Statistica Sinica **33** (2023), 705-727
doi:https://doi.org/10.5705/ss.202020.0446

ARC-SIN TRANSFORMATION FOR BINOMIAL SAMPLE PROPORTIONS IN SMALL AREA ESTIMATION

Masayo Y. Hirose, Malay Ghosh and Tamal Ghosh

Kyushu University, University of Florida and Citibank, N.A.

Abstract: The 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 returning an estimate 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 an explicit second-order correct approximation of the mean squared errors (MSEs) of such estimators, as well as second-order correct estimators of these MSEs. Moreover, the proposed EB estimators and corresponding MSE estimators outperform their competitors in terms of the bias and variance, as demonstrated in a simulation study. We 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 from both from the public and the private sectors. An important example is the 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, refer to Ghosh and Rao (1994), Pfeffermann (2002, 2013), and Rao and Molina (2015).

Small area estimation can either be at the area level or at the unit level. The former is more popular, because in most instances, unit level data are not available to secondary users of survey data. The classic area level model is attributed to Fay and Herriot (1979), and is essentially a mixed effects normal linear model, with

Corresponding author: Masayo Y. Hirose, Institute of Mathematics for Industry, Kyushu University, Nishi-ku, Fukuoka, Japan. E-mail: masayo@imi.kyushu-u.ac.jp.

the area level effect being the random effect. There are two variance components, the sampling error variance and the random effect variance. Owing to the 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.

We are interested in the analysis of binomial sample proportions. Then, the normality assumption of the original data can only be justified when the 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 that achieves the dual purpose of a closer approximation to normality of the transformed data and a known variance.

In small area estimation, it is important to consider situations in which the sample size within an area is quite small, as is common in several research fields, such as survey studies at the early stage of a pandemic in epidemiology. Even in such situations, the arc-sin transformation can very often justify the assumption of a 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. It is important to transform back properly to the original scale to arrive at the final conclusion. However, the natural back transformation could produce a severe bias, especially when the sample size within an area is not sufficiently large.

The arc-sin transformation, a variance-stabilizing transformation, is a special case of the general variable transformation. A popular choice is the log-transformation of skewed data, resulting in more symmetrical transformed data, and a readily implementable procedure based on the log-normal distribution. Slud and Maiti (2006), Ghosh, Kubokawa and Kawakubo (2015), and Molina and Martin (2018) adopted 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 for this bias correction. Sugasawa and Kubokawa (2017) suggested a non-explicit empirical Bayes (EB) estimator, and performed an analysis based on the general dual-power transformation in terms of the bias and the mean squared error (MSE).

In this study, for arc-sin transformed data, we find an explicit EB estimator that is also geared toward bias correction, but is more optimal in the present

context than is the multiplicative approach proposed by Slud and Maiti (2006) from an MSE point of view. Moreover, for the evaluation of the EB estimator, we obtain an explicit second-order approximation of the MSE and its second-order unbiased estimator, maintaining strict positivity.

The remainder of the paper proceeds as follows. In Section 2, we introduce EB estimators for untransformed data and find an explicit EB estimator of the original parameters of interest for arc-sin transformed data. In Section 3, we analytically obtain the second-order MSE 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 the proposed method by 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 relegated to the online Supplementary Material.

2. EB Estimation for Arc-sin Transformation

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

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

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

where $g(\cdot)$ is a smoothed monotone function of the original data $y = (y_1, \ldots, y_m)'$, with m small areas. In small area estimation, y_i is referred to as the direct estimate, and is obtained from data for the ith area only. In the level-1 model, θ_i and D_i denote the true mean and the sampling variance, respectively, of $g(y_i)$ 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 a Bayes estimator $\hat{\theta}_i^B$ of θ_i that minimizes the 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 toward $x_i'\beta$. Instead of $\hat{\theta}_i^B$, EB estimator of θ_i is used in practice, which replaces the unknown parameters β and A in $\hat{\theta}_i^B$ with $\hat{\beta}(\hat{A}) = (X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}y$ and some consistent estimators of A, respectively 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 a moment estimator of A, as suggested in Prasad and Rao (1990). Other options are the maximum likelihood (ML), the residual maximum likelihood (REML) estimator of Datta and Lahiri (2000), and some adjusted likelihood (AL) estimators (Li and Lahiri (2010); Yoshimori and Lahiri (2014); Hirose and Lahiri (2018)). In this study, to establish our theoretical results, we consider an estimator $\hat{A}$ of A such that

i) $\hat{A} \equiv \hat{A}(g(y))$ is even and translation invariant for arbitrary $g(y)$ and $X\beta$, as in Kackar 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 was developed by Prasad and Rao (1990), Datta and Lahiri (2000), Datta, Rao and Smith (2005), and others. For instance, Prasad and Rao (1990) obtained the second-order approximation of the MSE of $\hat{\theta}_i^{EB}$ and its second-order unbiased estimator using the explicit moment estimator of A, namely $\hat{A}_{PR}$, as follows:

