(\sigma_u^2), g_{2i}(\sigma_u^2), g_{3i}(\sigma_u^2)$, *and* $g_4(\sigma_u^2)$ *are defined in Theorem 1.*

Proof. By Theorem A.3 in Prasad and Rao (1990), $E[g_{1i}(\hat{\sigma}_u^2)+g_{2i}(\hat{\sigma}_u^2)+2g_{3i}(\hat{\sigma}_u^2)] = g_{1i}(\sigma_u^2) + g_{2i}(\sigma_u^2) + g_{3i}(\sigma_u^2) + o(m^{-1})$. In addition, we consider $E[g_4(\hat{\sigma}_u^2)]$, where $g_4(\sigma_u^2) = \sum_{i=1}^m w_i^2 B_i^2 V_i - \sum_{i=1}^m \sum_{j=1}^m w_i w_j B_i B_j h_{ij}^V =: g_{41}(\sigma_u^2) + g_{42}(\sigma_u^2)$. We first show that the derivatives of $g_{41}(\sigma_u^2)$ and $g_{42}(\sigma_u^2)$ satisfy assumption (4.1). Let $T = [0, \infty)$. Consider

$$\sup_{\sigma_u^2 \geq 0} \left| \frac{\partial g_{41}(\sigma_u^2)}{\partial \sigma_u^2} \right| = \sup_{\sigma_u^2 \geq 0} \sum_{i=1}^m w_i^2 B_i^2 = O(m^{-1}).$$

It can be shown that $\dfrac{\partial B_i B_j}{\partial \sigma_u^2} = -B_i B_j^2 D_j^{-1} - B_i^2 B_j D_i^{-1}$ and $(X^T V^{-1} X)^{-1} \leq (X^T V^{-2} X)^{-1} D_L^{-1}$. Observe that

$$\begin{aligned}
\left| \frac{\partial g_{42}(\sigma_u^2)}{\partial \sigma_u^2} \right| &\leq \sum_{i=1}^m \sum_{j=1}^m w_i w_j \Big[|B_i D_L^{-1} h_{ij}^V| + |B_j D_L^{-1} h_{ij}^V| \\
&\quad + B_i B_j \boldsymbol{x}_i^T (X^T V^{-1} X)^{-1} X^T V^{-2} X (X^T V^{-1} X)^{-1} \boldsymbol{x}_i \Big] \\
&\leq 3m^2 (\max_{1\leq i\leq m} w_i)^2 D_L^{-1} B_i (\sigma_u^2 + D_U)(\max_{1\leq i\leq m} h_i) \\
&\leq 3m^2 (\max_{1\leq i\leq m} w_i)^2 D_L^{-1} D_U (\sigma_u^2 + D_L)^{-1} (\sigma_u^2 + D_U)(\max_{1\leq i\leq m} h_i) \\
&= 3m^2 (\max_{1\leq i\leq m} w_i)^2 D_L^{-1} D_U (1 + D_U D_L^{-1})(\max_{1\leq i\leq m} h_i) = O(m^{-1}).
\end{aligned}$$

This implies that $\sup_{\sigma_u^2\geq 0}\left|\frac{\partial g_{42}(\sigma_u^2)}{\partial\sigma_u^2}\right| = O(m^{-1})$. Since the derivatives of $g_{41}(\sigma_u^2)$ and $g_{42}(\sigma_u^2)$ satisfy assumption (4.1), we know that $E[g_4(\hat{\sigma}_u^2)] = g_4(\sigma_u^2) + o(m^{-1})$. $\square$

5. Parametric Bootstrap Estimator of the MSE of the Benchmarked Empirical Bayes Estimator

In this section, we extend the methods of Butar and Lahiri (2003) to find a parametric bootstrap estimator of the MSE of the benchmarked EB estimator. Under the proposed model, the expectation of the proposed measure of uncertainty of the benchmarked EB estimator is correct up to order $O(m^{-1})$.

To introduce the parametric bootstrap method, consider the following model:

$$\begin{aligned} \hat{\theta}_i^*|u_i^* &\overset{ind}{\sim} N(\boldsymbol{x}_i^T\tilde{\boldsymbol{\beta}} + u_i^*, D_i) \\ u_i^* &\overset{ind}{\sim} N(0, \hat{\sigma}_u^2). \end{aligned} \tag{5.1}$$

As explained in Butar and Lahiri (2003), we use the parametric bootstrap twice. We first use it to estimate $g_{1i}(\sigma_u^2), g_{2i}(\sigma_u^2)$, and $g_4(\sigma_u^2)$ by correcting the bias of $g_{1i}(\hat{\sigma}_u^2), g_{2i}(\hat{\sigma}_u^2)$, and $g_4(\hat{\sigma}_u^2)$. We then use it again to estimate $E[(\hat{\theta}_i^{EB} - \hat{\theta}_i^B)^2] = g_{3i}(\sigma_u^2) + o(m^{-1})$.

Butar and Lahiri (2003) derived a parametric bootstrap estimator for the MSE of the EB estimator under the Fay and Herriot (1979) model. Using Theorem A.1 of their paper, they show that the bootstrap estimator V_i^{BOOT} is

$$V_i^{\text{BOOT}} = 2[g_{1i}(\hat{\sigma}_u^2) + g_{2i}(\hat{\sigma}_u^2)] - E_*\left[g_{1i}(\hat{\sigma}_u^{*2}) + g_{2i}(\hat{\sigma}_u^{*2})\right] + E_*[(\hat{\theta}_i^{EB*} - \hat{\theta}_i^{EB})^2], \tag{5.2}$$

where E_* denotes the expectation computed with respect to the model given in (5.1), and $\hat{\theta}_i^{EB*} = (1 - B_i(\hat{\sigma}_u^{*2}))\hat{\theta}_i + B_i(\hat{\sigma}_u^{*2})\boldsymbol{x}_i^T\hat{\boldsymbol{\beta}}$. Following their work, we propose a parametric bootstrap estimator of the MSE of the benchmarked EB estimator which is a simple extension of (5.2).

We propose to estimate $g_{1i}(\sigma_u^2) + g_{2i}(\sigma_u^2) + g_4(\sigma_u^2)$ by

$$2[g_{1i}(\hat{\sigma}_u^2) + g_{2i}(\hat{\sigma}_u^2) + g_4(\hat{\sigma}_u^2)] - E_*\left[g_{1i}(\hat{\sigma}_u^{*2}) + g_{2i}(\hat{\sigma}_u^{*2}) + g_4(\hat{\sigma}_u^{*2})\right]$$

and then estimate $E[(\hat{\theta}_i^{EB} - \hat{\theta}_i^B)^2]$ by $E_*[(\hat{\theta}_i^{EB*} - \hat{\theta}_i^{EB})^2]$. Thus, our proposed estimator of $\text{MSE}[\hat{\theta}_i^{\text{EBM1}}]$ is given by

$$V_i^{\text{B-BOOT}} = 2[g_{1i}(\hat{\sigma}_u^2) + g_{2i}(\hat{\sigma}_u^2) + g_4(\hat{\sigma}_u^2)] - E_*\left[g_{1i}(\hat{\sigma}_u^{*2}) + g_{2i}(\hat{\sigma}_u^{*2}) + g_4(\hat{\sigma}_u^{*2})\right] + E_*[(\hat{\theta}_i^{EB*} - \hat{\theta}_i^{EB})^2].$$

We now show that the expectation of $V_i^{\text{B-BOOT}}$ is correct up to $O(m^{-1})$.

