Statistica Sinica Preprint No: SS-2020-0446	
Title	Arc-Sin Transformation for Binomial Sample Proportions in Small Area Estimation
Manuscript ID	SS-2020-0446
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202020.0446
Complete List of Authors	Masayo Hirose, Malay Ghosh and Tamal Ghosh
Corresponding Author	Masayo Hirose
E-mail	masayo@imi.kyushu-u.ac.jp
Notice: Accepted version subject to English editing.	

Statistica Sinica

Arc-Sin Transformation for Binomial Sample Proportions in Small Area Estimation

Masayo Y. Hirose*, Malay Ghosh† and Tamal Ghosh†

Kyushu University, University of Florida*†

Abstract: Arc-sin transformation has long been used as a variance-stabilizer for the binomial sample proportion arising out of binary data. The natural back-transformed function is useful for making an estimate back to the original scale of the parameter of interest. However, it is known that such a transformation leads to bias when estimating the original parameter of interest. In this study, we find explicit asymptotic bias-adjusted empirical Bayes (EB) estimators for binomial sample proportions in the context of small area estimation. We obtain explicit second-order correct approximation of mean squared errors (MSE's) of such estimators also second-order correct estimators of these MSE's. Moreover, the proposed empirical Bayes estimators and the corresponding MSE estimators have superior performance in comparison with their competitors in terms of bias and variance as demonstrated in a simulation study. We further apply our methodology to real data associated with Coronavirus Disease 2019 (COVID-19) for each prefecture in Japan.

Key words and phrases: Area Level Model; COVID-19; Linear Mixed Model; Mean Squared Error Estimation

1. Introduction

Small area estimation is receiving increasing attention in recent years both from the public and private sectors. One highly important example of the day is small area estimation of poverty and income undertaken by the United States Bureau of the Census. Federal agencies are often mandated to produce reliable estimates for small areas such as counties, census tracts and school districts. Of equal importance is to provide reliable small domain estimates cross-classified by age, sex, race and ethnicity. For more details on small area estimation, one may refer to the review articles of Ghosh and Rao (1994), Pfeffermann (2002, 2013) and the more recent book of Rao and Molina (2015).

Small area estimation can either be at the area level or at the unit level. The former is more popular than the latter since in most instances unit level data are not available to secondary users of survey data. The classic area level model is due to Fay–Herriot (1979) which is essentially a mixed effects normal linear model with the area level effect being the random effect. There are two variance components, the sampling error variance and the random effect variance. Due to non-availability of microdata, in order to avoid non-identifiability, the sampling error variance is often assumed to be known, whereas in reality, it is only an estimate.

In this paper, we are interested in the analysis of binomial sample proportions. Then the normality assumption of the original data can only be justified when sample size within an area is very large. The sampling variance in a binomial model is a function of the unknown sample mean, and can hardly be assumed to be known.

The arc-sin transformation (Anscombe, 1952; Efron and Morris, 1975) is a classical transformation which achieves the dual purpose of closer approximation to normality of the transformed data as well as a known variance.

In small area estimation, it is also important to consider situations when sample size within an area is quite small. Such situation may often occur in several research fields. For example, survey studies at early stage of pandemic in epidemiology. Even in such situations, the arc-sin transformation can very often justify the assumption of known sampling variance. Casas-Cordero, Encina and Lahiri (2015) also use this transformation for poverty mapping.

We therefore focus on this transformation and use the Fay–Herriot model for this transformed data. However, it is important to transform back properly to the original scale to arrive at the final conclusion. The natural back transformation could produce a severe bias especially when

sample size within an area is not large enough.

The arc-sin transformation, a variance stabilizing transformation, is a special case of general variable transformation. A very popular one is the log-transformation of skewed data, resulting in closer to symmetry of the transformed data, and readily implementable procedure based on the log-normal distribution. Slud and Maiti (2006), Ghosh, Kubokawa and Kawakubo (2015), and Molina and Martin (2018) took this approach, providing results for the back-transformed original parameters. In contrast, Sugasawa and Kubokawa (2017) focused on a more general dual power one to one transformation of the original data, originally proposed in Sugasawa and Kubokawa (2015).

Once the back transformation is made, Slud and Maiti (2006) proposed a multiplicative method towards this bias correction. Sugasawa and Kubokawa (2017) suggested a non-explicit empirical Bayes estimator and performed an analysis based on the general dual power transformation both in terms of bias and mean squared error.

In this paper, for arc-sin transformed data, we find an explicit empirical Bayes estimator which is also geared towards bias correction, but is more optimal in the present context than the multiplicative approach proposed by Slud and Maiti (2006) from the mean squared error (MSE) point of

view. Moreover, for evaluation of the empirical Bayes estimator, we obtain explicit second-order approximation of MSE and its second-order unbiased estimator, maintaining strict positivity.

The outline of the remaining sections is as follows. In Section 2 of this paper, we introduce empirical Bayes estimators for untransformed data and then find an explicit empirical Bayes estimator of the original parameters of interest for arc-sin transformed data. In Section 3, we analytically obtain the second-order mean squared error approximation and its explicit second-order unbiased estimator for a large number of small areas. We evaluate our method numerically in Section 4, by comparing it with other existing methods. In Section 5, we illustrate one example in predicting the positive rate in Polymerase Chain Reaction (PCR) testing for Coronavirus Disease 2019 (COVID-19) for each prefecture in Japan. All technical proofs are separately given in the supplemental file.

2. Empirical Bayes Estimation for Arc-sin Transformation

The Fay–Herriot model (1979) is a well-known area level model for small-area estimation, and it is given as follows:

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

$$\begin{aligned} &\text{Level 1} : g(y_i)|\theta_i \sim^{ind.} N(\theta_i, D_i), \\ &\text{Level 2} : \theta_i \sim^{ind.} N(x_i'\beta, A), \end{aligned} \tag{2.1}$$

where $g(\cdot)$ is a smoothed monotone function of the original data $y = (y_1, \ldots, y_m)'$ with m number of small areas. In small area estimation, y_i is referred to as the direct estimate which is obtained from data only for the i-th area. In the level 1 model, θ_i and D_i denote the true mean and the sampling variance of $g(y_i)$ respectively for each area i. The area-specific auxiliary variables $x_i = (x_{i1}, \ldots, x_{ip})'$ can be linked to θ_i for each area i, where x_i are p-dimensional vectors with $p < m$. The unknown parameters are the coefficient vector $\beta \in \mathbb{R}^p$ and the model variance parameter A. The sampling variance D_i is assumed to be known in the Fay–Herriot model (2.1) to avoid non-identifiability.

In this section, we first recall some well-known results for untransformed data, that is, $g(y_i) = y_i$. We obtain Bayes estimator $\hat{\theta}_i^B$ of θ_i, that minimizes the mean squared error (MSE) $E[(\hat{\theta}_i - \theta_i)^2]$, among all predictors $\hat{\theta}_i$, where the expectation E is defined with respect to the joint distribution of y and $\theta = (\theta_1, \ldots, \theta_m)'$:

$$\hat{\theta}_i^B \equiv \hat{\theta}_i^B(\beta, A) = (1 - B_i)y_i + B_i x_i'\beta,$$

where the shrinkage factor $B_i = D_i/A + D_i$ shrinks y_i towards $x_i'\beta$. Instead of $\hat{\theta}_i^B$, Empirical Bayes estimator (EB) of θ_i is used in practice, which replaces the unknown parameters β and A in $\hat{\theta}_i^B$ by $\hat{\beta}(\hat{A}) = (X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}y$ and some consistent estimators of A for large m. Here $X = (x_1, \ldots, x_m)'$ and $V = diag\{A + D_1, \ldots, A + D_m\}$:

$$\hat{\theta}_i^{EB} \equiv \hat{\theta}_i^{EB}(\hat{A}, \hat{\beta}) = (1 - \hat{B}_i)y_i + \hat{B}_i x_i'\hat{\beta}.$$

One can adopt the iterative moment based approach of Fay and Herriot (1979) or an explicit method of moment estimator of A as suggested in Prasad and Rao (1990). Other options are maximum likelihood (ML), residual maximum likelihood (REML) estimator of Datta and Lahiri (2000). Also, some adjusted likelihood (AL) estimators (Li and Lahiri, 2010; Yoshimori and Lahiri, 2014; Hirose and Lahiri, 2018) can be adopted. In this study, for establishing some theoretical results, we consider an estimator $\hat{A}$ of A such that

i) $\hat{A} \equiv \hat{A}(g(y))$ is an even and translation invariant for arbitrary $g(y)$ and $X\beta$ as in Kacker and Harville (1981, 1984).

ii) $\hat{A}(g(y)+\mu) = \hat{A}(g(y)) + r(g(y), \mu)$ where $r(g(y), \mu)$ is such that $E[r(g(y), \mu)^2] = O(m^{-2})$.

iii) $E[\hat{A} - A] = O(m^{-1})$ and $E[(\hat{A} - A)^8] = O(m^{-4})$ for large m.

where $g(y) = (g(y_1), \ldots, g(y_m))'$.

These conditions hold for the estimators mentioned above under certain regularity conditions.