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

respectively, 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 $\hat{A}_{RE}$, is also widely used in practice. Datta and Lahiri (2000) and Das, Jiang and Rao (2004) obtained the second-order approximation of the MSE of $\hat{\theta}_i^{EB}$ using $\hat{A}_{RE}$, and its second-order unbiased

estimator, as follows:

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

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

In the present scenario, the responses $y_1 \ldots, y_m$ from m local areas are modeled as

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

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 direct-proportion estimates, as in Efron and Morris (1975).

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

$$z_i|\theta_i \overset{ind.}{\sim} N(\theta_i, D_i),$$
$$\theta_i \overset{ind.}{\sim} N(x_i'\beta, A). \quad (2.4)$$

where $D_i = 1/n_i$.

We now state a basic lemma and its corollary.

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

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

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

The above lemma is proved in Section S1.1 of the Supplementary Material, and leads immediately to the following corollary.

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

$$(i)\ E[\cos(c'W_n)] = \exp\left(-\frac{c'\Sigma c}{2}\right)\cos(c'\mu),$$
$$(ii)\ E[\sin(c'W_n)] = \exp\left(-\frac{c'\Sigma c}{2}\right)\sin(c'\mu).$$

Using the model-based approach, we have the posterior

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

where $B_i = 1/(1+n_iA)$.
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, we have

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

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

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

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 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, this still does not take bias into account under the above regularity conditions. For this model, the bias-corrected empirical predictor $\hat{p}_i^{SM}$ can be obtained explicitly using 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

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

To obtain ρ, we use Corollary 1. However, this approach, unlike the log-normal case used in Slud and Maiti (2006), does not provide an optimal estimator from an MSE perspective.

3. Bias and MSE evaluations of the EB estimator for the Arc-Sin Transformation

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

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

Here, we establish first a theorem for the asymptotic bias and the 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*

$$M_{1i}(\lambda) = \frac{1}{8}(1 - \exp(-g_{1i}(A)))(1 + \exp(-2A + g_{1i}(A))\cos(2x_i'\beta));$$

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

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 Supplementary Material. Note that $\sup_i M_i$ tends to zero when $\inf_i n_i$ and 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 the 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. 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 are replaced by $V_A^{DL}(A)$ and $g_{3i}^{DL}(A)$, respectively. Furthermore, 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 the second-order MSE estimator of $\hat{p}^{EB}(\hat{A})$. The following theorem 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, Jiang and Rao (2004, p.831)), as given after equation (4.7). In the above, note that*

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

and

$$\begin{aligned} b_M(\lambda) = &- \frac{1}{8}\exp(-g_{1i}(A))\left(g_{3i}(A) - b_A B_i^2 + \frac{B_i^4}{2}V_A\right) \\ &- \frac{1}{8}\exp(-2A + g_{1i}(A))\cos(2x_i'\beta) \\ &\left\{\frac{2g_{2i}(A)}{B_i^2} + g_{3i}(A) - b_A(B_i^2 - 2) - \frac{(B_i^2 - 2)^2}{2}V_A\right\} \\ &- \frac{1}{8}\exp(-2A)\cos(2x_i'\beta)\left(2V_A - 2b_A - \frac{2g_{2i}(A)}{B_i^2}\right). \end{aligned} \tag{3.1}$$

These proofs are deferred to Section S2.2 of the Supplementary Material. Note too that the regularity conditions R1–R3 are quite standard. See, for example, Prasad and Rao (1990) and Datta and Lahiri (2000).

