An Estimation of Variance of Random Effects to Solve Multiple Problems in Small Area Estimation

Masayo Y. Hirose

Institute of Mathematics for Industry, Kyushu University,

(The research was conducted with Prof. Partha Lahiri at the University of Maryland, College Park.)

[Hirose and Lahiri, 2018, AoS]

May 29, 2024

Outline

1. Small Area Estimation and Aggregated Level Model
2. Empirical Best Linear Unbiased Predictor under aggregated level model
3. A new variance component estimation for achieving desired properties
4. Monte Carlo simulation study
5. SAIPE data analysis
6. Conclusion

What is a small area estimation problem?

- Subpopulation inference is also very important, not only for the total population
- Direct estimates are constructed based only on each domain's sample data (Example: An estimation of Poverty rate: $\hat{p}_i^D = \sum_i w_{ij} y_{ij}$, where $y_{ij} \in \{0, 1\}$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n_i$.)

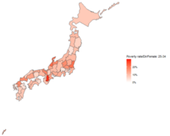


Figure: One example of Poverty mapping for Prefectures of Japan using Official microdata (Hirose and Oka in progress)

Note: The results in the analysis differ from published statistics in Japan.

- The small sample size may cause a large variation.
- refers to it as a Small area estimation problem

Aggregated level model for The Fay–Herriot Bayesian Model

There are two well-known kinds of explicit small-area models.

- Unit level model
- Aggregated level model

These models have played a critical role in the theory and practice of small-area estimation.

The unit-level model often requires unit-level data from confidential microdata.

Implementing an aggregated-level model does not tend to require confidential microdata compared with the unit-level model.

- Aggregate statistics are modeled; the chance of disclosing information about a given individual is low.
- Aggregate statistics are modeled; the relatively easier accessibility of aggregate statistics

The Fay-Herriot Bayesian Model

Fay and Herriot (1979)

For $i = 1, \cdots, m$,

Level 1: (Sampling model): $g(y_i)|\theta_i \sim N(\theta_i, D_i)$;

Level 2: (Linking model): $\theta_i \sim N(\boldsymbol{x}_i'\boldsymbol{\beta}, A)$

where

- m : number of small area;
- y_i : direct survey estimate;
- $g(y_i)$: transformed direct estimates using a smoothed monotone function g;
- θ_i : a true mean in transformed scale for area i;
- $\boldsymbol{x}_i$: p-vector of known auxiliary variables;
- D_i: *known* sampling variance of the direct estimate;
- The p-vector of regression coefficients $\boldsymbol{\beta}$ and model variance A are unknown.

Note: Hereafter, we focus on $g(\cdot) = (\cdot)$.

The Fay-Herriot Model As a Linear Mixed Model

The Fay-Herriot Bayesian model can be viewed as the following linear mixed model:

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

where

- $X = (x_1', \ldots, x_m')'$
- $u = (u_1, \cdots, u_m)'$ and $e = (e_1, \cdots, e_m)'$ are independent with $u \sim N(0, AI)$, $e \sim N(0, D)$
- I: an identity matrix of dimension m;
- $D = diag(D_1, \cdots, D_m)$

We are interested in predicting

$$\boldsymbol{\theta} = (\theta_1, \cdots, \theta_m)' = X\boldsymbol{\beta} + u,$$

where $\theta_i = \boldsymbol{x}_i'\boldsymbol{\beta} + u_i,\ i = 1, \cdots, m.$

The Best Linear Unbiased Predictor (BLUP) of θ_i

When A is known, the following BLUP of θ_i is obtained by minimizing $MSE(\hat{\theta}_i)$ among all linear unbiased predictors of θ_i, where $MSE(\hat{\theta}_i) = E[(\hat{\theta}_i - \theta_i)^2]$ and E is the expectation with respect to Fay Herriot model:

$$\hat{\theta}_i^{BLUP} = (1 - B_i)y_i + B_i \boldsymbol{x}_i' \hat{\boldsymbol{\beta}},$$

where

- $B_i \equiv B_i(A) = \frac{D_i}{A+D_i}$
- $\hat{\boldsymbol{\beta}} \equiv \hat{\boldsymbol{\beta}}(A) = (X'V^{-1}X)^{-1}X'V^{-1}y$ where $V \equiv V(A) = diag(A + D_1, \cdots, A + D_m)$.