Second-order unbiased MSE estimation has been developed by Prasad and Rao (1990), Datta and Lahiri (2000), Datta, Rao and Smith (2005) and others. For instance, Prasad and Rao (1990) obtained that the second-order approximation of MSE of $\hat{\theta}_i^{EB}$ and its second-order unbiased estimator in using the explicit moment estimator of A, namely $\hat{A}_{PR}$ as follows:

$$E[(\hat{\theta}_i^{EB}(\hat{A}_{PR}) - \theta_i)^2] = g_{1i}(A) + g_{2i}(A) + g_{3i}^{PR}(A) + o(m^{-1}), \tag{2.2}$$

$$E[g_{1i}(\hat{A}_{PR}) + g_{2i}(\hat{A}_{PR}) + 2g_{3i}^{PR}(\hat{A}_{PR})] = E[(\hat{\theta}_i^{EB}(\hat{A}) - \theta_i)^2] + o(m^{-1}),$$

where $g_{1i}(A) = D_i(1 - B_i)$, $g_{2i}(A) = B_i^2 x_i'(X'V^1X)^{-1}x_i$ and $g_{3i}^{PR}(A) = B_i^2 V_A^{PR}/(A + D_i)$ with $V_A^{PR} = 2\sum_i (A + D_i)^2/m^2$.

The residual maximum likelihood estimator of A, namely $\hat{A}_{RE}$, also has been widely used in practice. Datta and Lahiri (2000) and Das, Jiang and Rao (2004) obtained the second-order approximation of MSE of $\hat{\theta}_i^{EB}$ in using $\hat{A}_{RE}$ and its second-order unbiased estimator as follows:

$$E[(\hat{\theta}_i^{EB}(\hat{A}_{RE}) - \theta_i)^2] = g_{1i}(A) + g_{2i}(A) + g_{3i}^{DL}(A) + o(m^{-1}), \tag{2.3}$$

$$E[g_{1i}(\hat{A}_{RE}) + g_{2i}(\hat{A}_{RE}) + 2g_{3i}^{DL}(\hat{A}_{RE})] = E[(\hat{\theta}_i^{EB}(\hat{A}_{RE}) - \theta_i)^2] + o(m^{-1}),$$

where $g_{3i}^{DL}(A) = B_i^2 V_A^{DL}/(A + D_i)$ with $V_A^{DL} = 2/tr[V^{-2}]$.

We are now back to the present scenario, where responses $y_1 \ldots, y_m$ from m local areas are modeled as

$$y_i|p_i \sim^{ind.} Bin(n_i, p_i), \quad (i = 1, \ldots, m).$$

The arc-sin transformation is given by $z_i \equiv g(y_i) = \sin^{-1}(2y_i - 1)$ with the corresponding parameters $\theta_i = \sin^{-1}(2p_i - 1)$. This transformation has been employed as a variance stabilizing transformation for the direct proportion estimates as in Efron and Morris (1975).

Our interest lies in estimation of the proportion p_i. We also consider the Fay-Herriot model for the transformed data z_i, namely

$$\begin{aligned} z_i|\theta_i &\sim^{ind.} N(\theta_i, D_i), \\ \theta_i &\sim^{ind.} N(x_i'\beta, A). \end{aligned} \tag{2.4}$$

where $D_i = 1/(4n_i)$.

We now state one basic lemma and its corollary.

Lemma 1. *Let some n dimensional random vector $W_n \sim N(0, \Sigma)$ with non-singular matrix Σ and let $f(W_n)$ be some integrable function such that $f(W_n) \in \mathbb{R}$. Then*

$$(i) E[\cos(c'W_n)f(W_n)] = \frac{1}{2}\exp(-c'\Sigma c/2)\{E[f(W_n + i_u\Sigma c)] + E[f(W_n - i_u\Sigma c)]\},$$
$$(ii) E[\sin(c'W_n)f(W_n)] = \frac{1}{2i_u}\exp(-c'\Sigma c/2)\{E[f(W_n + i_u\Sigma c)] - E[f(W_n - i_u\Sigma c)]\},$$

where c denotes some n-dimensional vectors of which components are all constants and $i_u = \sqrt{-1}$ denote some constant vectors and the imaginary unit, respectively.

The above lemma is proved in Section S1.1 of the supplemental file. And it provides immediately the following corollary.

Corollary 1. *Let the n dimensional random vector $W_n \sim N(\mu, \Sigma)$ with non-singular matrix Σ. Then we have,*

$$(i) E[\cos(c'W_n)] = \exp(-c'\Sigma c/2)\cos(c'\mu),$$

$$(ii) E[\sin(c'W_n)] = \exp(-c'\Sigma c/2)\sin(c'\mu).$$

Using the model based approach, we have the posterior

$$\theta_i | z_i \sim^{ind.} N((1-B_i)z_i + B_i x_i'\beta, (4n_i)^{-1}(1-B_i)), \tag{2.5}$$

where $B_i = (4n_i)^{-1}/\{(4n_i)^{-1} + A\}$.

For notational convenience, henceforth we write $\hat{\theta}_i^B = (1-B_i)z_i + B_i x_i'\beta$. The Bayes estimator of p_i is then given by

$$\hat{p}_i^B = E[p_i|z_i] = \frac{1}{2}(1 + E[\sin\theta_i|z_i]), \ (i = 1, \ldots, m).$$

By (2.5) and Corollary 1, one gets

$$\hat{p}_i^B = \frac{1}{2}\left[1 + \exp(-g_{1i}(A)/2)\sin(\hat{\theta}_i^B)\right]. \tag{2.6}$$

To find an empirical Bayes (EB) estimator of p_i, we continue to consider the estimator of model variance parameter A mentioned in the previous section. We now propose the explicit EB estimator of p_i as follows:

$$\hat{p}_i^{EB} = \frac{1}{2}\left[1 + \exp(-g_{1i}(\hat{A})/2)\sin(\hat{\theta}_i^{EB})\right], \tag{2.7}$$

where $\hat{\theta}_i^{EB} = (1-\hat{B}_i)z_i + \hat{B}_i x_i'\hat{\beta}$ with $Z = (z_1, \ldots, z_m)'$ and $\hat{\beta} = (X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}Z$.

Hereafter, we further assume the following regularity conditions:

R1 $rank(X) = p$ is fixed for large m;

R2 $\sup_{i\geq 1} h_{ii} = O(m^{-1})$ for large m, where $h_{ii} = x_i'(X'X)^{-1}x_i$;

R3 $0 < \inf_{i\geq 1} n_i \leq \sup_{i\geq 1} n_i < \infty,\ \ 0 < A < \infty$.

One may consider using the following empirical natural back transformed predictor $\hat{p}^N$ of p_i:

$$\hat{p}_i^N(\hat{\theta}_i^{EB}) = \frac{1}{2}[1 + \sin(\hat{\theta}_i^{EB})],\ \ (i = 1, \ldots, m). \tag{2.8}$$

However, it still does not take into account of bias under the above regularity conditions. For this model, the bias corrected empirical predictor $\hat{p}_i^{SM}$ can also be obtained explicitly with the definition $\lambda = (A, x_i'\beta)$, as suggested in Slud and Maiti (2006):

$$\hat{p}_i^{SM} = \rho(\hat{\lambda})\hat{p}_i^N(\hat{\theta}_i^{EB}),\ \ (i = 1, \ldots, m), \tag{2.9}$$

where

$$\rho(\lambda) = \frac{E[\hat{p}_i^N(\theta_i)]}{E[\hat{p}_i^N(\hat{\theta}_i^B)]} = \frac{1 + \sin(x_i'\beta)\exp(-A/2)}{1 + \sin(x_i'\beta)\exp\{-A^2/(2(A + D_i))\}}.$$

For obtaining ρ, we use Corollary 1. However, this approach, unlike the lognormal case as in Slud and Maiti (2006), does not provide an optimal estimator from the MSE consideration.

3. Bias and MSE evaluations of Empirical Bayes estimator for Arc-Sin Transformation

3.1 Bias and MSE approximations of $\hat{p}_i^{EB}$

Our next objective is to evaluate asymptotic bias and obtain the second-order approximation of mean squared error (MSE) of $\hat{p}_i^{EB}$.

Here we establish first theorem for asymptotic bias and MSE of $\hat{p}_i^{EB}$.

Theorem 1. *Under the regularity conditions R1-R3, we have, for large m,*

(i) $E(\hat{p}_i^{EB} - p_i) = O(m^{-1})$,

(ii) $E[(\hat{p}_i^{EB} - p_i)^2] = M_i(\lambda) + o(m^{-1})$,

where $M_i(\lambda) = M_{1i}(\lambda) + M_{2i}(\lambda)$ with

$$
\begin{aligned}
M_{1i}(\lambda) &= \frac{1}{8}(1-\exp(-g_{1i}(A)))(1+\exp(-2A+g_{1i}(A))\cos(2x_i'\beta));\\
M_{2i}(\lambda) &= \frac{1}{8}\exp(-g_{1i}(A))\left\{g_{2i}(A)+g_{3i}(A)+\frac{B_i^4}{4}V_A\right\}\\
&\quad +\frac{1}{8}\cos(2x_i'\beta)\exp(-2A+g_{1i}(A))\left\{g_{2i}(A)+g_{3i}(A)-\frac{B_i^2(B_i-2)^2}{4}V_A\right\}
\end{aligned}
$$

In the above, $\lambda = (A, x_i'\beta)$ and $E[(\hat{A}-A)^2] = V_A + o(m^{-1})$. These proofs are given in Section S2.1 of the supplemental file. It may be noted that $\sup_i M_i$ tends to zero when $\inf_i n_i$ as well as m tend to infinity.