Theorem 3. *$E[V_i^{B\text{-}BOOT}] = MSE[\hat{\theta}_i^{EBM1}] + o(m^{-1})$.*

Proof. First, by Theorem A.1 in Butar and Lahiri (2003), we note that

$$E_*[g_{1i}(\hat{\sigma}_u^{*2})] = g_{1i}(\hat{\sigma}_u^2) - g_{3i}(\hat{\sigma}_u^2) + o_p(m^{-1}),$$
$$E_*[g_{2i}(\hat{\sigma}_u^{*2})] = g_{2i}(\hat{\sigma}_u^2) + o_p(m^{-1}), \quad \text{and}$$
$$E_*[(\hat{\theta}_i^{EB*} - \hat{\theta}_i^{EB})^2] = g_{5i}(\hat{\sigma}_u^2) + o_p(m^{-1}),$$

where $g_{5i}(\hat{\sigma}_u^2) = [B_i(\hat{\sigma}_u^2)]^4 D_i^{-2}\left(\hat{\theta}_i - \boldsymbol{x}_i^T \tilde{\boldsymbol{\beta}}(\hat{\sigma}_u^2)\right)^2$. Also, $E_*[g_4(\hat{\sigma}_u^{*2})] = g_4(\hat{\sigma}_u^2) + o_p(m^{-1})$, which follows along the lines of the proof of Theorem A.2(b) of Datta and Lahiri (2000). Applying these results and Theorem 2 of this paper, we find

$$V_i^{\text{B-BOOT}} = g_{1i}(\hat{\sigma}_u^2) + g_{2i}(\hat{\sigma}_u^2) + g_{3i}(\hat{\sigma}_u^2) + g_4(\hat{\sigma}_u^2) + g_{5i}(\hat{\sigma}_u^2) + o_p(m^{-1}).$$

This implies that

$$E[V_i^{\text{B-BOOT}}] = g_{1i}(\sigma_u^2) + g_{2i}(\sigma_u^2) + g_{3i}(\sigma_u^2) + g_4(\sigma_u^2) + o(m^{-1})$$

since $E[g_{5i}(\hat{\sigma}_u^2)] = g_{3i}(\sigma_u^2) + o(m^{-1})$ by Butar and Lahiri (2003), and applying the results of Prasad and Rao (1990). □

6. Two Applications

In this section, we consider two examples (one real data set and one simulation study) where in both we are looking to compare the performance of the estimator of the MSE of the benchmarked EB estimator and the parametric bootstrap estimator of the MSE of the benchmarked EB estimator. All tables and figures that correspond with this section can be found in Appendix A.

We consider data from the Small Area Income and Poverty Estimates (SAIPE) program at the U.S. Census Bureau, which produces model-based estimates of the number of poor school-aged children (5–17 years old) at the national, state, county, and district levels. The school district estimates are benchmarked to the state estimates by the Department of Education to allocate funds under the No Child Left Behind Act of 2001. Specifically, we consider year 1997. In the SAIPE program, the model-based state estimates are benchmarked to the national school-aged poverty rate using the benchmarked estimator in (2.3). The number of poor school-aged children has been collected from the Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS) from 1995 to 2004, while the American Community Survey (ACS) estimates have been used beginning in 2005. Additionally, the model-based county estimates are benchmarked to the model-based state estimates using the the benchmarked estimator in (2.3).

In the SAIPE program, the state model for poverty rates in school-aged children follows the basic Fay-Herriot (1979) framework where $\hat{\theta}_i = \theta_i + e_i$ and $\theta_i = \boldsymbol{x}_i^T\boldsymbol{\beta} + u_i$ where θ_i is the true state level poverty rate, $\hat{\theta}_i$ is the direct survey estimate (from CPS ASEC), e_i is the sampling error term with assumed known variance $D_i > 0$, $\boldsymbol{x}_i$ are the predictors, $\boldsymbol{\beta}$ is the unknown vector of regression coefficients, and u_i is the model error with unknown variance σ_u^2. The explanatory variables in the model are IRS income tax–based pseudo-estimate of the child poverty rate, IRS non-filer rate, food stamp rate, and the residual term from the regression of the 1990 Census estimated child poverty rate. We estimate $\boldsymbol{\beta}$ using the weighted least squares type estimator $\tilde{\boldsymbol{\beta}}(\hat{\sigma}_u^2) = (X'V^{-1}X)^{-1}X'V^{-1}\hat{\boldsymbol{\theta}}$, and we estimate σ_u^2 using the modified moment estimator $\hat{\sigma}_u^2$ from Section 2.

As shown in Table A.1, the estimated MSE of the EB estimator, $\text{mse}(\hat{\theta}_i^{EB})$, compared to the estimated MSE of the benchmarked EB estimator, $\text{mse}(\hat{\theta}_i^{EBM1})$, differs by the constant $g_4(\sigma_u^2)$, which is 0.025. This constant is effectively the increase in MSE that we suffer from benchmarking, and we see that in this case this quantity is small (compared to the values of the MSEs). Generally speaking this quantity is expected to be small since $g_4(\sigma_u^2) = O(m^{-1})$.

In Table A.1, we define mse^B and mse^{BB} as the bootstrap estimates of the MSE of the EB estimator and the benchmarked EB estimator respectively. As already mentioned, we consider year 1997 for illustrative purposes. When we perform the bootstrapping, we resample $\tilde{\sigma}_u^{*2}$ $10,000$ times in order to calculate mse^B and mse^{BB}. This is best understood by considering the concept behind our bootstrapping approach. Consider the behavior of $g_{1i}(\sigma_u^2)$, the only term that is $O(1)$. Ordinarily, $g_{1i}(\hat{\sigma}_u^2)$ underestimates $g_{1i}(\sigma_u^2)$, and $E_*[g_{1i}(\hat{\sigma}_u^2)]$ underestimates $g_{1i}(\hat{\sigma}_u^2)$. The basic idea is that we use the amount by which $E_*[g_{1i}(\hat{\sigma}_u^2)]$ underestimates $g_{1i}(\hat{\sigma}_u^2)$ as an approximation of the amount by which $g_{1i}(\hat{\sigma}_u^2)$ underestimates $g_{1i}(\sigma_u^2)$.

We run into a problem with the 1997 data, where $g_{1i}(\hat{\sigma}_u^2)$ is 0, since in this case $E_*[g_{1i}(\hat{\sigma}_u^2)]$ overestimates $g_{1i}(\hat{\sigma}_u^2)$. Recall that

$$V_i^{\text{B-BOOT}} = g_{1i}(\hat{\sigma}_u^2) + \{g_{1i}(\hat{\sigma}_u^2) - E_*[g_{1i}(\hat{\sigma}_u^{*2})]\} + O(m^{-1}).$$

Since $g_{1i}(\hat{\sigma}_u^2)$ is 0 and is the dominating term of $V_i^{\text{B-BOOT}}$, many of the estimated MSEs of the benchmarked bootstrapped estimator (mse^{BB}) will be negative. Also, observe this same behavior holds true for the bootstrapped estimator proposed by Butar and Lahiri (2003), which we denote by mse^{B}. Hence, we do not recommend using bootstrapping when $\hat{\sigma}_u^2$ is too close to zero because of the form of $\hat{\sigma}_u^2$. We also note that the MSE of the benchmarked EB estimator is always non-negative as explained in Remark 1 of Section 3.

In the second example, we consider a simulation study, where we take the same covariates from the SAIPE dataset from year 1997. We generate our data from the following model:

$$\begin{aligned} \hat{\theta}_i|\theta_i &\overset{ind}{\sim} N(\theta_i, D_i) \\ \theta_i &\overset{ind}{\sim} N(X^T\boldsymbol{\beta}, \sigma_u^2), \end{aligned} \tag{6.1}$$

where D_i comes from the SAIPE dataset. We first simulate 10,000 sets of values for θ_i and $\hat{\theta}_i$ using model (6.1). We then use each set of $\hat{\theta}_i$ values as the data and compute the EB and benchmarked EB estimators according to equation (2.3) and the EB formula given below it. In order to use EB, we assume that $\boldsymbol{\beta} = (-3, 0.5, 1, 1, 0.5)^T$ and $\sigma_u^2 = 5$.

In Figure A.1, we compare the estimator of the theoretical MSE of the benchmarked EB estimator and the bootstrap estimator of the MSE of the benchmarked EB estimator with the true value, i.e., the average of the squared difference between the estimator values and the true θ_i which was generated according to model (6.1). In the upper plot, we find that the estimator of the theoretical MSE of the benchmarked EB estimator overshoots the truth very slightly, which shows that our estimator is slightly conservative. We find the opposite behavior to be true of the bootstrap estimator of the MSE of the benchmarked Bayes estimator, meaning that it undershoots the truth slightly.