Let

$$\hat{M}_i^0(\hat{\lambda}) = M_i(\hat{\lambda}) - b_M(\hat{\lambda}), \tag{3.2}$$

where $\hat{M}_i(\lambda)$ and $b_M(\lambda)$ are given in Theorem 1 and (3.1), 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 the 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 to maintain the strict positivity of A given in Yoshimori and Lahiri (*2014*), denoted by $\hat{A}_{YL}$ hereafter. This form of the MSE estimator is the same as that using $\hat{A}_{RE}$, except $\hat{A} = \hat{A}_{YL}$.*

*An alternative adjusted residual maximum likelihood estimator may change the form of the MSE estimator. For example, if we let $\hat{A}_{HL}$ be the estimator, as suggested in Hirose and Lahiri (*2018*), the explicit form of $\hat{M}_i^0$ is similar in form to that of $\hat{A}_{RE}$, except that $\hat{A} = \hat{A}_{HL}$ and $b_A = 2/\{tr[V^{-2}](A+D_i)\}$.*

All of these MSE estimators may result in negative estimates. To circumvent this problem, $\hat{M}_i^0$ is replaced with an arbitrary strictly positive value or some strictly positive estimate $\hat{M}_i^*$ when $\hat{M}_i^0$ results in 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 value C and a small positive value s. For example, $M_{1i}(\hat{\lambda})$ could be adapted as M_i^* because $0 < \hat{M}_{1i}(\hat{\lambda}) < 1/4 < \infty$ holds almost surely with either $\hat{A}_{YL}$ or $\hat{A}_{HL}$.

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

$$\hat{M}_i(\hat{\lambda}) = \begin{cases} \hat{M}_i^0(\hat{\lambda}) \ (0 < \hat{M}_i^0(\hat{\lambda}) < Cm^s) \\ \hat{M}_i^* \qquad (otherwise), \end{cases} \tag{3.3}$$

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

The following theorem ensures that our estimator attains two desired properties in terms of the MSE estimation, namely, 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 Supplementary Material. The proof of part (ii) is trivial.

4. Simulation Study

In this section, we implement two finite-sample simulation studies to evaluate the performance of several back-transformed EB estimators using a Monte Carlo simulation under the Fay–Herriot model (2.4).

To assess the 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 with the following three patterns of sampling variances Ds1–Ds3:

(Ds1) $D_i = 1/n$, for all i, with $n = 8$;

(Ds2) $D_i = 1/n$, for all i, with $n = 16$;

(Ds3) $D_i = 1/n$, for all i, with $n = 32$.

We also considered three patterns of 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 $x_i'\beta \neq 0$. Therefore, it is also important to investigate the effect on the bias after changing β with x_{1i}, fixed at one.

4.1. Bias and MSE

We first compare the 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 EB estimator $\hat{p}_i^{N.RE}$ defined in (2.8) when $\hat{A}_{RE}$ estimates A (denoted by "NBT.RE");

2) Natural back-transformed EB 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 EB estimator $\hat{p}_i^{EB.RE}$ suggested in (2.7) when $\hat{A}_{RE}$ estimates A (denoted by "EB.RE");

6) Bias-adjusted EB 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 shown in Yoshimori and Lahiri (2014). In this simulation study, when the REML solution is negative, we let the REML estimate be 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 rth replication under model (2.4). Furthermore, $\hat{p}_i^{(r)}$ denotes one of the above six estimators based on the rth 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 comprises three sub-figures for the sampling variance patterns Ds1–Ds3, and each x-axis indicates β. These figures show that all six estimators exhibit good performance in terms of bias in the setting $\beta = 0$, because the 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 show that the natural back-transformed empirical predictors $\hat{p}_i^N$ have relatively larger biases. Although the left side of Figure 1 illustrates similar performance in terms of the