Empirical Best Linear Unbiased Predictor (EBLUP) of θ_i

Let $\hat{A}$ be a consistent estimator of model variance parameter A for large m.

An EBLUP of θ_i is given by

$$\hat{\theta}_i^{EBLUP} = (1 - \hat{B}_i) y_i + \hat{B}_i \boldsymbol{x}_i' \hat{\boldsymbol{\beta}}.$$

where

- $\hat{B}_i = \frac{D_i}{\hat{A}+D_i}$
- $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}(\hat{A})$
- e.g., $\hat{A}$: PR estimator (Prasad and Rao, 1990), FH estimator (Fay and Herriot, 1979), ML, REML

Estimation of A: Likelihood-Based Methods

Profile Maximum Likelihood estimator (ML estimator)

$$\hat{A}_{ML} = \arg\max_{0 \le A < \infty} L_p(A|\boldsymbol{y}),$$

where

- $L_p(A, \boldsymbol{y}) = K|V|^{-1/2} \exp\{-\frac{1}{2} y' P y\}$, where K is a generic constant free from A;
- $P \equiv P(A) = V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}$.

Residual Maximum Likelihood estimator (REML estimator)

$$\hat{A}_{RE} = \arg\max_{0 \le A < \infty} L_{RE}(A|y),$$

where $L_{RE}(A|y) = h_{RE}(A) L_p(A|\boldsymbol{y})$ with $h_{RE}(A) = |X'V^{-1}(A)X|^{-1/2}$.

Remarks: Over-shrinkage problem for an estimation of B_i; $\hat{B}_i = 1$. In such case, EBLUP gets over-shrinking to the regression estimator.

Estimation of A: Likelihood-Based Methods

Adjusted Maximum Likelihood Methods for avoiding zero estimates

- Li and Lahiri (2010)
  - The Li-Lahiri adjusted ML estimator (LL.ML):

$$\hat{A}_{LL.ML} = \underset{0<A<\infty}{\arg\max}\, h_{LL}(A)L_p(A|y), \text{ where } h_{LL}(A) = A.$$

  - The Li-Lahiri adjusted REML estimator (LL.RE):

$$\hat{A}_{LL.RE} = \underset{0<A<\infty}{\arg\max}\, h_{LL}(A)L_{RE}(A|y)$$

.

Remarks:

- That is, these methods provide the following property under mild regularity conditions; $0 < \inf_{i\geq 1} \hat{B}_i(\hat{A}) \leq \sup_{i\geq 1} \hat{B}_i(\hat{A}) < 1$.
- The MSE of $\hat{A}$ is all equivalent, up to order $O(m^{-1})$. The bias of $\hat{A}_{LL.ML}$ is of order $O(m^{-1})$ that is the same as the order of $\hat{A}_{ML}$. But the bias of $\hat{A}_{LL.RE}$ is?

Estimation of A: Likelihood-Based Methods

Adjusted Maximum Likelihood Methods for avoiding zero estimates

- Yoshimori and Lahiri (2014a, JMVA)
  - The Yoshimori-Lahiri adjusted ML estimator (YL.ML):

$$\hat{A}_{YL.ML} = \underset{0<A<\infty}{\arg\max}\, h_{YL}(A) L_p(A|\boldsymbol{y}), \text{ where } h_{YL}(A) = \arctan\left[\sum_{i}^{m}(1-B_i)\right]^{1/m}.$$

  - The Yoshimori-Lahiri adjusted REML estimator (YL.RE):

$$\hat{A}_{YL.RE} = \underset{0<A<\infty}{\arg\max}\, h_{YL}(A) L_{RE}(A|\boldsymbol{y})$$

Remarks:

- That is, these methods provide the following property under mild regularity conditions; $0 < \inf_{i\geq 1} \hat{B}_i(\hat{A}) \leq \sup_{i\geq 1} \hat{B}_i(\hat{A}) < 1$.
- Not only the MSE, but also these estimators of A enjoy the same asymptotic properties of ML and REML, up to the order of $O(m^{-1})$, respectively.

Mean Squared Error (MSE) of EBLUP

The MSE of BLUP under the Fay-Herriot model is derived as,

$$MSE_i^{BLUP} \equiv MSE(\hat{\theta}_i^{BLUP}) = g_{1i}(A) + g_{2i}(A),$$

where $g_{1i}(A) = \frac{AD_i}{A+D_i}$ and $g_{2i}(A) = \frac{D_i^2}{(A+D_i)^2} x_i'(X'V^{-1}X)^{-1}x_i$.