In practice, it seems safer to use a MSE estimator which overestimates than one that underestimates, and hence, we recommend using our proposed MSE estimator over the bootstrapped MSE estimator. Using the lower plot, we compare the theoretical Prasad Rao (PR) MSE estimator with the associated true value. We find the same behavior in the PR estimator as we did in our proposed theoretical MSE of the benchmarked EB estimator. The overshoot occurs in the terms that the estimators have in common, i.e., $g_{1i}(\sigma_u^2); g_{2i}(\sigma_u^2)$; and $g_{3i}(\sigma_u^2)$. We see that for this particular simulation study where m is particularly large at 10,000, the difference between the two MSEs is indistinguishable.

7. Summary and Conclusion

We have shown via theoretical calculations that the increase in MSE due to benchmarking under our modeling assumptions is quite small for the Fay-Herriot model, specifically $O(m^{-1})$. We have derived an asymptotically unbiased estimate of the MSE of the benchmarked EB estimator (EBLUP) under these same assumptions which is correct to order $O(m^{-1})$. We have derived a parametric bootstrap estimator of the benchmarked EB estimator based on work done by Butar and Lahiri (2003). Furthermore, we have illustrated our methodology for a data set for fixed m using U.S. Census data. Since our theoretical estimator of the MSE under benchmarking is guaranteed to be positive, we recommend this choice over the one derived by bootstrapping. We also perform a simulation study that suggests use of the theoretical estimator of the MSE under benchmarking. In closing, it is important to pursue further work for more complex models, and in particular, when it is necessary to achieve multi-stage benchmarking. We believe that the present work will be the genesis for future work in this direction.

Acknowledgment

The research was partially supported by NSF Grant SES 1026165 and the United States Census Bureau Dissertation Fellowship Program. The views expressed here are those of the authors and do not reflect those of the U.S. Census Bureau. We would like to express our thanks to the Associate Editor, the referees, and Professor J.N.K. Rao for their helpful suggestions.

Appendix A

Table A.1: Table of estimates for 1997

i	$\hat{\theta}_i$	$\hat{\theta}_i^{EB}$	$\hat{\theta}_i^{EBM1}$	mse($\hat{\theta}_i$)	mse($\hat{\theta}_i^{EB}$)	mse($\hat{\theta}_i^{EBM1}$)	mse^{B}	mse^{BB}
1	25.16	21.38	21.56	15.72	1.38	1.41	0.02	0.04
2	10.99	14.94	15.11	10.44	2.12	2.14	0.66	0.68
3	23.35	20.89	21.06	11.84	1.68	1.70	0.00	0.01
4	23.32	22.18	22.35	13.85	1.90	1.92	0.37	0.38
5	23.55	22.71	22.88	2.39	5.92	5.94	1.12	1.13
6	9.14	13.12	13.29	6.38	2.19	2.22	0.36	0.38
7	10.34	13.39	13.56	9.85	2.08	2.10	0.39	0.41
8	15.54	13.06	13.23	17.56	0.91	0.94	-0.47	-0.45
9	35.85	32.43	32.60	32.35	4.92	4.95	3.49	3.50
10	18.34	19.59	19.76	3.70	3.71	3.74	0.40	0.41
11	23.52	20.53	20.70	12.93	1.16	1.19	-0.38	-0.37
12	18.98	13.72	13.89	20.87	2.45	2.48	1.24	1.26
13	17.56	13.64	13.82	12.38	1.70	1.73	0.23	0.25
14	14.57	15.72	15.89	3.56	3.45	3.47	-0.06	-0.05
15	11.07	12.53	12.70	7.58	1.84	1.86	-0.23	-0.22
16	11.09	11.21	11.38	8.49	1.74	1.76	-0.24	-0.22
17	11.01	13.48	13.65	9.34	1.61	1.63	-0.15	-0.14
18	23.12	20.78	20.95	13.98	1.37	1.40	-0.12	-0.11
19	21.08	24.15	24.32	15.19	1.80	1.82	0.40	0.42
20	13.18	12.44	12.61	13.63	2.09	2.11	0.56	0.57
21	9.90	13.16	13.33	9.28	1.65	1.67	-0.03	-0.01
22	19.66	14.38	14.56	7.66	2.46	2.48	1.02	1.04
23	13.78	16.86	17.03	4.04	3.11	3.13	0.38	0.39
24	14.34	10.11	10.28	9.91	1.64	1.67	0.16	0.17
25	20.58	22.30	22.47	15.07	2.42	2.45	0.97	0.99
26	18.90	15.11	15.28	15.24	1.00	1.03	-0.37	-0.35
27	17.00	18.60	18.77	12.95	1.37	1.40	-0.21	-0.19

Table A.1: Table of estimates for 1997 (continued)

i	$\hat{\theta}_i$	$\hat{\theta}_i^{EB}$	$\hat{\theta}_i^{EBM1}$	mse($\hat{\theta}_i$)	mse($\hat{\theta}_i^{EB}$)	mse($\hat{\theta}_i^{EBM1}$)	mse^B	mse^{BB}
28	9.72	9.62	9.79	7.18	2.24	2.26	0.09	0.10
29	14.06	12.94	13.12	10.23	1.71	1.74	-0.06	-0.04
30	10.94	6.72	6.89	11.35	1.88	1.91	0.50	0.52
31	14.66	13.28	13.45	5.52	2.48	2.51	-0.03	-0.01
32	29.69	24.44	24.61	13.18	2.62	2.65	1.38	1.40
33	23.76	22.85	23.02	3.10	4.76	4.79	0.94	0.95
34	13.90	16.58	16.75	5.70	2.29	2.31	-0.01	0.01
35	18.19	13.64	13.81	11.92	1.81	1.84	0.48	0.50
36	13.91	13.64	13.81	3.95	3.07	3.10	-0.25	-0.23
37	16.09	21.50	21.68	11.14	1.52	1.54	0.24	0.26
38	12.60	13.43	13.60	10.35	2.53	2.56	0.83	0.84
39	14.61	13.92	14.09	3.73	3.40	3.42	-0.01	0.00
40	20.37	14.60	14.77	18.53	1.04	1.07	-0.15	-0.14
41	18.74	21.21	21.38	14.57	1.49	1.52	0.02	0.04
42	12.87	15.77	15.94	12.94	1.98	2.01	0.46	0.47
43	16.09	16.10	16.27	11.94	1.92	1.95	0.28	0.30
44	21.95	21.38	21.55	3.38	4.05	4.07	0.38	0.40
45	11.27	9.76	9.93	9.45	2.28	2.31	0.50	0.51
46	11.15	10.10	10.27	11.95	2.45	2.48	0.86	0.88
47	16.40	14.96	15.13	11.51	1.20	1.22	-0.49	-0.47
48	12.26	13.17	13.34	9.33	1.85	1.87	0.01	0.02
49	18.76	22.25	22.42	13.73	3.81	3.83	2.46	2.48
50	7.60	11.87	12.04	6.41	2.74	2.76	0.97	0.98
51	11.74	11.70	11.87	8.86	2.08	2.10	0.17	0.19