The above result leads to the following corollaries.

Corollary 2. (a) *If we estimate A by $\hat{A}_{PR}$, then, a second-order approximation of MSE of $\hat{p}^{EB}(\hat{A}_{PR})$ (i.e. correct up to order $O(m^{-1})$) is obtained by replacing $V(A)$ and $g_{3i}(A)$ in $M_i(\lambda)$ with $V_A^{PR}(A)$ and $g_{3i}^{PR}(A)$ respectively. We may recall that $V_A^{PR}(A)$ and $g_{3i}^{PR}(A)$ are defined in* (2.2).

(b) *If we estimate A by $\hat{A}_{RE}$, then, $V(A)$ and $g_{3i}(A)$ in the second-order approximation M_i is replaced by $V_A^{DL}(A)$ and $g_{3i}^{DL}(A)$, respectively. Also we recall that $V_A^{DL}(A)$ and $g_{3i}^{DL}(A)$ are defined in* (2.3).

3.2 MSE estimation

Next, we find an explicit form of second-order MSE estimator of $\hat{p}^{EB}(\hat{A})$. Next theorem would be a helps to construct an explicit second-order unbiased MSE estimator.

Theorem 2. *We have, for large m, under regularity conditions R1-R3,*

(i) $E[M_{1i}(\hat{\lambda}) - M_{1i}(\lambda)] = b_M(\lambda) + o(m^{-1})$,

(ii) $E[M_{1i}(\hat{\lambda}) - b_M(\hat{\lambda}) + M_{2i}(\hat{\lambda})] = M_i(\lambda) + o(m^{-1})$,

provided $\max(|\hat{M}_{1i}|, |b_M(\hat{\lambda})|, |M_{2i}(\hat{\lambda})|) < Cm^s$ with $0 < s < 1$. Otherwise, we need condition similar to Das et al. (2004; p831) as given after the equation (4.7). In the above, note that

$$\lambda = (A, \beta),\ E(\hat{A} - A) = b_A + o(m^{-1}),\ E[(\hat{A} - A)^2] = V_A + o(m^{-1});$$

and

$$\begin{aligned} b_M(\lambda) = &-\frac{1}{8}\exp(-g_{1i}(A))\left(g_{3i}(A) - b_A B_i^2 + \frac{B_i^4}{2}V_A\right) \\ &-\frac{1}{8}\exp(-2A + g_{1i}(A))\cos(2x_i'\beta)\left\{\frac{2g_{2i}(A)}{B_i^2} + g_{3i}(A) - b_A(B_i^2 - 2) - \frac{(B_i^2-2)^2}{2}V_A\right\} \\ &-\frac{1}{8}\exp(-2A)\cos(2x_i'\beta)\left(2V_A - 2b_A - \frac{2g_{2i}(A)}{B_i^2}\right). \end{aligned} \tag{3.10}$$

These proofs are deferred to Section S2.2 of the supplemental file. Also note that the regularity conditions R1-R3 are quite standard. See for example, Prasad and Rao (1990) and Datta and Lahiri (2000).

Let

$$\hat{M}_i^0(\hat{\lambda}) = M_i(\hat{\lambda}) - b_M(\hat{\lambda}), \tag{3.11}$$

where $\hat{M}_i(\lambda)$ and $b_M(\lambda)$ are given in Theorem 1 and (3.10) respectively. Theorem 2 ensures $\hat{M}_i^0(\hat{\lambda})$ is second-order unbiased.

The estimator $\hat{A}$ can be replaced with other estimators as mentioned in Section 2. From the above theorem, we prove the following corollary.

Corollary 3. (a) *If A is estimated by $\hat{A}_{PR}$ given in Prasad and Rao (1990), the explicit form of MSE estimator $\hat{M}_i^0(\hat{\lambda})$ is obtained with $\hat{A} = \hat{A}_{PR}$, $b_A = 0$, $V_A = V_A^{PR}$ and $g_{3i}(A) = g_{3i}^{PR}(A)$, respectively.*

(b) *If A is estimated by the residual maximum likelihood estimator $\hat{A}_{RE}$, given in Datta and Lahiri (2000), then the explicit form of the MSE estimator $\hat{M}_i^0(\hat{\lambda})$ is obtained with $\hat{A} = \hat{A}_{RE}$, $b_A = 0$, $V_A = V_A^{DL}$ and $g_{3i}(A) = g_{3i}^{DL}(A)$.*

(c) *One may use the adjusted residual maximum likelihood for maintaining a strict positivity of A given in Yoshimori and Lahiri (2014), denoted by $\hat{A}_{YL}$ hereafter. This form of MSE estimator is the same while using $\hat{A}_{RE}$, except $\hat{A} = \hat{A}_{YL}$.*

Alternative adjusted residual maximum likelihood estimator may change the form of the MSE estimator. For example, if we let $\hat{A}_{HL}$ as

the estimator as suggested in Hirose and Lahiri (2018), the explicit form of $\hat{M}_i^0$ *has a similar form as* $\hat{A}_{RE}$*, except that* $\hat{A} = \hat{A}_{HL}$ *and* $b_A = 2/\{tr[V^{-2}](A + D_i)\}$.

All these MSE estimators may result in negative estimates. To circumvent such problem, $\hat{M}_i^0$ is also replaced with arbitrary strictly positive value or some strictly positive estimate $\hat{M}_i^*$ when $\hat{M}_i^0$ gets negative estimates. As mentioned in Das, Jiang and Rao (2004), even for theoretical considerations, one can use $\hat{M}_i^*$ instead when the condition $|\hat{M}_i^0| < Cm^s$ does not hold with some general positive constant values C and small positive value s. For example, $M_{1i}(\hat{\lambda})$ could be adapted as M_i^* because of the fact that $0 < \hat{M}_{1i}(\hat{\lambda}) < 1/4 < \infty$ holds almost surely while either $\hat{A}_{YL}$ or $\hat{A}_{HL}$.

To maintain strict positivity of MSE, we finally suggest the MSE estimator $\hat{M}_i$ such that

$$\hat{M}_i(\hat{\lambda}) = \begin{cases} \hat{M}_i^0(\hat{\lambda}) & (0 < \hat{M}_i^0(\hat{\lambda}) < Cm^s) \\ \hat{M}_i^* & (otherwise). \end{cases} \tag{3.12}$$

where s is such that $0 < s < 3/5$.

The following theorem ensures that our estimator attains two desired properties in terms of MSE estimation, which are the second-order unbiasedness for large m and strict positivity.

Theorem 3. *Under the regularity conditions, we have for large m,*

(i) $E[\hat{M}_i - M_i] = o(m^{-1})$,

(ii) $\hat{M}_i > 0$ *with probability one,*

provided $|\hat{M}_i(\hat{\lambda})| < Cm^s$ with $0 < s < 3/5$ and generic constant C, where $M_i = E[(\hat{p}_i^{EB} - p_i)^2] + o(m^{-1})$.

The proof of Part (i) is given in Section S2.3 of the supplemental file. Proof of part (ii) is trivial.

4. Simulation Study

In this section, we implement two finite sample simulation studies in order to evaluate the performance of several back transformed empirical Bayes estimators through Monte Carlo simulation under the Fay–Herriot model (2.4).

To assess effects of the number of small areas and the sample size n within small areas, with $A = 0.125$, we set the simulation setting such that $m = 15$ and 50 and the following three patterns of sampling variances Ds1-Ds3.

(Ds1) $D_i = 1/(4n)$ for all i with $n = 2$;

(Ds2) $D_i = 1/(4n)$ for all i with $n = 4$;

(Ds3) $D_i = 1/(4n)$ for all i with $n = 8$.

We also considered three patterns of the regression coefficients $\beta \in \{-0.5, 0, 0.5\}$ for each case with fixed $x_{1i} = 1$. This setting comes from the fact that the natural back transformed empirical predictor (2.7) has bias even for a large number of small areas m when $x_i'\beta \neq 0$. For that reason, it is also important to investigate an effect on bias from $x_i'\beta$ by changing β with x_{1i} fixed at 1.

4.1 Bias and MSE

We first compare bias and uncertainty of the following six back transformed estimators of p_i for each combination of (m, β, Ds) where Ds indicates one of three patterns of sampling variances:

1) Natural Back Transformed empirical Bayes estimator $\hat{p}_i^{N.RE}$ defined in (2.8) when $\hat{A}_{RE}$ estimates A (denoted by "NBT.RE");

2) Natural Back Transformed empirical Bayes estimator $\hat{p}_i^{N.YL}$ defined in (2.8) when $\hat{A}_{YL}$ estimates A (denoted by "NBT.YL");

3) Slud-Maiti-type Bias corrected estimator $\hat{p}_i^{SM.RE}$ defined in (2.9) when $\hat{A}_{RE}$ estimates A. (denoted by "SM.RE");

4) Slud-Maiti-type Bias corrected estimator $\hat{p}_i^{SM.YL}$ defined in (2.9) when $\hat{A}_{YL}$ estimates A (denoted by"SM.YL");

5) Bias Adjusted Empirical Bayes estimator $\hat{p}_i^{EB.RE}$ suggested in (2.7) when $\hat{A}_{RE}$ estimates A (denoted by "EB.RE");

6) Bias Adjusted Empirical Bayes estimator $\hat{p}_i^{EB.YL}$ suggested in (2.7) when $\hat{A}_{YL}$ estimates A (denoted by "EB.YL");

Recall that the REML estimator $\hat{A}_{RE}$ can produce zero estimates. On the other hand, $\hat{A}_{YL}$ maintains strict positivity as given in Yoshimori and Lahiri (2014). In this simulation study, when REML solution being negative, we let REML estimates zero.