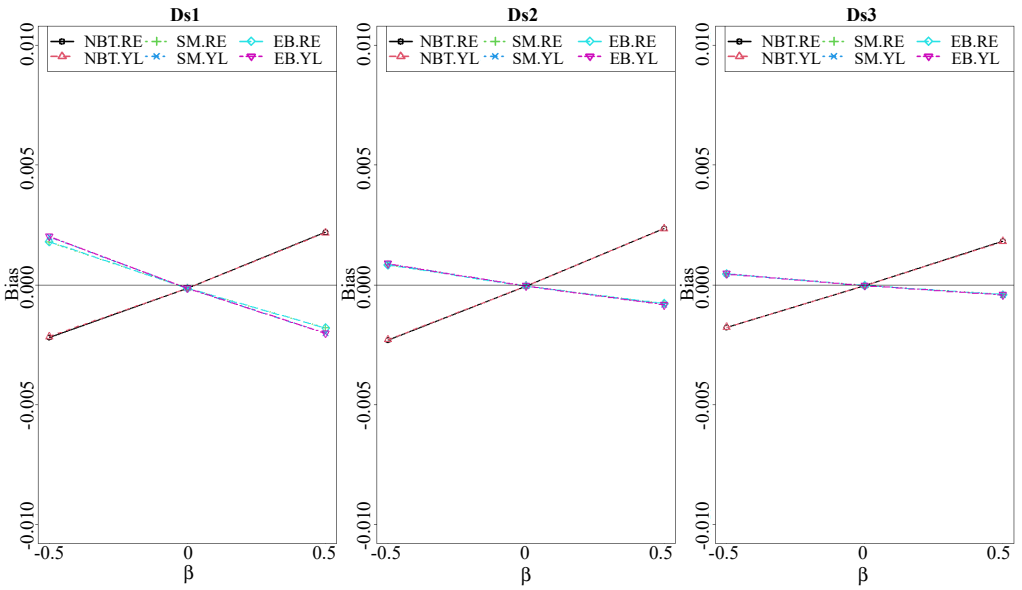

Figure 1. Simulated biases (SB) of six empirical predictors in the case $m = 15$; each sub-figure shows the results for the three sampling variance patterns $Ds1$(Left), $Ds2$(Center), and $Ds3$(Right), with each x-axis indicating $\beta \in \{-0.5, 0, 0.5\}$.

simulated absolute bias, the simulated biases of the bias-corrected estimators 3)–6) decrease as the sample size n_i and the number of areas m increase. These findings imply that the bias-adjusted EB estimators $\hat{p}_i^{EB}$ and the Slud–Maiti-type of bias-corrected estimators $\hat{p}_i^{SM}$ outperform the natural back-transformed empirical predictors in terms of bias.

Figures 3 and 4 show the SMSEs 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 decrease as the sample size n increases in all situations, agreeing with one's intuition. The left sub-figure in Figure 3 indicates the result for the smallest sizes of n and m in this simulation setting. It demonstrates the superiority of the EB estimator (especially $\hat{p}_i^{EB.YL}$) in terms of the MSE, whereas the other cases yield very similar performance for the six estimators.

In summary, the simulation results show that the back-transformed EB estimators $\hat{p}_i^{EB}$ have the best performance of all the candidates in terms of the bias and MSE. This is especially evident where $\hat{p}_i^{EB.YL}$ based on $\hat{A}_{YL}$ outperforms $\hat{p}_i^{EB.RE}$ using the REML method in terms of the MSE.

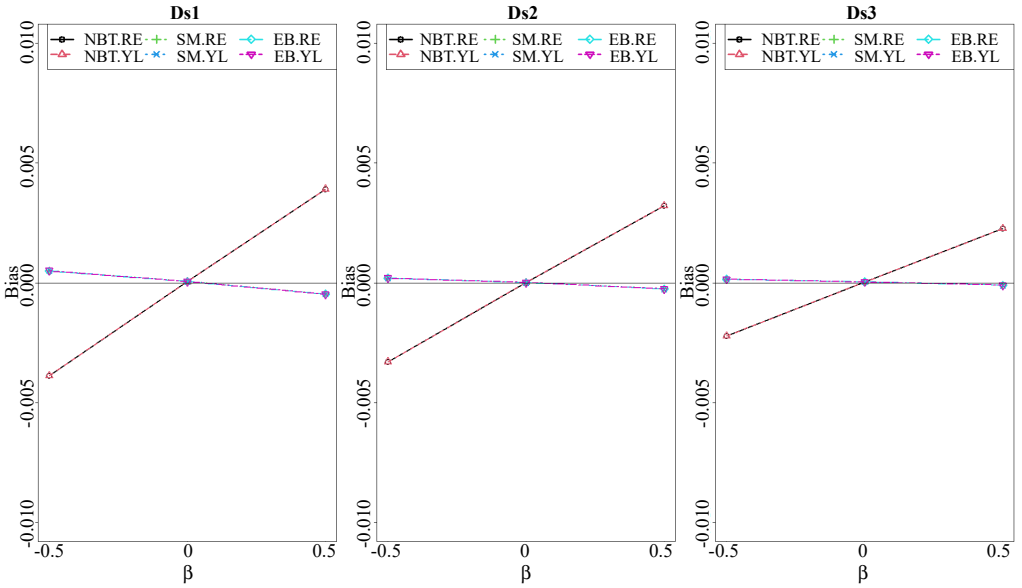

Figure 2. Simulated biases (SB) of six empirical predictors in the case $m = 50$; each sub-figure shows the results for the three sampling variance patterns $Ds1$(Left), $Ds2$(Center), and $Ds3$(Right), with each x-axis indicating $\beta \in \{-0.5, 0, 0.5\}$.