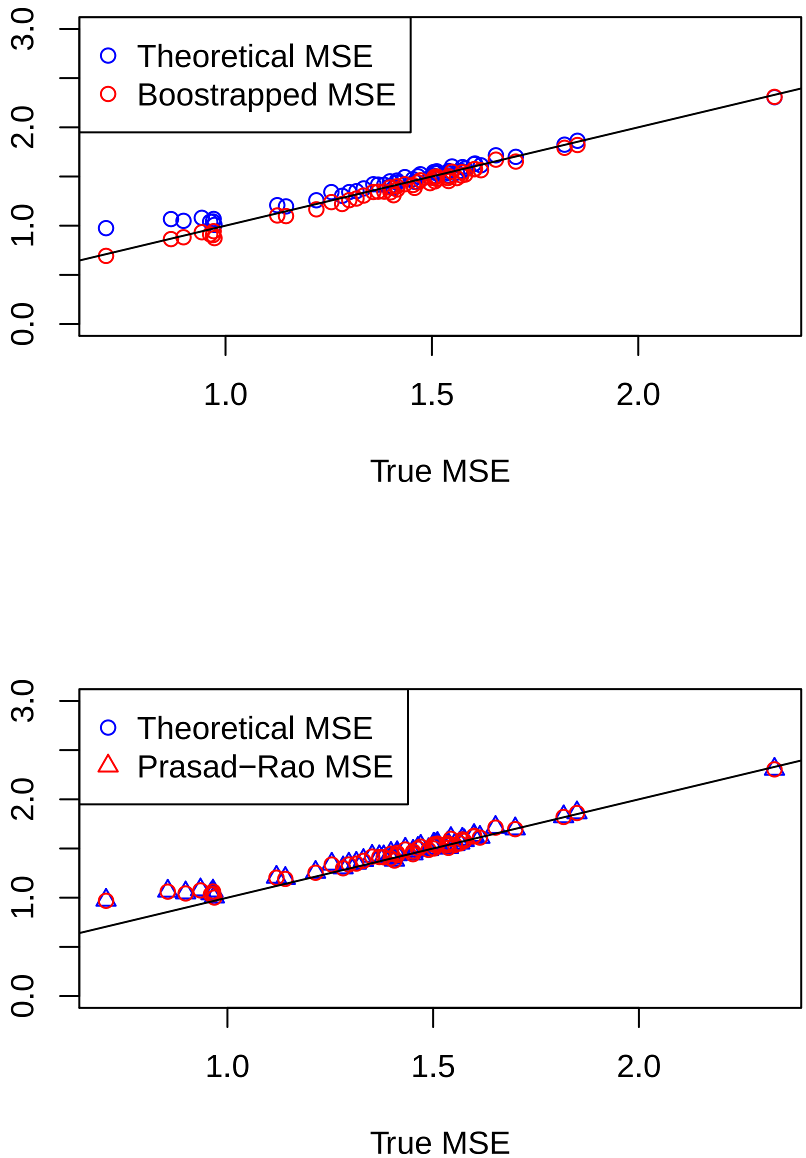

Figure A.1: Comparing Simulated MSEs with True MSEs

Appendix B

We now include the proof of Theorem 1.

Proof of Theorem 1. Observe

$$\begin{aligned}
E[(\hat{\theta}_i^{EBM1} - \theta_i)^2] &= E[(\hat{\theta}_i^{B} - \theta_i)^2] + E[(\hat{\theta}_i^{EBM1} - \hat{\theta}_i^{B})^2] \\
&= E[(\hat{\theta}_i^{B} - \theta_i)^2] + E[(\hat{\theta}_i^{B} - \hat{\theta}_i^{EB} - t + \bar{\hat{\theta}}_w^{EB})^2] \\
&= E[(\hat{\theta}_i^{B} - \theta_i)^2] + E[(\hat{\theta}_i^{B} - \hat{\theta}_i^{EB} + \bar{\hat{\theta}}_w^{EB} - \bar{\hat{\theta}}_w^{B} + \bar{\hat{\theta}}_w^{B} - t)^2] \\
&= E[(\hat{\theta}_i^{B} - \theta_i)^2] + E[(\hat{\theta}_i^{EB} - \hat{\theta}_i^{B})^2] + E[(\bar{\hat{\theta}}_w^{B} - t)^2] + E[(\bar{\hat{\theta}}_w^{EB} - \bar{\hat{\theta}}_w^{B})^2] \\
&\quad - 2E[(\hat{\theta}_i^{EB} - \hat{\theta}_i^{B})(\bar{\hat{\theta}}_w^{EB} - \bar{\hat{\theta}}_w^{B})] - 2E[(\hat{\theta}_i^{EB} - \hat{\theta}_i^{B})(\bar{\hat{\theta}}_w^{B} - t)] \\
&\quad + 2E[(\bar{\hat{\theta}}_w^{EB} - \bar{\hat{\theta}}_w^{B})(\bar{\hat{\theta}}_w^{B} - t)]. \qquad \text{(B.1)}
\end{aligned}$$

Next, we observe that $E[(\hat{\theta}_i^{B} - \theta_i)^2] + E[\hat{\theta}_i^{EB} - \hat{\theta}_i^{B}]^2 = g_{1i}(\sigma_u^2) + g_{2i}(\sigma_u^2) + g_{3i}(\sigma_u^2) + o(m^{-1})$ by Prasad and Rao (1990), where

$$\begin{aligned}
g_{1i}(\sigma_u^2) &= B_i \sigma_u^2 \\
g_{2i}(\sigma_u^2) &= B_i^2 h_{ii}^V \\
g_{3i}(\sigma_u^2) &= B_i^3 D_i^{-1} \mathrm{Var}(\tilde{\sigma}_u^2).
\end{aligned}$$

It may be noted that while $g_{1i}(\sigma_u^2) = O(1)$, both $g_{2i}(\sigma_u^2)$ and $g_{3i}(\sigma_u^2)$ are of order $O(m^{-1})$ as shown in Prasad and Rao (1990). We will show that $E[(\bar{\hat{\theta}}_w^{B} - t)^2] = g_4(\sigma_u^2) = O(m^{-1})$, whereas the remaining four terms of expression (B.1) are of order $o(m^{-1})$.

First we show that $E[(\bar{\hat{\theta}}_w^{B} - t)^2] = g_4(\sigma_u^2)$. We observe $\bar{\hat{\theta}}_w^{B} - t = -\sum_{i=1}^m w_i B_i (\hat{\theta}_i - \boldsymbol{x}_i^T \tilde{\boldsymbol{\beta}})$ and consider

$$\begin{aligned}
E[(\bar{\hat{\theta}}_w^{B} - t)^2] &= E\left[\left\{\sum_{i=1}^m w_i B_i (\hat{\theta}_i - \boldsymbol{x}_i^T \tilde{\boldsymbol{\beta}})\right\}^2\right] \\
&= \sum_{i=1}^m w_i^2 B_i^2 E[(\hat{\theta}_i - \boldsymbol{x}_i^T \tilde{\boldsymbol{\beta}})^2] + \sum_{i \neq j} w_i w_j B_i B_j E[(\hat{\theta}_i - \boldsymbol{x}_i^T \tilde{\boldsymbol{\beta}})(\hat{\theta}_j - \boldsymbol{x_j}^T \tilde{\boldsymbol{\beta}})] \\
&= \sum_{i=1}^m w_i^2 B_i^2 (V_i - h_{ii}^V) + \sum_{i \neq j} w_i w_j B_i B_j (-h_{ij}^V) \\
&= \sum_{i=1}^m w_i^2 B_i^2 V_i - \sum_{i=1}^m \sum_{j=1}^m w_i w_j B_i B_j h_{ij}^V. \qquad \text{(B.2)}
\end{aligned}$$

We may note that the expression on the right hand side of (B.2) is $O(m^{-1})$ since $\max_{1\leq i\leq m} h_{ii} = O(m^{-1})$, which implies that $\max_{1\leq i\leq j\leq m} h_{ij}^V = O(m^{-1})$.

Next, we return to (B.1) and show that $E[(\bar{\hat{\theta}}_w^{EB} - \bar{\hat{\theta}}_w^B)^2] = o(m^{-1})$. Consider

$$E[(\bar{\hat{\theta}}_w^{EB} - \bar{\hat{\theta}}_w^B)^2] = \sum_i w_i^2 E\left[(\hat{\theta}_i^{EB} - \hat{\theta}_i^B)^2\right] + 2\sum_{i=1}^{m-1}\sum_{j=i+1}^{m} w_i w_j E\left[(\hat{\theta}_i^{EB} - \hat{\theta}_i^B)(\hat{\theta}_j^{EB} - \hat{\theta}_j^B)\right]$$

$$= 2\sum_{i=1}^{m-1}\sum_{j=i+1}^{m} w_i w_j E\left[(\hat{\theta}_i^{EB} - \hat{\theta}_i^B)(\hat{\theta}_j^{EB} - \hat{\theta}_j^B)\right] + o(m^{-1}) \quad \text{(B.3)}$$