We evaluate the following simulated biases (SB) and MSEs (SMSE) using $R = 10^5$ replications. We define SB and SMSE as follows:

$$SB \equiv \frac{1}{mR}\sum_{i=1}^{m}\sum_{r=1}^{R}(\hat{p}_i^{(r)} - p_i^{(r)}),$$

$$SMSE \equiv \frac{1}{mR}\sum_{i=1}^{m}\sum_{r=1}^{R}(\hat{p}_i^{(r)} - p_i^{(r)})^2,$$

where $p_i^{(r)} = (1+\sin(\theta_i^{(r)}))/2$ is constructed using the r-th replication under the model (2.4). Furthermore, $\hat{p}_i^{(r)}$ denotes one of the above six estimators based on the r-th replication.

We display the results of the simulated biases (SB) in Figures 1 and 2 in the case $m = 15$ and 50, respectively. Each figure is consisting of three sub-figures for each sampling variance patterns Ds1-Ds3 and each x axis indicates β. These figures show that all six estimators have good performance

in terms of bias in the setting $\beta = 0$ since these simulated biases are all very close to zero. These results are reasonable because all six estimators are unbiased when $\beta = 0$. In contrast, for the other setting of $\beta \in \{-0.5, 0.5\}$, these figures also suggest that the natural back transformed empirical predictors $\hat{p}_i^N$ have relatively larger biases than others. Although the left side of Figure 1 illustrates similar performances in terms of simulated absolute bias, the decreasing simulated biases of the bias corrected estimators 3)–6) are shown as larger sample size n_i or the number of areas m from the other figures. The findings also imply that the empirical best unbiased predictors $\hat{p}_i^{EB}$ and Slud-Maiti-type of bias corrected estimators $\hat{p}_i^{SM}$ outperform the natural back transformed empirical predictors in terms of bias.

Respective Figures 3 and 4 provide the simulated MSEs (SMSE) against the combinations of β and the sampling variance patterns Ds in the simulation setting $m = 15$ and 50. We can see that these MSEs are getting smaller as sample size n getting larger in all situations, agreeing with one's intuition. The left sub-figure in Figure 3 indicates the result in the case with smallest sizes of n and m in this simulation setting. And it demonstrates superiority of the empirical Bayes estimator (especially $\hat{p}_i^{EB.YL}$) against others in terms of the MSE, whereas other cases yeild very similar performances of the six estimators.

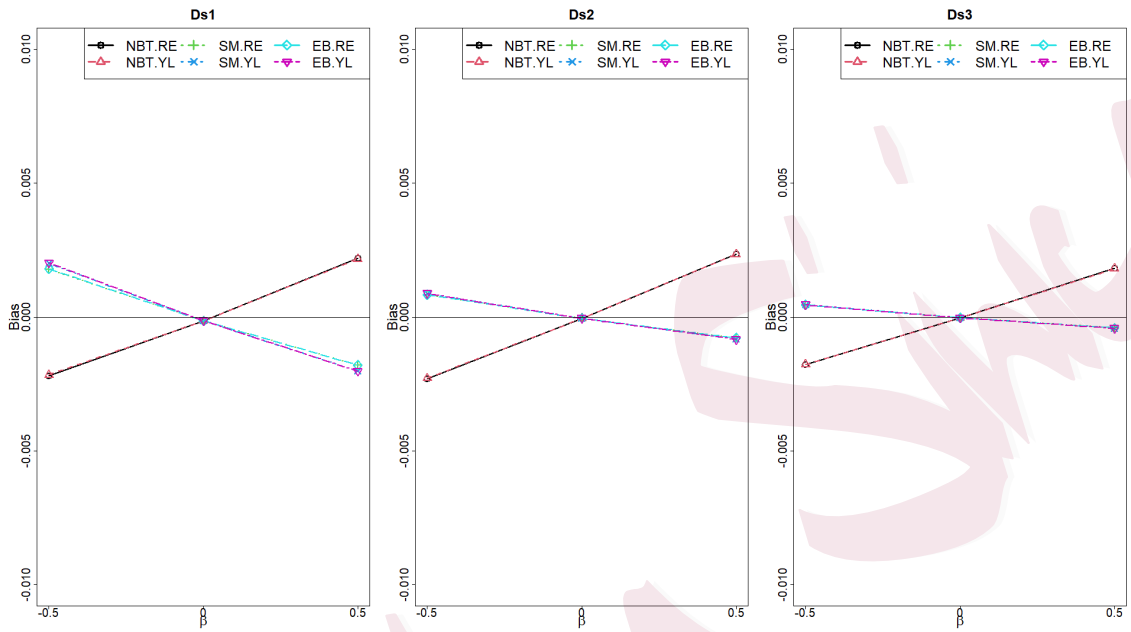

Figure 1: Simulated biases (SB) of six empirical predictors in the case $m = 15$; each sub-figure shows the results for the three sampling variance patterns $Ds1$(Left), $Ds2$(Center), $Ds3$(Right) with each x-axis indicating $\beta \in \{-0.5, 0, 0.5\}$

In summary, it seems from those simulation results that the back transformed empirical Bayes estimators $\hat{p}_i^{EB}$ seem to have best performance among all candidates in terms of bias and MSE. This is especially evident $\hat{p}_i^{EB.YL}$ based on $\hat{A}_{YL}$ outperforms than $\hat{p}_i^{EB.RE}$ using the REML method in terms of MSE.

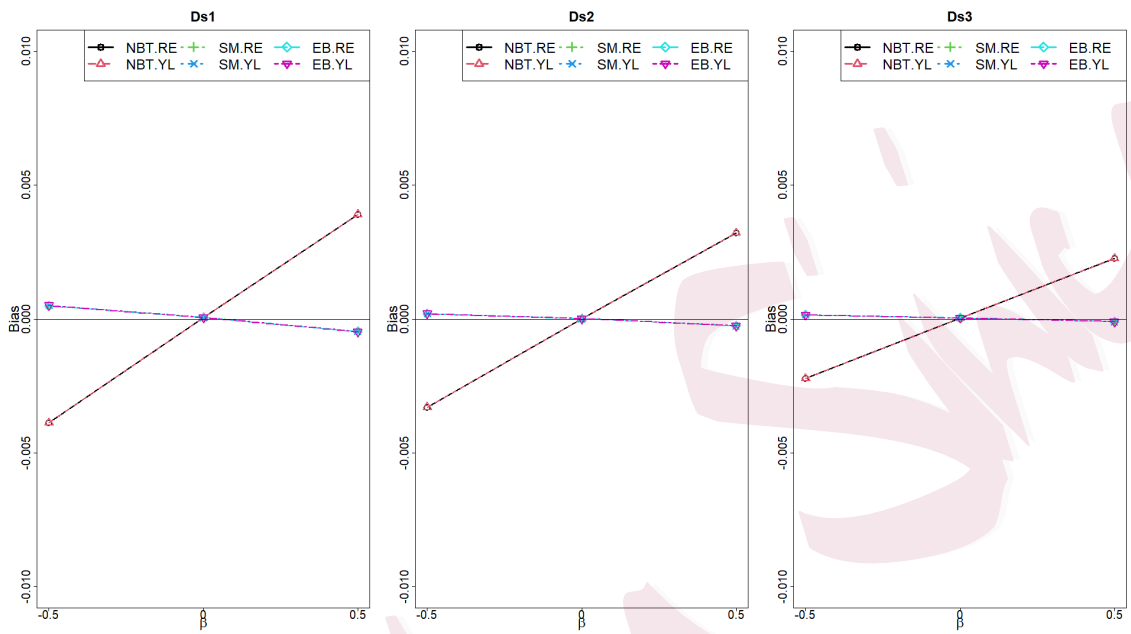

Figure 2: Simulated biases (SB) of six empirical predictors in the case $m = 50$; each sub-figure shows the results for the three sampling variance patterns $Ds1$(Left), $Ds2$(Center), $Ds3$(Right) with each x-axis indicating $\beta \in \{-0.5, 0, 0.5\}$

4.2 Assessment of MSE estimation

Next, we evaluate the efficiencies of several MSE estimators of MSE of the empirical Bayes estimators $\hat{p}_i^{EB.YL}$.