The MSE of EBLUP under the Fay-Herriot model is approximated for large m as,

$$MSE_i^{EBLUP} \equiv MSE[\hat{\theta}_i^{EBLUP}(\hat{A})] = g_{1i}(A) + g_{2i}(A) + g_{3i}(A) + o(m^{-1}),$$

where $g_{3i}(A) = \frac{2D_i^2}{(A+D_i)^3 tr[V^{-2}]}$ and $\hat{A} \in \{\hat{A}_{ML},\ \hat{A}_{RE},\ \hat{A}_{LL.ML}, \hat{A}_{LL.RE}, \hat{A}_{YL.ML}, \hat{A}_{YL.RE}\}$.

still depends on an unknown parameter...

A second-order unbiased estimator of Mean Squared Error (MSE) for EBLUP

Definition: A second-order unbiased MSE estimator for true MSE, $\widehat{MSE}$

$\widehat{MSE}$ is satisfying that $E[\widehat{MSE} - MSE] = o(m^{-1})$ for large m.

The naive estimator: plugged $\hat{A}$ into MSE of BLUP does not satisfy.

$$E\left[MSE_i(\hat{\theta}_i^{BLUP}(A))\Big|_{A=\hat{A}} - MSE\right] = O(m^{-1}), \qquad (1)$$

where $\hat{A} \in \{\hat{A}_{ML},\ \hat{A}_{RE},\ \hat{A}_{LL.ML}, \hat{A}_{LL.RE}, \hat{A}_{YL.ML}, \hat{A}_{YL.RE}\}$.

A second-order unbiased estimator of Mean Squared Error (MSE) for EBLUP

Definition: A second-order unbiased MSE estimator for true MSE, $\widehat{MSE}$

$\widehat{MSE}$ is satisfying that $E[\widehat{MSE} - MSE] = o(m^{-1})$ for large m.

Bias correction terms are required:

1. **Taylor linearization** (Prasad and Rao, 1990; Datta and Lahiri, 2000; Datta et al., 2004; Das et al., 2004;Li and Lahiri, 2010; Yoshimori and Lahiri, 2014b)
   Using $\hat{A}_{ML}$ $/\hat{A}_{LL.ML}$ $/\hat{A}_{LL.RE}/\hat{A}_{YL.ML}$
   $$\widehat{MSE}_i \equiv \widehat{MSE}_i[\hat{\theta}_i(\hat{A})] = g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + 2g_{3i}(\hat{A}) - b(\hat{A})\hat{B}_i^2,$$
   where $b(A)$ is a bias of $\hat{A}$, up to the order $O(m^{-1})$.

   Using $\hat{A}_{RE}$/ $\hat{A}_{YL.RE}$
   $$\widehat{MSE}_i \equiv \widehat{MSE}_i[\hat{\theta}_i(\hat{A})] = g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + 2g_{3i}(\hat{A}).$$

2. **Jackknife method** (Jiang, Lahiri and Wan, 2002; Chen and Lahiri, 2008): In this presentation, we won't focus on the Jackknife method.

Parametric Bootstrap estimator of Mean Squared Error for EBLUP

3. **Parametric Bootstrap method [Single]** (Butar and Lahiri, 2003)

$$\widehat{MSE}_i^{BL} \equiv \widehat{MSE}_i[\hat{\theta}_i(\hat{A})] = 2[g_{1i}(\hat{A}) + g_{2i}(\hat{A})] - \frac{1}{B}\sum_{b=1}^{B}[g_{1i}(\hat{A}^{(b)}) + g_{2i}(\hat{A}^{(b)})] + \frac{1}{B}\sum_{b=1}^{B}[\hat{\theta}_i(y, \hat{\beta}^{(b)}, \hat{A}^{(b)}) - \hat{\theta}_i(y, \hat{\beta}^{(b)}, \hat{A}^{(b)})]^2.$$

Remark

They could be negative MSE estimates due to their bias corrections.