since $\sum_i w_i^2 E[(\hat{\theta}_i^{EB} - \hat{\theta}_i^B)^2] = o(m^{-1})$. The latter holds because $E[(\hat{\theta}_i^{EB} - \hat{\theta}_i^B)^2] = g_{2i}(\sigma_u^2) + g_{3i}(\sigma_u^2) = O(m^{-1})$, $\max_{1\leq i\leq m} w_i = O(m^{-1})$, and $\sum_i w_i = 1$. Thus, it suffices to show $E\left[(\hat{\theta}_i^{EB} - \hat{\theta}_i^B)(\hat{\theta}_j^{EB} - \hat{\theta}_j^B)\right] = o(m^{-1})$ for all $i \neq j$, and we do so by expanding $\hat{\theta}_i^{EB}$ about $\hat{\theta}_i^B$. For simplicity of notation, denote $\dfrac{\partial \hat{\theta}_i^B}{\partial \sigma_u^2} = \dfrac{\partial \hat{\theta}_i^B(\sigma_u^2)}{\partial \sigma_u^2}$ and $\dfrac{\partial^2 \hat{\theta}_{i*}^B}{\partial(\sigma_u^2)^2} = \dfrac{\partial^2 \hat{\theta}_i^B(\sigma_u^{*2})}{\partial(\sigma_u^2)^2}$. This results in

$$\hat{\theta}_i^{EB} - \hat{\theta}_i^B = \frac{\partial \hat{\theta}_i^B}{\partial \sigma_u^2}(\hat{\sigma}_u^2 - \sigma_u^2) + \frac{1}{2}\frac{\partial^2 \hat{\theta}_{i*}^B}{\partial(\sigma_u^2)^2}(\hat{\sigma}_u^2 - \sigma_u^2)^2$$

for some σ_u^{*2} between σ_u^2 and $\hat{\sigma}_u^2$. The expansion of $\hat{\theta}_j^{EB}$ about $\hat{\theta}_j^B$ is similar.

We now consider $E[(\hat{\theta}_i^{EB} - \hat{\theta}_i^B)(\hat{\theta}_j^{EB} - \hat{\theta}_j^B)]$ for $i \neq j$. Notice that

$$\begin{aligned} E[(\hat{\theta}_i^{EB} - \hat{\theta}_i^B)(\hat{\theta}_j^{EB} - \hat{\theta}_j^B)] &= E\left[\frac{\partial \hat{\theta}_i^B}{\partial \sigma_u^2}\frac{\partial \hat{\theta}_j^B}{\partial \sigma_u^2}(\hat{\sigma}_u^2 - \sigma_u^2)^2\right] + \frac{1}{2}E\left[\frac{\partial \hat{\theta}_i^B}{\partial \sigma_u^2}\frac{\partial^2 \hat{\theta}_{j*}^B}{\partial(\sigma_u^2)^2}(\hat{\sigma}_u^2 - \sigma_u^2)^3\right] \\ &+ \frac{1}{2}E\left[\frac{\partial^2 \hat{\theta}_{i*}^B}{\partial(\sigma_u^{*2})^2}\frac{\partial \hat{\theta}_j^B}{\partial \sigma_u^2}(\hat{\sigma}_u^2 - \sigma_u^2)^3\right] + \frac{1}{4}E\left[\frac{\partial^2 \hat{\theta}_{i*}^B}{(\partial\sigma_u^2)^2}\frac{\partial^2 \hat{\theta}_{j*}^B}{\partial^2(\sigma_u^2)^2}(\hat{\sigma}_u^2 - \sigma_u^2)^4\right] \\ &:= R_0 + R_1 + R_2 + R_3. \end{aligned}$$

In R_1, consider that

$$\begin{aligned} E\left[\frac{\partial \hat{\theta}_i^B}{\partial \sigma_u^2}\frac{\partial^2 \hat{\theta}_{j*}^B}{\partial(\sigma_u^2)^2}(\hat{\sigma}_u^2 - \sigma_u^2)^3\right] &= E\left[\frac{\partial \hat{\theta}_i^B}{\partial \sigma_u^2}\frac{\partial^2 \hat{\theta}_{j*}^B}{\partial(\sigma_u^2)^2}(\tilde{\sigma}_u^2 - \sigma_u^2)^3 I(\tilde{\sigma}_u^2 > 0)\right] \\ &- E\left[\frac{\partial \hat{\theta}_i^B}{\partial \sigma_u^2}\frac{\partial^2 \hat{\theta}_{j*}^B}{\partial(\sigma_u^2)^2}(\sigma_u^2)^3 I(\tilde{\sigma}_u^2 \leq 0)\right]. \end{aligned} \quad \text{(B.4)}$$

Observe that

$$
\begin{aligned}
E\left[\frac{\partial\hat{\theta}_i^B}{\partial\sigma_u^2}\frac{\partial^2\hat{\theta}_{j*}^B}{\partial(\sigma_u^2)^2}(\sigma_u^2)^3 I(\tilde{\sigma}_u^2\le 0)\right] &\le \sigma_u^6 E^{\frac{1}{4}}\left[\left\{\frac{\partial\hat{\theta}_i^B}{\partial\sigma_u^2}\right\}^4\right] E^{\frac{1}{4}}\left[\left\{\frac{\partial^2\hat{\theta}_{j*}^B}{\partial(\sigma_u^2)^2}\right\}^4\right] P^{\frac{1}{2}}(\tilde{\sigma}_u^2\le 0)\\
&\le \sigma_u^6 E^{\frac{1}{4}}\left[\left\{\frac{\partial\hat{\theta}_i^B}{\partial\sigma_u^2}\right\}^4\right] E^{\frac{1}{4}}\left[\sup_{\sigma_u^2\ge 0}\left\{\frac{\partial^2\hat{\theta}_{j}^B}{\partial(\sigma_u^2)^2}\right\}^4\right] P^{\frac{1}{2}}(\tilde{\sigma}_u^2\le 0)\\
&= o(m^{-r})
\end{aligned}
$$

for all $r>0$ by Lemma 1 (ii) and 2, which we have proved in Appendix A. Also, $P(\tilde{\sigma}_u^2\le 0)=O(m^{-r})\ \forall\, r>0$ as proved in Lemma A.6 of Prasad and Rao (1990). We next consider

$$
\begin{aligned}
E\left[\frac{\partial\hat{\theta}_i^B}{\partial\sigma_u^2}\frac{\partial^2\hat{\theta}_{j*}^B}{\partial(\sigma_u^2)^2}(\tilde{\sigma}_u^2-\sigma_u^2)^3 I(\tilde{\sigma}_u^2> 0)\right] &= E\left[\frac{\partial\hat{\theta}_i^B}{\partial\sigma_u^2}\frac{\partial^2\hat{\theta}_{j*}^B}{\partial(\sigma_u^2)^2}(\tilde{\sigma}_u^2-\sigma_u^2)^3\right]\\
&\quad - E\left[\frac{\partial\hat{\theta}_i^B}{\partial\sigma_u^2}\frac{\partial^2\hat{\theta}_{j*}^B}{\partial(\sigma_u^2)^2}(\tilde{\sigma}_u^2-\sigma_u^2)^3 I(\tilde{\sigma}_u^2\le 0)\right],
\end{aligned}
\tag{B.5}
$$

where the second term expression in (B.5) is $O(m^{-r})$ since $P(\tilde{\sigma}_u^2\le 0)=O(m^{-r})\ \forall\, r>0$. Continuing along, we next observe that

$$
\begin{aligned}
E\left[\frac{\partial\hat{\theta}_i^B}{\partial\sigma_u^2}\frac{\partial^2\hat{\theta}_{j*}^B}{\partial(\sigma_u^2)^2}(\tilde{\sigma}_u^2-\sigma_u^2)^3\right] &\le E^{\frac{1}{4}}\left[\left\{\frac{\partial\hat{\theta}_i^B}{\partial\sigma_u^2}\right\}^4\right] E^{\frac{1}{4}}\left[\left\{\frac{\partial^2\hat{\theta}_{j*}^B}{\partial(\sigma_u^2)^2}\right\}^4\right] E^{\frac{1}{2}}[(\tilde{\sigma}_u^2-\sigma_u^2)^6]\\
&\le E^{\frac{1}{4}}\left[\left\{\frac{\partial\hat{\theta}_i^B}{\partial\sigma_u^2}\right\}^4\right] E^{\frac{1}{4}}\left[\sup_{\sigma_u^2\ge 0}\left\{\frac{\partial^2\hat{\theta}_{j}^B}{\partial(\sigma_u^2)^2}\right\}^4\right] E^{\frac{1}{2}}[(\tilde{\sigma}_u^2-\sigma_u^2)^6]\\
&= O(m^{-3/2})
\end{aligned}
$$