We set the same nine simulation settings of (β, Ds) for each $m = 15$ and 50 as in Section 4.1 and reconsider the following six estimators of MSE's of

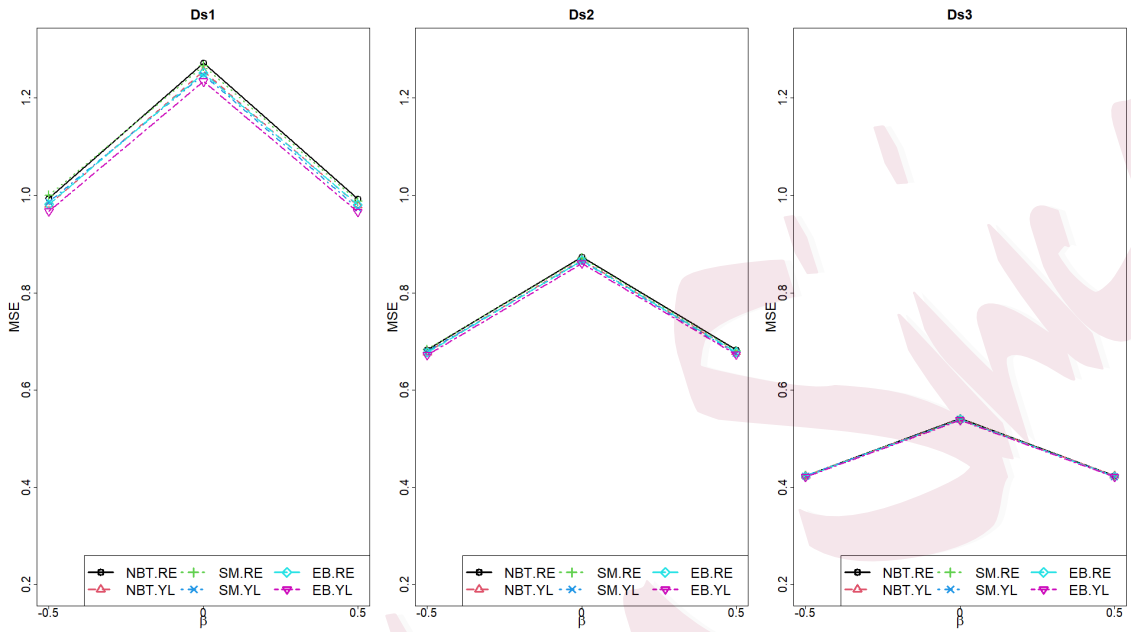

Figure 3: Simulated MSEs (SMSE) of six empirical predictors in the case $m = 15$; each sub-figure shows the results for the three sampling variance patterns $Ds1$(Left), $Ds2$(Center), $Ds3$(Right) with each x-axis indicating $\beta \in \{-0.5, 0, 0.5\}$

our estimator $\hat{p}_i^{EB.YL}$:

1) Second-order unbiased MSE estimator for untransformed data, based on $\hat{A}_{RE}$ estimator: $\hat{M}_i^{NBT.RE} \equiv g_{1i}(\hat{A}_{RE}) + g_{2i}(\hat{A}_{RE}) + 2g_{3i}(\hat{A}_{RE})$ mentioned in (2.3) (denoted by "NBT.RE");

2) Second-order unbiased MSE estimator for untransformed data, based on

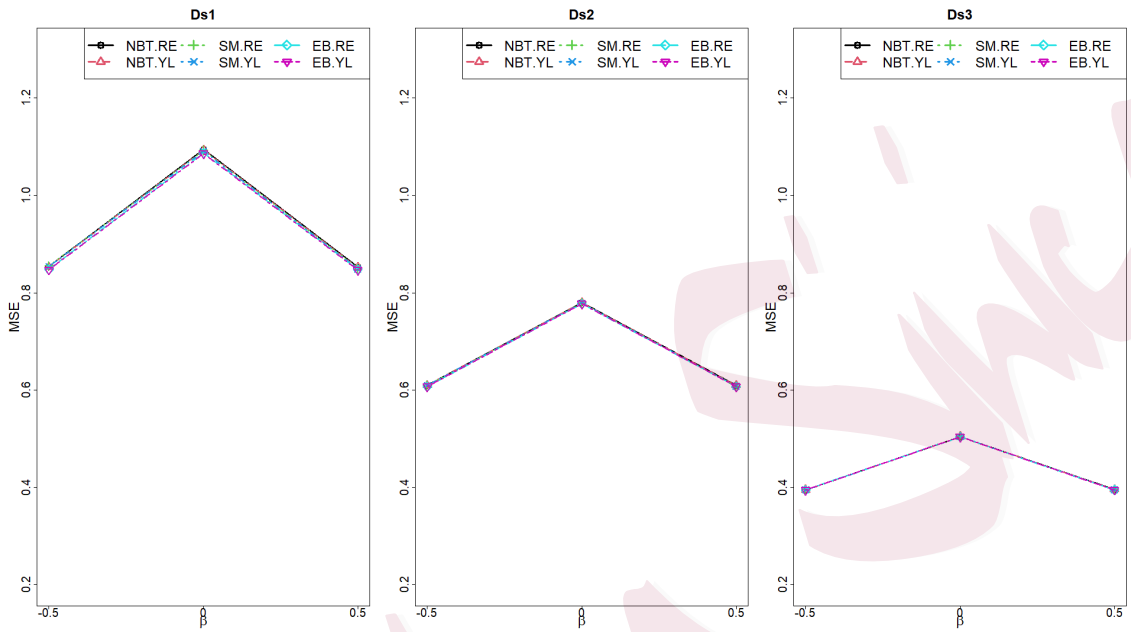

Figure 4: Simulated MSEs (SMSE) of six empirical predictors in the case $m = 50$; each sub-figure shows the results for the three sampling variance patterns $Ds1$(Left), $Ds2$(Center), $Ds3$(Right) with each x-axis indicating $\beta \in \{-0.5, 0, 0.5\}$

$\hat{A}_{YL}$ estimator: $\hat{M}_i^{NBT.RE} \equiv g_{1i}(\hat{A}_{RE}) + g_{2i}(\hat{A}_{RE}) + 2g_{3i}(\hat{A}_{YL})$ mentioned in (2.3) (denoted by "NBT.YL");

3) First-order unbiased MSE estimator $\hat{M}_{1i}^{RE} \equiv M_{1i}(\hat{\lambda})$ when adopting $\hat{A}_{RE}$, where $M_{1i}(\lambda)$ is given in Theorem 1 (denoted by "M1.RE");

4) First-order unbiased MSE estimator $\hat{M}_{1i}^{YL} \equiv M_{1i}(\hat{\lambda})$ using $\hat{A}_{YL}$, where

$M_{1i}(\lambda)$ is given in Theorem 1 (denoted by "M1.YL");

5) Second-order unbiased MSE estimator $\hat{M}_i^{RE} \equiv \hat{M}_i(\hat{\lambda})$ when using $\hat{A}_{RE}$, defined in (3.12) (denoted by "Ms.RE");

6) Second-order unbiased MSE estimator $\hat{M}_i^{YL} \equiv \hat{M}_i(\hat{\lambda})$ using $\hat{A}_{YL}$, defined in (3.12) (denoted by "Ms.YL");

For evaluation of these MSE estimators, we calculated percentages of relative bias (PRB) and the percentages of relative RMSE (PRRMSE) estimators with replication number $R = 10^5$. The PRB and PRRMSE are defined as

$$PRB = \frac{1}{mR}\sum_{i=1}^{m}\sum_{r=1}^{R}\frac{\hat{M}_i^{(r)} - M_i}{M_i} \times 100,$$

$$PRRMSE = \frac{1}{m}\sum_{i=1}^{m}\frac{\sqrt{\frac{1}{R}\sum_{r=1}^{R}(\hat{M}_i^{(r)} - M_i)^2}}{M_i} \times 100,$$

where $\hat{M}_i^{(r)}$ is one of the MSE estimators in the above using the r-th replication and M_i denotes $SMSE_i$ of $\hat{p}^{EB}$ based on $\hat{A}_{YL}$.

We report PRB and $PRRMSE$ in Figures 5–8 for comparing MSE estimators (M1.RE,M1.YL,Ms.RE,Ms.YL), whereas two untransformed MSE estimators (NBT.RE,NBT.YL) were not appearing in these figures, due to their considerable large values (over 300) for all situations. From these results, in terms of the relative bias and the relative RMSE, it is seems that

our proposed MSE estimators $\hat{M}_i(\hat{\lambda})$ are performing better than others.