Parametric Bootstrap estimator of Mean Squared Error for θ_i

4. **Parametric Bootstrap method [Double]**
   (Hall and Maiti, 2006; Chatterjee and Lahiri, 2007).
   e.g., one of the estimators of Hall and Maiti (2006) is given by the following;

$$\widehat{MSE}_i^{HM1} = \begin{cases} 2\hat{u} - \hat{v} & (\hat{u} \geq \hat{v}) \\ \exp[-(\hat{v} - \hat{u})/\hat{v}]\hat{u} & (\hat{u} < \hat{v}) \end{cases}$$

where $\hat{u} = \frac{1}{B}\sum_{b=1}^{B}\left[\hat{\theta}_i^{(b)}(y^{(b)}, \hat{\beta}^{(b)}, \hat{A}^{(b)}) - \theta_i^{(b)}\right]^2$,
$\hat{v} = \frac{1}{B}\sum_{b=1}^{B}\left[\frac{1}{C}\sum_{c=1}^{C}\left[\hat{\theta}_i^{(bc)}(y^{(bc)}, \hat{\beta}^{(bc)}, \hat{A}^{(bc)}) - \theta_i^{(bc)}\right]^2\right]$.

These MSE estimators are strictly positive, but the double bootstrap method is more computer-intensive than the single bootstrap method. And not sure about the second-order unbiasedness (Jiang et al., 2016)

Research Question

What are desired properties?

- For θ_i, We need to focus on estimating the shrinkage factor B_i, rather than that of A.
- We wish to protect EBLUP from over-shrinking to the regression estimator.
- There is also a desire to use a simple second-order unbiased MSE estimator to maintain the MSE estimator's strict positivity for practical users.

Research Question

What are desired properties?

Desired properties

1. Obtain a second-order unbiased estimator of B_i;
   $E(\hat{B}_i) = B_i + o(m^{-1})$ in maintaining equivalent identical variance of other likelihood-based methods, up to the order $O(m^{-1})$.
2. $0 < \inf_{m \geq 1} \hat{B}_i \leq \sup_{m \geq 1} \hat{B}_i < 1$ for protecting EBLUP from over-shrinking to the regression estimator;
3. Obtain a simple second-order unbiased Taylor series MSE estimator of EBLUP without any bias correction; that is, $\widehat{MSE}_i = g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + g_{3i}(\hat{A})$;
4. Produce a strictly positive second-order unbiased single parametric bootstrap MSE estimator without bias correction.

$$\widehat{MSE}_i^{PB} = E_*[(\hat{\theta}_i^{EBLUP}(\hat{A}^*, y^{(*)}) - \theta_i^*)^2];$$

Research Question

Then, can we achieve these four desired properties simultaneously?

To address such an issue, we propose an area-specific estimator of A, say $\hat{A}_i$, that simultaneously satisfies these multiple desirable properties under certain mild regularity conditions.

A new adjusted maximum likelihood estimator of A

The residual maximum likelihood estimator of A is defined as:

$$\hat{A}_{RE} = \arg\max_{0 \le A < \infty} L_{RE}(A|y).$$

Note that $\hat{A}_{RE}$ does not satisfy any of the four desirable properties.

To find a likelihood-based estimator of A that satisfies all the four desirable properties, we start by setting up a general adjusted maximum likelihood estimator of A defined as:

$$\hat{A}_i = \arg\max_{0 < A < \infty} h_i(A) L_{RE}(A), \quad (2)$$

where $h_i(A)$ is not specified.

A new adjusted maximum likelihood estimator of A

We first find the adjustment factor $h_i(A)$ that satisfies Property 1.
Under the mild regularity conditions, we have, for large m,

$$E(\hat{B}_i) = B_i + \left[\frac{\partial B_i}{\partial A}\frac{\partial \log h_i(A)}{\partial A} + \frac{1}{2}\frac{\partial^2 B_i}{\partial A^2}\right]\frac{2}{tr[V^{-2}]} + o(m^{-1}).$$

Thus, Property 1 is satisfied if we have

$$\frac{\partial B_i}{\partial A}\frac{\partial \log h_i(A)}{\partial A} + \frac{1}{2}\frac{\partial^2 B_i}{\partial A^2} = 0.$$

Thus, an adequate adjustment factor is given by

$$\mathbf{h_{i0}(A) = (A + D_i)}.$$

This adjustment factor is indeed the unique solution, up to the order of $O(1)$ for large m.

A new adjusted maximum likelihood estimator of A

The resulting estimator is given by,

$$\hat{A}_i = \arg\max_{0<A<\infty} \tilde{h}_0(A) L_{RE}(A).$$

Interestingly, it turns out that such an adjusted maximum likelihood estimator also satisfies Properties 3 and 4.