since $E[(\tilde{\sigma}_u^2-\sigma_u^2)^{2r}]=O(m^{-r})$ for any $r\ge 1$ by Lemma A.5 in Prasad and Rao (1990). This proves that $R_1=o(m^{-1})$ since $\max_{1\le i\le m} w_i=O(m^{-1})$. By symmetry, R_2 is also $o(m^{-1})$. Finally, we show that R_3 is $o(m^{-1})$. Using a similar calculation involving R_1, we can show that

$$
E\left[\frac{\partial^2\hat{\theta}_{i*}^B}{(\partial\sigma_u^2)^2}\frac{\partial^2\hat{\theta}_{j*}^B}{\partial^2(\sigma_u^2)^2}(\hat{\sigma}_u^2-\sigma_u^2)^4\right] = E\left[\frac{\partial^2\hat{\theta}_{i*}^B}{(\partial\sigma_u^2)^2}\frac{\partial^2\hat{\theta}_{j*}^B}{\partial^2(\sigma_u^2)^2}(\tilde{\sigma}_u^2-\sigma_u^2)^4\right]+o(m^{-r}).
\tag{B.6}
$$

Observe now that

$$\begin{aligned} E\left[\frac{\partial^2\hat{\theta}^B_{i*}}{(\partial\sigma^2_u)^2}\frac{\partial^2\hat{\theta}^B_{j*}}{\partial^2(\sigma^2_u)^2}(\tilde{\sigma}^2_u-\sigma^2_u)^4\right] &\leq E^{\frac{1}{4}}\left[\left\{\frac{\partial^2\hat{\theta}^B_{i*}}{(\partial\sigma^2_u)^2}\right\}^4\right]E^{\frac{1}{4}}\left[\left\{\frac{\partial^2\hat{\theta}^B_{j*}}{\partial^2(\sigma^2_u)^2}\right\}^4\right]E^{\frac{1}{2}}\left[(\tilde{\sigma}^2_u-\sigma^2_u)^8\right] \\ &\leq E^{\frac{1}{4}}\left[\sup_{\sigma^2_u\geq 0}\left\{\frac{\partial^2\hat{\theta}^B_{i}}{(\partial\sigma^2_u)^2}\right\}^4\right]E^{\frac{1}{4}}\left[\sup_{\sigma^2_u\geq 0}\left\{\frac{\partial^2\hat{\theta}^B_{j}}{\partial^2(\sigma^2_u)^2}\right\}^4\right]E^{\frac{1}{2}}\left[(\tilde{\sigma}^2_u-\sigma^2_u)^8\right] \\ &= O(m^{-2}). \end{aligned}$$

Plugging this back into the expression in (B.6), we find that $E\left[\frac{\partial^2\hat{\theta}^B_{i*}}{(\partial\sigma^2_u)^2}\frac{\partial^2\hat{\theta}^B_{j*}}{\partial^2(\sigma^2_u)^2}(\hat{\sigma}^2_u-\sigma^2_u)^4\right]=$ $o(m^{-1})$. Hence, R_3 is $o(m^{-1})$. Finally, by calculations similar to those used for expression (B.4), we find that

$$R_0 = E\left[\frac{\partial\hat{\theta}^B_i}{\partial\sigma^2_u}\frac{\partial\hat{\theta}^B_j}{\partial\sigma^2_u}(\hat{\sigma}^2_u-\sigma^2_u)^2\right] = E\left[\frac{\partial\hat{\theta}^B_i}{\partial\sigma^2_u}\frac{\partial\hat{\theta}^B_j}{\partial\sigma^2_u}(\tilde{\sigma}^2_u-\sigma^2_u)^2\right] + o(m^{-r}).$$

Define $\Sigma = V - X(X^TV^{-1}X)^{-1}X^T = (I-P^V_X)V$, where $P_X = X(X^TV^{-1}X)^{-1}X^T$. Also, define $P^V_X = X(X^TV^{-1}X)^{-1}X^TV^{-1}$ and let $\boldsymbol{e_i}$ represent the ith unit vector. We can show $\frac{\partial\hat{\theta}^B_i}{\partial\sigma^2_u} = B_i\boldsymbol{e_i}^T\Sigma V^{-2}\tilde{\boldsymbol{u}}$, where $\tilde{\boldsymbol{u}} = \hat{\boldsymbol{\theta}} - X\tilde{\boldsymbol{\beta}}$. Define $A_{ij} = B_iB_jV^{-2}\Sigma\boldsymbol{e_i}\boldsymbol{e_j}^T\Sigma V^{-2}$and consider

$$\begin{aligned} E\left[\frac{\partial\hat{\theta}^B_i}{\partial\sigma^2_u}\frac{\partial\hat{\theta}^B_j}{\partial\sigma^2_u}(\tilde{\sigma}^2_u-\sigma^2_u)^2\right] &= E[\tilde{\boldsymbol{u}}^TA_{ij}\tilde{\boldsymbol{u}}(\tilde{\sigma}^2_u-\sigma^2_u)^2] \\ &= \text{Cov}(\tilde{\boldsymbol{u}}^TA_{ij}\tilde{\boldsymbol{u}}, (\tilde{\sigma}^2_u-\sigma^2_u)^2) + E[\tilde{\boldsymbol{u}}^TA_{ij}\tilde{\boldsymbol{u}}]E[(\tilde{\sigma}^2_u-\sigma^2_u)^2]. \end{aligned}$$

Using Lemma 3 and the relation $(I - P_X)\Sigma = (I - P_X)V$,

$$\begin{aligned}
&\mathrm{Cov}(\tilde{\boldsymbol{u}}^T A_{ij}\tilde{\boldsymbol{u}}, (\tilde{\sigma}_u^2 - \sigma_u^2)^2) \\
&\quad = (m-p)^{-2}\mathrm{Cov}(\tilde{\boldsymbol{u}}^T A_{ij}\tilde{\boldsymbol{u}}, [\tilde{\boldsymbol{u}}^T (I - P_X)\tilde{\boldsymbol{u}} - \mathrm{tr}\{(I - P_X)V\}]^2) \qquad \text{(B.7)} \\
&\quad = (m-p)^{-2}\mathrm{Cov}(\tilde{\boldsymbol{u}}^T A_{ij}\tilde{\boldsymbol{u}}, [\tilde{\boldsymbol{u}}^T (I - P_X)\tilde{\boldsymbol{u}}]^2) \\
&\quad - 2(m-p)^{-2}\mathrm{Cov}(\tilde{\boldsymbol{u}}^T A_{ij}\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{u}}^T (I - P_X)\tilde{\boldsymbol{u}})\mathrm{tr}\{(I - P_X)V\} \\
&\quad = (m-p)^{-2}\Big\{4\mathrm{tr}\{A_{ij}V(I - P_X)V\}\mathrm{tr}\{(I - P_X)V\} \\
&\quad + 8\mathrm{tr}\{A_{ij}V(I - P_X)V(I - P_X)V\} \\
&\quad - 4\mathrm{tr}\{A_{ij}V(I - P_X)V\}\mathrm{tr}\{(I - P_X)V\}\Big\} \\
&\quad = 8(m-p)^{-2}\mathrm{tr}\{A_{ij}V(I - P_X)V(I - P_X)V\}. \\
&\quad = 8(m-p)^{-2}B_iB_j\boldsymbol{e_j}^T\Sigma V^{-1}(I - P_X)V(I - P_X)V^{-1}\Sigma\boldsymbol{e_i},
\end{aligned}$$

where tr denotes the trace. Observe that $(I - P_X)V^{-1}\Sigma = I - (P_X^V)^T$ and $(I - P_X^V)V(I - (P_X^V)^T) = \Sigma$. Then