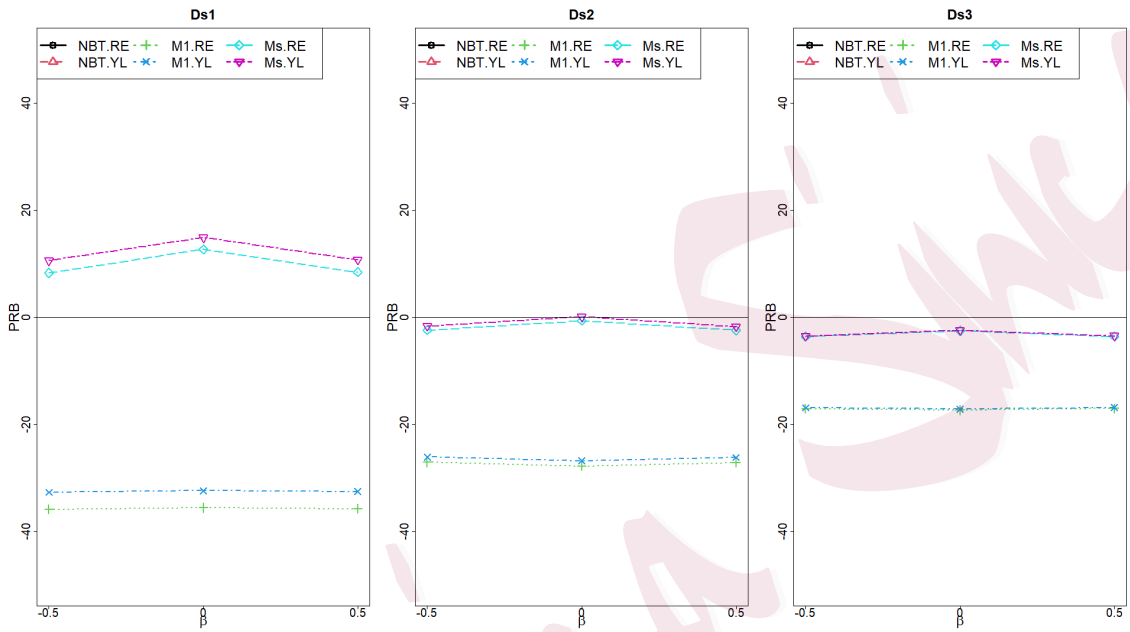

Figure 5: Percentage of relative bias (PRB) of six MSE estimators for MSE of empirical bayes estimator $\hat{p}_i^{EB.YL}$ in the case $m = 15$; each sub-figure shows the results in the three sampling variance patterns $Ds1$(Left), $Ds2$(Center), $Ds3$(Right) with each x-axis indicating $\beta \in \{-0.5, 0, 0.5\}$

In this simulation, there were no occurrence of negative estimates of MSE, but only $\hat{M}_{1i}^{RE}$ produced zero estimates. And both of our second-order unbiased MSE estimates always used $\hat{M}_i^0$ in (3.12). Finally, we report in Table 1 for the percentage of occurrence of zero estimates in using $\hat{M}_{1i}^{RE}$. The result were exactly the same as the simulated probability of REML

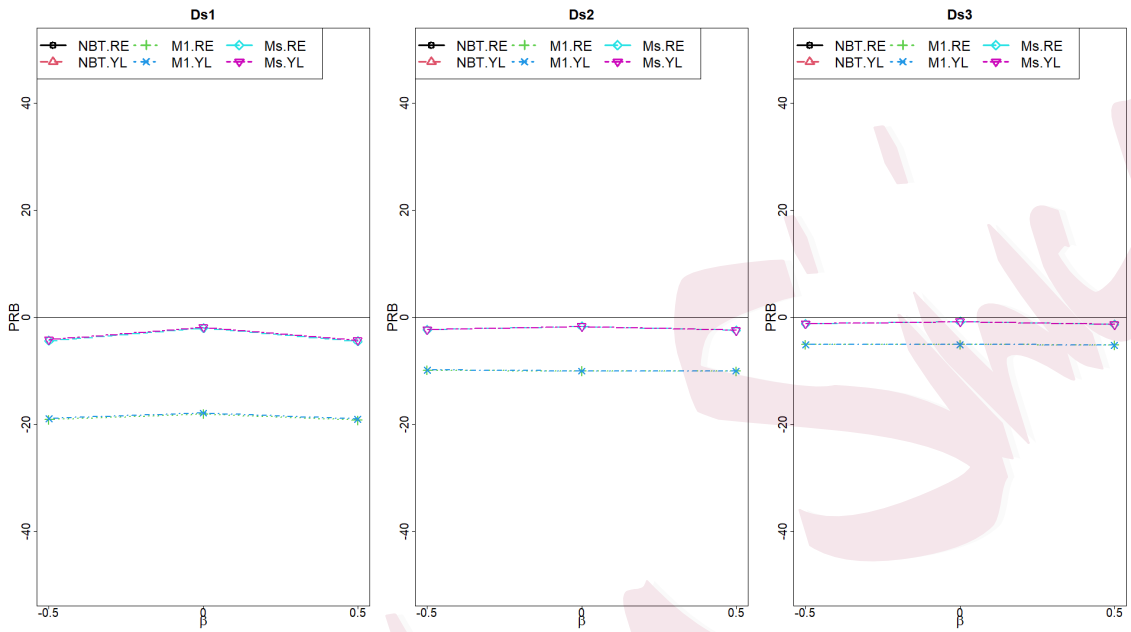

Figure 6: Percentage of relative bias (PRB) of six MSE estimators for MSE of empirical bayes estimator $\hat{p}_i^{EB.YL}$ in the case $m = 50$; each sub-figure shows the results in the three sampling variance patterns $Ds1$(Left), $Ds2$(Center), $Ds3$(Right) with each x-axis indicating $\beta \in \{-0.5, 0, 0.5\}$

being zero estimates. As a summary, it implies the estimator $\hat{M}_{1i}^{RE}$ may bring unrealistic estimates of MSE.

5. Data Analysis

COVID-19 has become a global pandemic since 2020. It is particular of interest to provide some results related to COVID-19. In this study, we

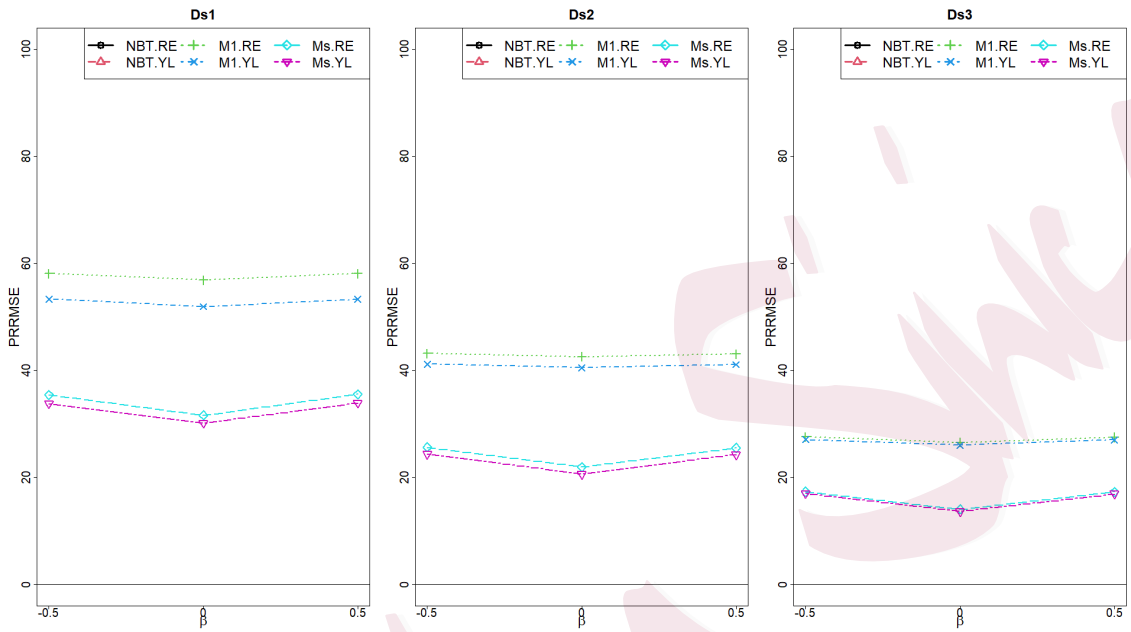

Figure 7: Percentatage of relative RMSE (PRRMSE) of six MSE estimators for MSE of empirical bayes estimator $\hat{p}_i^{EB.YL}$ in the case $m = 15$; each sub-figure shows the results in the three sampling variance patterns $Ds1$(Left), $Ds2$(Center), $Ds3$(Right) with each x-axis indicating $\beta \in \{-0.5, 0, 0.5\}$

illustrate our methodology to COVID-19 real data. The purpose of this study is, as one example, to predict the positive rate in PCR testing for each 47 prefectures in Japan.

For this purpose, we use real data having "the number of positive cases", "the number of people who have taken the PCR test" for each prefecture at date of April 21st in 2021. The data are obtained from website of Ministry of