"$\hat{A}_i$ satisfy Property 1, 3, 4 but not Property 2..."

We propose our final estimator of A for $m > p + 2$ as:

$$\hat{A}_{i;MG} = \arg\max_{0<A<\infty} \tilde{h}_i(A) L_{RE}(A),$$

where $\tilde{h}_i(A) = h_+(A) h_{i0}(A)$ with the additional adjustment $h_+(A)$ satisfying several conditions.
The choice of $h_+(A)$ is generally not unique. One can use the choice h_{YL} given in Yoshimori and Lahiri (2014a, JMVA).

A new adjusted maximum likelihood estimator of A

Theorem 1

Under some mild regularity conditions, we have, for large m,

$$(i) E[\hat{B}_{i;MG} - B_i] = o(m^{-1});\ Var(\hat{B}_{i;MG}) = \frac{2D_i^2}{(A+D_i)^4 tr[V^{-2}]} + o(m^{-1});$$

$$(ii)\ 0 < inf_{m\geq 1}\hat{B}_{i;MG} \leq sup_{m\geq 1}\hat{B}_{i;MG} < 1,\ for\ m > p+2;$$

$$(iii) E[\widehat{MSE}_{i;MG} - MSE_i(\hat{\theta}^{EB}_{i;MG})] = o(m^{-1});$$

$$(iv)\ E[\widehat{MSE}^{PB}_{i;MG} - MSE_i(\hat{\theta}^{EB}_{i;MG}) = o(m^{-1}),$$

where

$$\hat{B}_{i;MG} = B_i(\hat{A}_{i;MG});\quad \hat{\theta}^{EB}_{i;MG} = \hat{\theta}^{BLUP}_i(\hat{A}_{i;MG});$$

$$\widehat{MSE}_{i;MG} = g_{1i}(\hat{A}_{i;MG}) + g_{2i}(\hat{A}_{i;MG}) + g_{3i}(\hat{A}_{i;MG});$$

$$\widehat{MSE}^{PB}_{i;MG} = E_*[(\hat{\theta}_i(\hat{A}^*_{i;MG}, y^{(*)}) - \theta^*_i)^2].$$

Our approach also ensures the important dual properties of the MSE estimator — second-order unbiasedness and strict positivity.

Simulation set-up

We considered the SAIPE program of the U.S. Census Bureau to estimate the percentages of school-age children in poverty for the fifty states and the District of Columbia. (http://www.census.gov/did/www/saipe/about/index.html, Bell et al. 2015)
To compare the performances in using $\hat{A}_{RE}$ with that of $\hat{A}_{i;MG}$, we use x_i and D_i from the same SAIPE data set for the 1992 year, considered by Bell (1999).

- The 15 areas correspond to states with the largest sampling variances D_i.
- $A = 15.94$ which is the median of D_i for the 15 states.
- β: The weighted least squared estimate of β from the real data, including all 50 states and DC. ($p = 5$)

Result 1: RB and RRMSE of $\hat{B}_i$

$$\text{RB of } \hat{B}_i : \frac{\mathrm{E}(\hat{B}_i - B_i)}{B_i} \times 100;$$

$$\text{RRMSE of } \hat{B}_i : \frac{\sqrt{\mathrm{MSE}(\hat{B}_i)}}{B_i} \times 100.$$

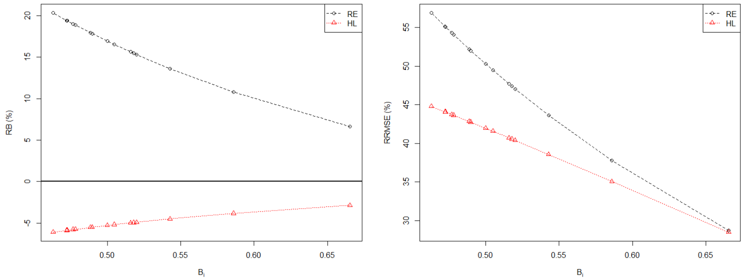
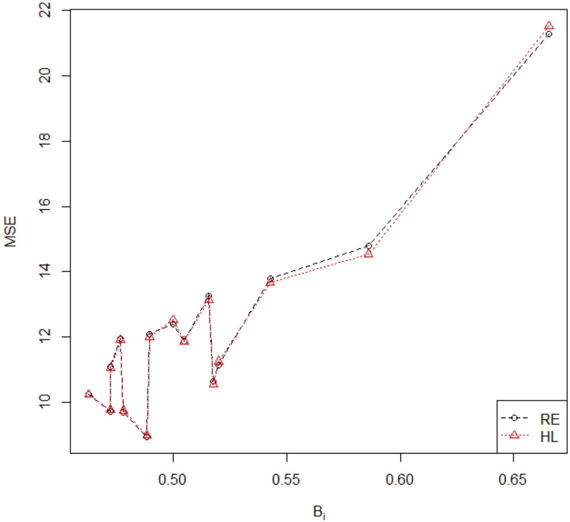