$$\begin{aligned}
\mathrm{Cov}(\tilde{\boldsymbol{u}}^T A_{ij}\tilde{\boldsymbol{u}}, (\tilde{\sigma}_u^2 - \sigma_u^2)^2) &= 8(m-p)^{-2}B_iB_j\boldsymbol{e_j}^T\Sigma V^{-1}(I - P_X)V(I - P_X)V^{-1}\Sigma\boldsymbol{e_i} \\
&= 8(m-p)^{-2}B_iB_j\boldsymbol{e_j}^T(I - P_X^V)V(I - (P_X^V)^T)\boldsymbol{e_i} \\
&= 8(m-p)^{-2}B_iB_j\boldsymbol{e_j}^T\Sigma\boldsymbol{e_i} \\
&= 8(m-p)^{-2}B_iB_j\boldsymbol{e_j}^T V\boldsymbol{e_i} + O(m^{-3}) = O(m^{-3})
\end{aligned}$$

since the first term is zero because $i \neq j$ and V is diagonal. We now calculate

$$E[\tilde{\boldsymbol{u}}^T A_{ij}\tilde{\boldsymbol{u}}] = \mathrm{tr}\{B_iB_jV^{-2}\Sigma\boldsymbol{e_i}\boldsymbol{e_j}^T\Sigma V^{-2}\Sigma\} = B_iB_j\boldsymbol{e_j}^T\Sigma V^{-2}\Sigma V^{-2}\Sigma\boldsymbol{e_i}.$$

Observe that $\Sigma V^{-2}\Sigma = I - (P_X^V)^T - P_X^V + P_X^V(P_X^V)^T$. Then after some computations, we find that $E[\tilde{\boldsymbol{u}}^T A_{ij}\tilde{\boldsymbol{u}}] = B_iB_j\boldsymbol{e_j}^T V^{-1}\boldsymbol{e_i} + O(m^{-1}) = O(m^{-1})$ since $i \neq j$. By Lemma 4, $E[(\tilde{\sigma}_u^2 - \sigma_u^2)^2] = 2(m-p)^{-2}\sum_{k=1}^m(\sigma_u^2 + D_k)^2 + O(m^{-2})$. Then

$$E[\tilde{\boldsymbol{u}}^T A_{ij}\tilde{\boldsymbol{u}}]E[(\tilde{\sigma}_u^2 - \sigma_u^2)^2] = o(m^{-1})$$

since $i \neq j$. This implies that $R_0 = o(m^{-1})$, which in turn implies that

$$E[(\hat{\theta}_i^{EB} - \hat{\theta}_i^B)(\hat{\theta}_j^{EB} - \hat{\theta}_j^B)] = o(m^{-1}) \text{ for } i \neq j, \qquad \text{(B.8)}$$

since R_0, R_1, R_2, and R_3 are all $o(m^{-1})$. Finally, this and (B.3) establishes that $E[(\bar{\hat{\theta}}_w^{EB} - \bar{\hat{\theta}}_w^B)^2] = o(m^{-1})$.

We now return to (B.1) and show that $E[(\hat{\theta}_i^{EB} - \hat{\theta}_i^{B})(\bar{\hat{\theta}}_w^{EB} - \bar{\hat{\theta}}_w^{B})] = o(m^{-1})$. By the Cauchy-Schwarz inequality, we find that

$$E[(\hat{\theta}_i^{EB} - \hat{\theta}_i^{B})(\bar{\hat{\theta}}_w^{EB} - \bar{\hat{\theta}}_w^{B})] \leq E^{\frac{1}{2}}\left[(\hat{\theta}_i^{EB} - \hat{\theta}_i^{B})^2\right] E^{\frac{1}{2}}\left[(\bar{\hat{\theta}}_w^{EB} - \bar{\hat{\theta}}_w^{B})^2\right] = o(m^{-1})$$

since the first term is $O(m^{-1/2})$ and the second term is $o(m^{-1/2})$.

For the next term of (B.1), we are interested in showing that $E[(\hat{\theta}_i^{EB} - \hat{\theta}_i^{B})(\bar{\hat{\theta}}_w^{B} - t)] = o(m^{-1})$. First, by Taylor expansion, we find that

$$\hat{\theta}_i^{EB} - \hat{\theta}_i^{B} = \frac{\partial \hat{\theta}_i^{B}}{\partial \sigma_u^2}(\hat{\sigma}_u^2 - \sigma_u^2) + \frac{1}{2}\frac{\partial^2 \hat{\theta}_{i*}^{B}}{\partial (\sigma_u^2)^2}(\hat{\sigma}_u^2 - \sigma_u^2)^2$$

for some σ_u^{*2} between σ_u^2 and $\hat{\sigma}_u^2$. Observe that $\bar{\hat{\theta}}_w^{B} - t = -\sum_i w_i B_i(\hat{\theta}_i - \boldsymbol{x}_i^T \tilde{\boldsymbol{\beta}})$. Then

$$\begin{aligned} E[(\hat{\theta}_i^{EB} - \hat{\theta}_i^{B})(\bar{\hat{\theta}}_w^{B} - t)] &= -\sum_j w_j B_j E\left[\frac{\partial \hat{\theta}_i^{B}}{\partial \sigma_u^2}(\hat{\sigma}_u^2 - \sigma_u^2)(\hat{\theta}_j - \boldsymbol{x_j}^T \tilde{\boldsymbol{\beta}})\right] \\ &\quad - \frac{1}{2}\sum_j w_j B_j E\left[\frac{\partial^2 \hat{\theta}_{i*}^{B}}{\partial (\sigma_u^2)^2}(\hat{\sigma}_u^2 - \sigma_u^2)^2(\hat{\theta}_j - \boldsymbol{x_j}^T \tilde{\boldsymbol{\beta}})\right] := R_4 + R_5. \end{aligned}$$

Observe

$$\begin{aligned} E\left[\frac{\partial \hat{\theta}_i^{B}}{\partial \sigma_u^2}(\hat{\sigma}_u^2 - \sigma_u^2)(\hat{\theta}_j - \boldsymbol{x_j}^T \tilde{\boldsymbol{\beta}})\right] &= -\sigma_u^2 E\left[\frac{\partial \hat{\theta}_i^{B}}{\partial \sigma_u^2}(\hat{\theta}_j - \boldsymbol{x_j}^T \tilde{\boldsymbol{\beta}}) I(\tilde{\sigma}_u^2 \leq 0)\right] \qquad \text{(B.9)} \\ &\quad + E\left[\frac{\partial \hat{\theta}_i^{B}}{\partial \sigma_u^2}(\tilde{\sigma}_u^2 - \sigma_u^2)(\hat{\theta}_j - \boldsymbol{x_j}^T \tilde{\boldsymbol{\beta}}) I(\tilde{\sigma}_u^2 > 0)\right] \\ &= E\left[\frac{\partial \hat{\theta}_i^{B}}{\partial \sigma_u^2}(\tilde{\sigma}_u^2 - \sigma_u^2)(\hat{\theta}_j - \boldsymbol{x_j}^T \tilde{\boldsymbol{\beta}}) I(\tilde{\sigma}_u^2 > 0)\right] + o(m^{-r}) \\ &= E\left[\frac{\partial \hat{\theta}_i^{B}}{\partial \sigma_u^2}(\tilde{\sigma}_u^2 - \sigma_u^2)(\hat{\theta}_j - \boldsymbol{x_j}^T \tilde{\boldsymbol{\beta}})\right] \\ &\quad - E\left[\frac{\partial \hat{\theta}_i^{B}}{\partial \sigma_u^2}(\tilde{\sigma}_u^2 - \sigma_u^2)(\hat{\theta}_j - \boldsymbol{x_j}^T \tilde{\boldsymbol{\beta}}) I(\tilde{\sigma}_u^2 \leq 0)\right] + o(m^{-r}) \\ &= E\left[\frac{\partial \hat{\theta}_i^{B}}{\partial \sigma_u^2}(\tilde{\sigma}_u^2 - \sigma_u^2)(\hat{\theta}_j - \boldsymbol{x_j}^T \tilde{\boldsymbol{\beta}})\right] + o(m^{-r}) \end{aligned}$$

since we may observe that $E\left[\frac{\partial \hat{\theta}_i^{B}}{\partial \sigma_u^2}(\sigma_u^2)(\hat{\theta}_j - \boldsymbol{x_j}^T \tilde{\boldsymbol{\beta}}) I(\tilde{\sigma}_u^2 \leq 0)\right] = o(m^{-r})$ and