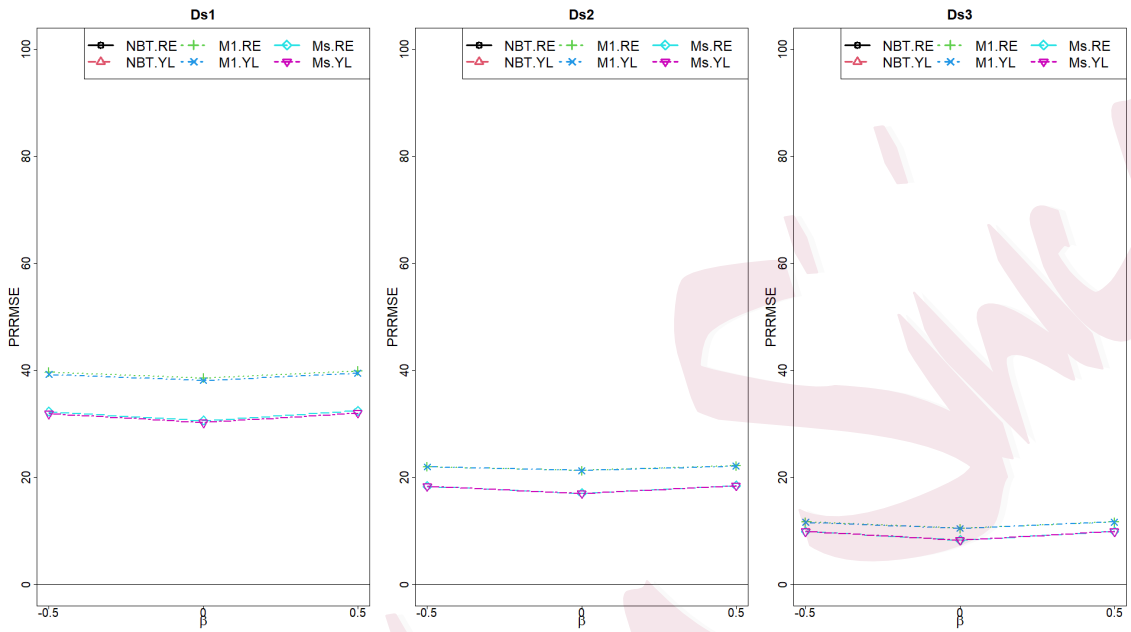

Figure 8: Percentatage of relative RMSE (PRRMSE) of six MSE estimators for MSE of empirical bayes estimator $\hat{p}_i^{EB.YL}$ in the case $m = 50$; each sub-figure shows the results in the three sampling variance patterns $Ds1$(Left), $Ds2$(Center), $Ds3$(Right) with each x-axis indicating $\beta \in \{-0.5, 0, 0.5\}$

Health, Labour and Welfare (`https://www.mhlw.go.jp/stf/covid-19/open-data.html`).

In this study, we assumed the model (2.4) and applied our methodology to this real data. Now let y_i and n_i be direct positive rate of those who took PCR test and the number of PCR test conducted for the i-th prefecture in Japan with $i \in \{1, \ldots, 47\}$. And also, the real auxiliary variable $x_i =$

Table 1: Percentage (%) of occurrence of zero estimates when $\hat{M}_{1i}^{RE}$ using in each combination case of (m, β, Ds)

Ds		$Ds1$			$Ds2$			$Ds3$		
β		-0.5	0	0.5	-0.5	0	0.5	-0.5	0	0.5
m	15	19.12	18.89	19.01	6.49	6.58	6.42	0.99	0.99	1.02
	50	3.53	3.49	3.65	0.11	0.15	0.14	0	0	0

$(1, x_{2i})'$ is used, where $x_{1i} = 1$ represents a dummy variable for the intercept term, while x_{i2} denotes $N_i \times 10^{-6}$ with population size N_i. We let N_i the population size in Japanese Census at 2015 which is the latest open census data. These are obtained from the website (`https://www.e-stat.go.jp/`). From the model assumption, the sampling variances D_i are assumed to be $1/(4 \times n_i)$ for the i-th prefectures. We call the pattern $D1$ hereafter. Moreover, we considered not only such real situation, but also tried one more hypothetical setting of sample size $n_i^* = \lceil n_i \times 10^{-4} \rceil$ while the real data (y_i, x_{2i}) used, where $\lceil n \rceil$ indicates the smallest integer greater than or equal to the value of n. Here, let D2: $D_i = 1/(4 \times n_i^*)$ be the hypothetical pattern of the sampling variance for all 47 prefectures in Japan. In this case, the range of n_i^* become from 1 to 194. We believe the situation is also important for prediction at early stages of the pandemic.

5.1 Predict Positive Rate in PCR testing

We first compare the six empirical Bayes estimates of p_i, introduced in Section 4.1, with direct proportion estimates (Direct). Figure 9 showed the result in each case of two sampling variance patterns $D1$(the left sub-figure) and $D2$(the right sub-figure). In each sub-figure, the x-axis and y-axis indicate 47 prefectures in Japan and each predict, respectively. And the resulting predicts are arranged in ascending order of the size n_i.

This left sub-figure indicates no large differences among all estimates in the real present situation. On the other hand, there seems to be considerable differences between direct estimates and other empirical Bayes estimates when sample sizes are small, as seen from the right figure. This may be due to the small sample size at early stage of the pandemic.

5.2 Estimates of Coefficients of Variation

Next, in order to more investigate the effect of MSE estimates based on several empirical Bayes estimators, we compare the following six estimators of the coefficient variation (CV):

1) $CV^{NBT.RE} \equiv \sqrt{\hat{M}_i^{NBT.RE}}/\hat{p}_i^{NBT.RE} \times 100$, constructed by the natural back transformed estimator $\hat{p}_i^{NBT.RE}$ and untransformed MSE estimator $\hat{M}_i^{NBT.RE}$, (denoted by "NBT.RE") ;

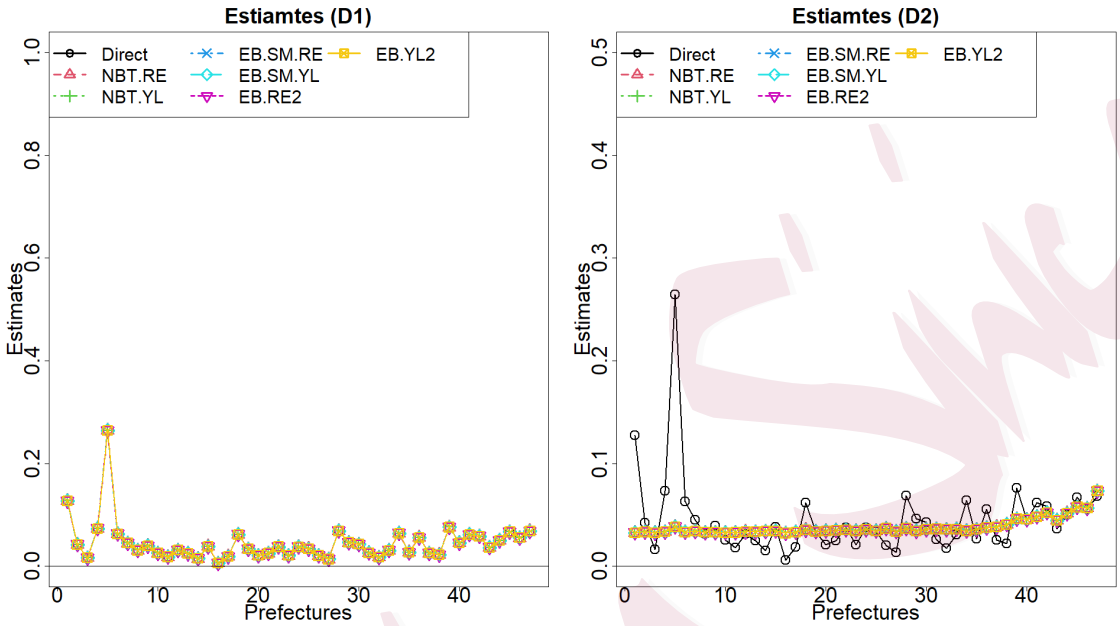

Figure 9: Seven predicted values of positive rates in PCR testing with two sampling variance patterns $D1 : D_i = 1/(4 \times n_i)$ (Left) and $D2 : D_i = 1/(4 \times n_i^*)$ (Right)

2) $CV^{NBT.YL} \equiv \sqrt{\hat{M}_i^{NBT.YL}}/\hat{p}_i^{NBT.YL} \times 100$, constructed by the natural back transformed estimator $\hat{p}_i^{NBT.YL}$ and untransformed MSE estimator $\hat{M}_i^{NBT.YL}$, (denoted by "NBT.YL");

3) $CV^{EB.RE.1} \equiv \sqrt{\hat{M}_{1i}^{RE}}/\hat{p}_i^{EB.RE} \times 100$, constructed by the bias adjusted empirical Bayes estimator $\hat{p}_i^{EB.RE}$ and the first-order unbiased MSE estimator $\hat{M}_{1i}^{RE}$, (denoted by "EB.RE.1");

4) $CV^{EB.YL.1} \equiv \sqrt{\hat{M}_{1i}^{YL}}/\hat{p}_i^{EB.YL} \times 100$, constructed by the bias adjusted empirical Bayes estimator $\hat{p}_i^{EB.YL}$ and the first-order unbiased MSE estimator $\hat{M}_{1i}^{YL}$, (denoted by "EB.YL.1");

5) $CV^{EB.RE.2} \equiv \sqrt{\hat{M}_i^{RE}}/\hat{p}_i^{EB.RE} \times 100$, constructed by the bias adjusted empirical Bayes estimator $\hat{p}_i^{EB.RE}$ and the second-order unbiased MSE estimator $\hat{M}_i^{RE}$, (denoted by "EB.RE.2");

6) $CV^{EB.YL.2} \equiv \sqrt{\hat{M}_i^{YL}}/\hat{p}_i^{EB.YL} \times 100$, constructed by the bias adjusted empirical Bayes estimator $\hat{p}_i^{EB.YL}$ and the second-order unbiased MSE estimator $\hat{M}_i^{YL}$, (denoted by "EB.YL.2").

For the estimators $CV^{EB.RE.2}$ and $CV^{EB.YL.2}$, we note that $\hat{M}_{1i}^{RE}$ and $\hat{M}_{1i}^{YL}$ are used as $\hat{M}_i^*$ in (3.12), respectively.