Figure: RB and RRMSE of $\hat{B}_i$

Result 2: MSE of EBLUP



We also report simulated RBs and RRMSE of different MSE estimators of EBLUPs that use $\hat{A}_{RE}$ and ours.

1. Naive MSE estimator (naive.RE): $g_{1i}(\hat{A}_{RE}) + g_{2i}(\hat{A}_{RE})$;
2. Single parametric bootstrap MSE estimator (PB.RE):
$$E_*[(\hat{\theta}_i(\hat{A}^*_{RE}, y^{(*)}) - \theta^*_i)^2];$$
3. DL.RE: $g_{1i}(\hat{A}_{RE}) + g_{2i}(\hat{A}_{RE}) + 2g_{3i}(\hat{A}_{RE})$;
4. Taylor.HL: the proposed Taylor series MSE estimator,
$$\widehat{MSE}_{i;MG} = g_{1i}(\hat{A}_{i;MG}) + g_{2i}(\hat{A}_{i;MG}) + g_{3i}(\hat{A}_{i;MG});$$
5. PB.HL: our proposed single parametric bootstrap MSE estimator,
$$\widehat{MSE}^{PB}_{i;MG} = E_*[(\hat{\theta}_i(\hat{A}^*_{i;MG}, y^{(*)}) - \theta^*_i)^2];$$
6. PB.BL:
$$2\{g_{1i}(\hat{A}_{RE}) + g_{2i}(\hat{A}_{RE})\} - E_*[g_{1i}(\hat{A}^*_{RE}) + g_{2i}(\hat{A}^*_{RE})] + E_*[\{\hat{\theta}^*_i(y_i, \hat{A}^*_{RE}, \hat{\beta}(\hat{A}^*_{RE}, y)) - \tilde{\theta}^*_i(y, \hat{A}_{RE}, \hat{\beta}(\hat{A}_{RE}, y_i))\}^2].$$

Result 3

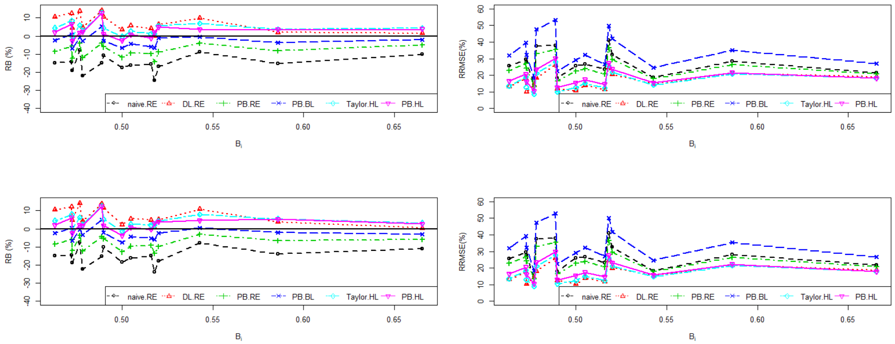


Figure: RB and RRMSE of MSE estimators for MSE of EB using REML(above) and HL(bottom); states are arranged in decreasing order of the sampling variances

Data Analysis

We consider 1992 and 1993 SAIPE data. In 1992, the REML estimate of A was zero, while in 1993, it was positive.
For this application, the small areas are 50 states and the District of Columbia of the United States, so $m = 51$.