$E\left[\frac{\partial\hat{\theta}_i^B}{\partial\sigma_u^2}(\tilde{\sigma}_u^2-\sigma_u^2)(\hat{\theta}_j-\boldsymbol{x_j}^T\tilde{\boldsymbol{\beta}})I(\tilde{\sigma}_u^2\le 0)\right]=o(m^{-r})$. Now, observe that $\frac{\partial\hat{\theta}_i^B}{\partial\sigma_u^2}=B_i\boldsymbol{e_i}^T\Sigma V^{-2}\tilde{\boldsymbol{u}}$ and define $D_{ij}=B_iV^{-2}\Sigma\boldsymbol{e_i}\boldsymbol{e_j}^T$. Then by calculations similar to those in expression (B.7), we find

$$\begin{aligned}
E\left[\frac{\partial\hat{\theta}_i^B}{\partial\sigma_u^2}(\tilde{\sigma}_u^2-\sigma_u^2)(\hat{\theta}_j-\boldsymbol{x_j}^T\tilde{\boldsymbol{\beta}})\right] &= \mathrm{Cov}(\tilde{\boldsymbol{u}}^TD_{ij}\tilde{\boldsymbol{u}},\tilde{\sigma}_u^2-\sigma_u^2)\\
&= (m-p)^{-1}\mathrm{Cov}(\tilde{\boldsymbol{u}}^TD_{ij}\tilde{\boldsymbol{u}},\tilde{\boldsymbol{u}}^T(I-P_X)\tilde{\boldsymbol{u}}-\mathrm{tr}\{(I-P_X)V\})\\
&= 2(m-p)^{-1}\mathrm{tr}\{D_{ij}V(I-P_X)V\}\\
&= 2(m-p)^{-1}\mathrm{tr}\{B_iV^{-2}\Sigma\boldsymbol{e_i}\boldsymbol{e_j}^TV(I-P_X)V\}\\
&= 2(m-p)^{-1}B_i\boldsymbol{e_j}^TV(I-P_X)V^{-1}\Sigma\boldsymbol{e_i}\\
&= 2(m-p)^{-1}B_i\boldsymbol{e_j}^TV(I-(P_X^V)^T)\boldsymbol{e_i}\\
&= 2(m-p)^{-1}B_i[\boldsymbol{e_j}^TV\boldsymbol{e_i}-h_{ij}^V]\\
&= 2(m-p)^{-1}B_i\boldsymbol{e_j}^TV\boldsymbol{e_i}+o(m^{-1}).
\end{aligned}$$

Using the expression derived above, we find that

$$\sum_j w_jB_jE\left[\frac{\partial\hat{\theta}_i^B}{\partial\sigma_u^2}(\tilde{\sigma}_u^2-\sigma_u^2)(\hat{\theta}_j-\boldsymbol{x_j}^T\tilde{\boldsymbol{\beta}})\right]=2(m-p)^{-1}B_i^2w_i(\sigma_u^2+D_i)+o(m^{-1})=o(m^{-1}).$$

Hence, R_4 is $o(m^{-1})$. We now show that $R_5=o(m^{-1})$. By calculations similar to those in expression (B.9),

$$\begin{aligned}
&\sum_j w_jB_jE\left[\frac{\partial^2\hat{\theta}_{i*}^B}{\partial(\sigma_u^2)^2}(\hat{\sigma}_u^2-\sigma_u^2)^2(\hat{\theta}_j-\boldsymbol{x_j}^T\tilde{\boldsymbol{\beta}})\right]\\
&\quad=\sum_j w_jB_jE\left[\frac{\partial^2\hat{\theta}_{i*}^B}{\partial(\sigma_u^2)^2}(\tilde{\sigma}_u^2-\sigma_u^2)^2(\hat{\theta}_j-\boldsymbol{x_j}^T\tilde{\boldsymbol{\beta}})\right]+o(m^{-r}).
\end{aligned}$$

Recall that $E\left[\left\{\sum_j w_j B_j(\hat{\theta}_j - \boldsymbol{x_j}^T\tilde{\boldsymbol{\beta}})\right\}^2\right] = O(m^{-1})$ by (B.2). Now consider

$$\begin{aligned}
&\sum_j w_j B_j E\left[\frac{\partial^2 \hat{\theta}_{i*}^B}{\partial(\sigma_u^2)^2}(\tilde{\sigma}_u^2 - \sigma_u^2)^2(\hat{\theta}_j - \boldsymbol{x_j}^T\tilde{\boldsymbol{\beta}})\right] \\
&\quad \le E^{\frac{1}{4}}\left[\left\{\frac{\partial^2 \hat{\theta}_{i*}^B}{\partial(\sigma_u^2)^2}\right\}^4\right] E^{\frac{1}{4}}\left[(\tilde{\sigma}_u^2 - \sigma_u^2)^8\right] E^{\frac{1}{2}}\left[\left\{\sum_j w_j B_j(\hat{\theta}_j - \boldsymbol{x_j}^T\tilde{\boldsymbol{\beta}})\right\}^2\right] \\
&\quad \le E^{\frac{1}{4}}\left[\left\{\sup_{\sigma_u^2 \ge 0}\frac{\partial^2 \hat{\theta}_{i}^B}{\partial(\sigma_u^2)^2}\right\}^4\right] E^{\frac{1}{4}}\left[(\tilde{\sigma}_u^2 - \sigma_u^2)^8\right] E^{\frac{1}{2}}\left[\left\{\sum_j w_j B_j(\hat{\theta}_j - \boldsymbol{x_j}^T\tilde{\boldsymbol{\beta}})\right\}^2\right] \\
&\quad = O(m^{-3/2})
\end{aligned}$$

by Lemma 1(ii), by Theorem A.5 of Prasad and Rao (1990), and by expression (B.2). Thus, R_5 is $o(m^{-1})$, and $E[(\hat{\theta}_i^{EB} - \hat{\theta}_i^B)(\bar{\hat{\theta}}_w^B - t)] = o(m^{-1})$.

For the last term in (B.1), we use the the Cauchy-Schwartz inequality to show

$$E[(\bar{\hat{\theta}}_w^{EB} - \bar{\hat{\theta}}_w^B)(\bar{\hat{\theta}}_w^B - t)] \le E^{\frac{1}{2}}[(\bar{\hat{\theta}}_w^{EB} - \bar{\hat{\theta}}_w^B)^2]E^{\frac{1}{2}}[(\bar{\hat{\theta}}_w^B - t)^2] = o(m^{-1}).$$

This concludes the proof of the theorem. □

References

Butar, F. and Lahiri, P. (2003). On measures of uncertainty of empirical Bayes small area estimators. *J. Statist. Plann. Inference*, **112** 63–76.

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

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

Fay, R. and Herriot, R. (1979). Estimates of income from small places: an application of James-Stein procedures to census data. *Journal of the American Stastical Association*, **74** 269–277.

Pfeffermann, D. and Barnard, C. (1991). Some new estimators for small area means with application to the assessment of farmland values. *Journal of Business and Economic Statistics*, **9** 31–42.

Pfeffermann, D. and Tiller, R. (2006). Small area estimation with state-space models subject to benchmark constraints. *Journal of the American Stastical Association*, **101** 1387–1397.

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

Rao, J. N. K. (2003). *Small Area Estimation.* Wiley, New York.

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

Wang, J., Fuller, W. and Qu, Y. (2008). Small area estimation under a restriction. *Survey Methodology*, **34** 29–36.

You, Y. and Rao, J. N. K. (2003). Pseudo hierarchical Bayes small area estimation combining unit level models and survey weights. *Journal of Statistical Planning and Inference*, **111** 197–208.

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

Department of Statistics, Carnegie Mellon University, Baker Hall 232K, Pittsburgh, PA 15213, U.S.A.

E-mail: (rsteorts@stat.cmu.edu)

Department of Statistics, University of Florida, P.O. Box 118545, Gainesville, FL 32611-8545, U.S.A.

E-mail: (ghoshm@stat.ufl.edu)