Figure 10 consists of four sub-figures to show the six CV estimates for all prefectures in two sampling variance patterns D1(the left two sub-figures) and D2 (the right two sub-figures). The x-axis and y-axis indicate 47 prefectures and each CV estimate, respectively. Each top and bottom sub-figure are all same except the scale changing in y-axis. The resulting estimates are arranged in ascending order of n_i as well as Figure 9.

The top left sub-figure in the sampling variance pattern D1, our result demonstrates that the overall differences of CV estimates 3)–6) from two

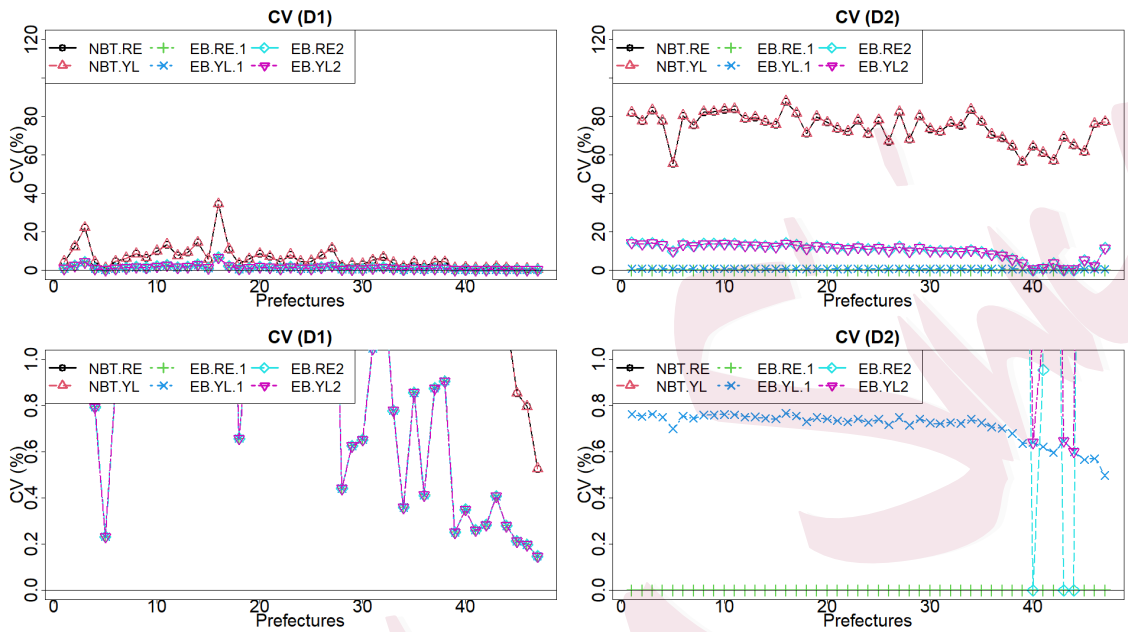

Figure 10: Six estimates of coefficient of variation (CV) of positive rate in PCR testing with two sampling variance patterns $D1 : D_i = 1/(4 \times n_i)$ (Left two figures) and $D2 : D_i = 1/(4 \times n_i^*)$ (Right two figures); Each top and bottom sub-figure are all same except the scale changing in y-axis.

naive CV estimates 1)–2) are getting smaller as the sample size n_i increases. Among CV estimates 3)–6), we cannot see large differences in the sub-figure.

We also reported the result for alternative situation with the pattern D2 in the top right sub-figure in Figure 10. The figure demonstrates that two naive CV estimates, $CV^{NBT.RE}$ and $CV^{NBT.YL}$, are much larger than others. In contrast, the estimates $CV^{EB.RE.1}$ and $CV^{EB.YL.1}$ provide much

smaller estimates than others for all prefectures. Nevertheless, these CV estimates need to be examined carefully due to small sample size. There are two reasons: first, each MSE estimator in four CV estimators 1)–4) does not have the second-order unbiasedness, and second, 2) $CV^{EB.RE.1}$ produced exactly zero CV estimates for all prefectures, as seen from the bottom of right sub-figure, in this data analysis. Incidentally, the latter unrealistic phenomenon is caused by REML estimate being zero. Also, the bottom of right sub-figure showed the $CV^{EB.RE.2}$ also provided zero estimates of CV for few prefectures although it is constructed with the second-order unbiased estimator of MSE. Note that it is not based on our final suggested MSE estimator in (3.12) since $M_i^* = M_{1i}^{RE}$ may not be strictly positive.

On the other hand, it does not report zero estimates of $CV^{EB.YL.2}$ in the bottom two figures of Figure 10. That is, in the calculation of $CV^{EB.YL.2}$ for all prefectures in this real data analysis, we used $\hat{M}_i^0$ as given in (3.12).

From the view of both theoretical and simulation results, we believe that $CV^{EB.YL.2}$ provide relatively precise and realistic CV estimates as compared to others.

6. Summary and Conclusion

In this study, we focus on arc-sin transformation for binomial sample proportions in the context of small area estimation. This specific variance stabilizing transformation avoids one important assumption of known sample variances as in the Fay–Herriot small area model. We find an explicit empirical Bayes estimator for such transformed data and then evaluate the asymptotic-order of its bias. Also, we obtain an explicit form of the second-order unbiased MSE estimator for large m based on arc-sin transformed data when $\sup_i n_i$ is bounded for large m. Moreover, we propose an explicit second-order unbiased MSE estimator which maintains strict positivity. Simulation study demonstrated the superiority of our proposed method over other methods in terms of efficiency. Furthermore, our methodology is potentially applicable to other research areas such as epidemiology, meta-analysis and others. As an example, we applied our methodology to predict the positive rate in PCR testing for each 47 prefectures in Japan. Although this is just one example to illustrate, our methodology, it may also contribute to quick studies at an early stage since the model does not require personal information in micro-data, and the estimates of the sampling variances, due to its aggregated model and variance stabilization, respectively.

Nevertheless, we may also consider another general model to treat com-

plex COVID-19 data. In the future, we intend to study further EB estimation under more general transformation models.

Supplementary Materials

All technical proofs are separetely given in the supplemental file.

Acknowledgements

The first author's research was partially supported by JSPS KAKENHI grant number 18K12758 and the ISM Cooperative Research Program (2018-ISMCRP -2057).

References

Anscombe, F. J. (1952). Large-sample theory of sequential estimation. In *Mathematical Proceedings of the Cambridge Philosophical Society*, Vol. 48, No. 4, pp. 600-607. Cambridge University Press.

Casas-Cordero, C., Encina, J. and Lahiri, P. (2015), Poverty Mapping for the Chilean Comunas, In *Analysis of Poverty Data by Small Area Estimation*, (Edited by Monica Pratesi,). Wiley Series in Survey Methodology.

Das, K., Jiang, J., and Rao, J. N. K. (2004). Mean squared error of empirical predictor. *The Annals of Statistics*, **32**, 818-840.

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

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.

Efron, B. and Morris, C. N. (1975). Data analysis using Stein's estimator and its generalizations. *Journal of the American Statistical Association* **70**: 311-319.

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

Ghosh, M., Kubokawa, T., and Kawakubo, Y. (2015). Benchmarked empirical Bayes methods in multiplicative area-level models with risk evaluation. *Biometrika*, **102**, 647-659.

Ghosh, M., and Rao, J. N. K. (1994). Small area estimation: an appraisal. *Statistical science*, **9**, 55-76.

Hirose, M. Y., and Lahiri, P. (2018). Estimating variance of random effects to solve multiple problems simultaneously. *The Annals of Statistics*, **46**, 1721-1741.

Kackar, R. N. and Harville, D. A. (1981). Unbiasedness of two-stage estimation and prediction procedures for mixed linear models. *Communications in statistics-theory and methods* **10**: 1249-1261.

Kackar, R. N., and Harville, D. A. (1984). Approximations for standard errors of estimators

of fixed and random effects in mixed linear models. *Journal of the American Statistical Association*, **79**, 853-862.

Li, H. and Lahiri, P. (2010). An adjusted maximum likelihood method for solving small area estimation problems, *Journal of Multivariate Analysis*, **101**, 882-892.

Molina, I., and Martin, N. (2018). Empirical best prediction under a nested error model with log transformation. *The Annals of Statistics*, **46**, 1961-1993.

Pfeffermann, D. (2002). Small area estimation—new developments and directions. *International Statistical Review*, **70**, 125-143.

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

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*, Wiley.

Slud, E. V., and Maiti, T. (2006). Mean-squared error estimation in transformed Fay–Herriot models. *Journal of the Royal Statistical Society: Series B*, **68**, 239-257.

Sugasawa, S., and Kubokawa, T. (2015). Parametric transformed Fay–Herriot model for small area estimation. *Journal of Multivariate Analysis*, **139**, 295-311.

Sugasawa, S., and Kubokawa, T. (2017). Transforming response values in small area prediction. *Computational Statistics & Data Analysis*, **114**, 47-60.

REFERENCES

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.

Institute of Mathematics for Industry, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka, Japan

E-mail: masayo@imi.kyushu-u.ac.jp

Department of Statistics, University of Florida, 223 Griffin-Floyd Hall, Gainesville, FL, U.S.A

E-mail: ghoshm@ufl.edu

Department of Statistics, University of Florida, 115A Griffin-Floyd Hall, Gainesville, FL, U.S.A

E-mail: tamalg@ufl.edu