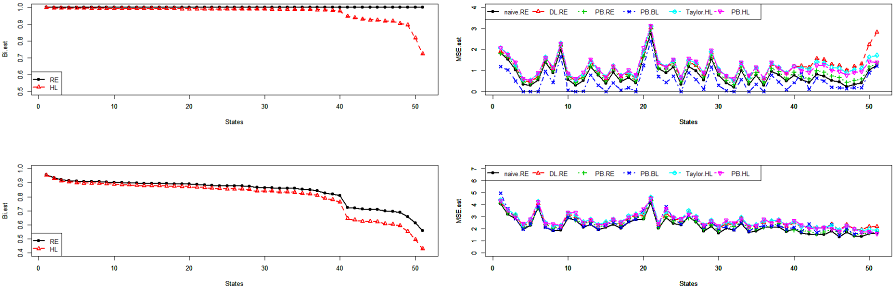


Figure: Estimates of B_i and MSE using all SAIPE data for 1992 (above) and 1993(bottom) year; states are arranged in decreasing order of the sampling variances

Conclusion

- Explanation of the basic EBLUP theory
- Proposed new variance estimator for achieving multiple goals simultaneously.
- Overall, we demonstrated that our proposed method offers reasonable results

Acknowledgment

- Supported by Grant-in-Aid for Research Activity start-up, JSPS Grant 26880011.
- Supported in part by the National Science Foundation Grant SES-1534413.

Reference

- Bell, W. R. (1999) Accounting for Uncertainty About Variances in Small Area Estimation, *Bulletin of the International Statistical Institute*, 52nd Session, Helsinki.
- Bell, William R., Basel, Wesley W., and Maples, Jerry J., (2015) "An Overview of the U.S. Census Bureau's Small Area Income and Poverty Estimates (SAIPE) Program," *in Analysis of Poverty Data by Small Area Methods*, ed. Monica Pratesi, Wiley.
- Butar, B.F., and Lahiri, P. (2003), "On measures of uncertainty of empirical Bayes small area estimators," Special issue II: Model Selection, Model Diagnostics, Empirical Bayes and Hierarchical Bayes, *Journal of Statistical Planning and Inference*, 112, 63-76.
- Chen, S., & Lahiri, P. (2008). On mean squared prediction error estimation in small area estimation problems. *Communications in Statistics-Theory and Methods*, 37(11), 1792-1798.
- Chatterjee, S., and Lahiri, P. (2007), "A simple computational method for estimating mean squared prediction error in general small-area model," *in Proceedings of the section on survey research methods, American Statistical Association Alexandria VA,* pp. 3486-3493.
- Datta, G. S., and Lahiri, P. (2000), "A unied measure of uncertainty of estimated best linear unbiased predictors in small area estimation problems," *Statistica Sinica*, 10, 613-627.
- Das, K., Jiang, J., and Rao, J.N.K. (2004), "Mean squared error of empirical predictor," *The Annals of Statistics,* 32, 818-840.
- Fay, R. E., and Herriot, R. A. (1979), "Estimates of income for small places: an application of James-Stein procedures to census data," *Journal of the American Statistical Association*, 74, 269-277.
- Hall, P., and Maiti, T. (2006), "On parametric bootstrap methods for small area prediction," *Journal of the Royal Statistical Society: Series B*, 68, 221-238.
- Hirose, M.Y. and Lahiri, P. ESTIMATING VARIANCE OF RANDOM EFFECTS TO SOLVE MULTIPLE PROBLEMS SIMULTANEOUSLY. *The Annals of Statistics*, to appear.
- Jiang, J., Lahiri, P., & Nguyen, T. (2016). A unified Monte-Carlo jackknife for small area estimation after model selection. *arXiv preprint arXiv:1602.05238*.
- Jiang, J., Lahiri, P., & Wan, S. M. (2002). A unified jackknife theory for empirical best prediction with M-estimation. *The Annals of Statistics*, 30(6), 1782-1810.
- Li, H., and Lahiri, P. (2010), "An adjusted maximum likelihood method for solving small area estimation problems," *Journal of Multivariate Analysis,* 101, 882-892.
- Prasad, N. G. N., and Rao, J. N. K. (1990), "The estimation of the mean squared error of small area estimators," *Journal of the American Statistical Association,* 85, 163-171.
- Rao, J. N. K., and Molina, I. (2015), *Small Area Estimation*, 2nd ed., Wiley, New York.
- Yoshimori, M., and Lahiri, P. (2014), "A new adjusted maximum likelihood method for the Fay–Herriot small area model," *Journal of Multivariate Analysis*, 124, 281-294.
- Yoshimori, M., and Lahiri, P. (2014b) Supplementary material to Yoshimori and Lahiri (2014), unpublished note.

Thank you for your listening!