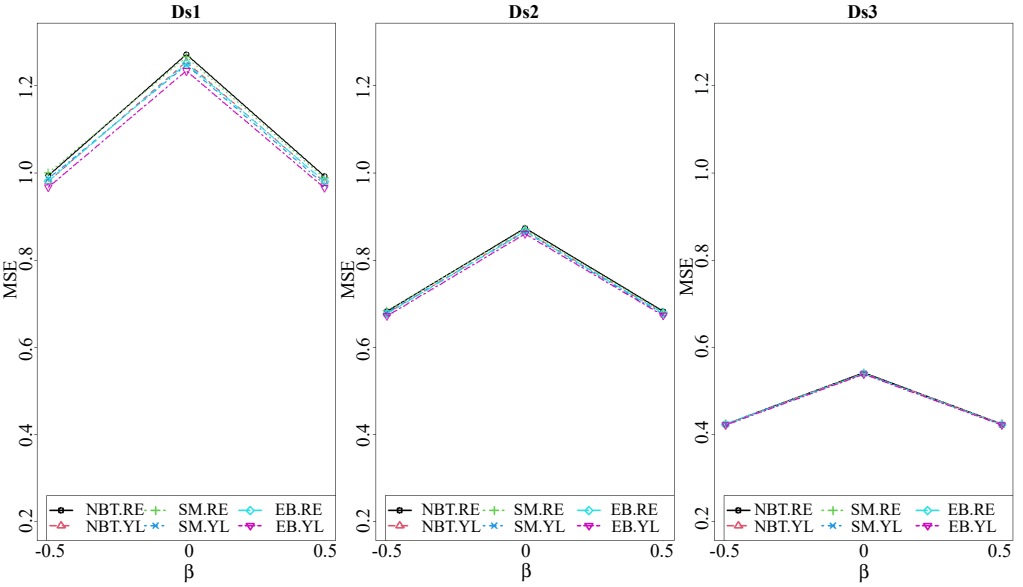

Figure 3. Simulated MSEs (SMSE) of six empirical predictors in the case $m = 15$; each sub-figure shows the results for the three sampling variance patterns $Ds1$(Left), $Ds2$(Center), and $Ds3$(Right), with each x-axis indicating $\beta \in \{-0.5, 0, 0.5\}$.

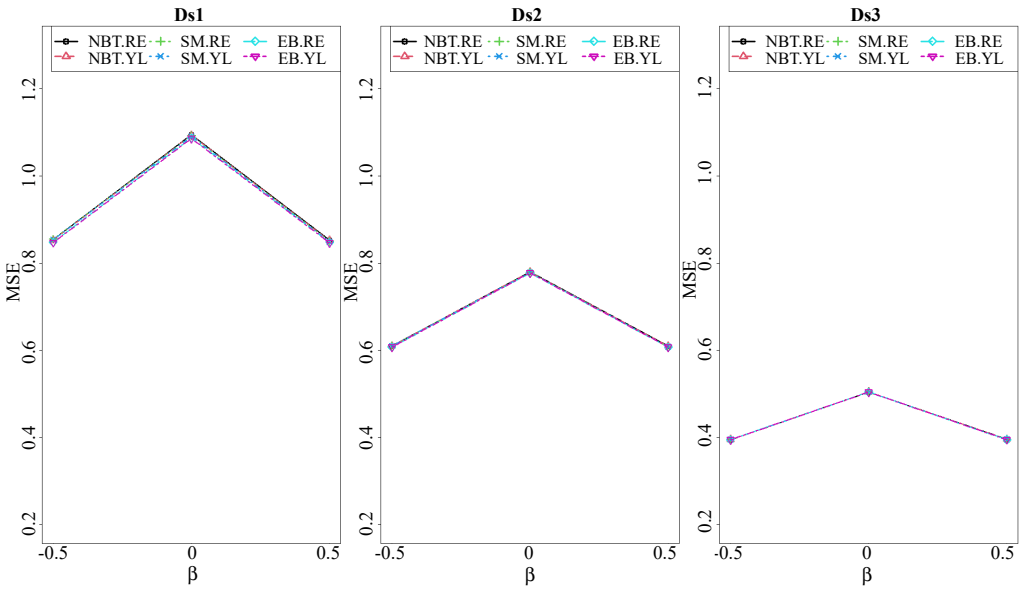

Figure 4. Simulated MSEs (SMSE) of six empirical predictors in the case $m = 50$; each sub-figure shows the results for the three sampling variance patterns $Ds1$(Left), $Ds2$(Center), and $Ds3$(Right), with each x-axis indicating $\beta \in \{-0.5, 0, 0.5\}$.

4.2. Assessment of MSE estimation

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

1) A second-order unbiased MSE estimator for untransformed data, based on the $\hat{A}_{RE}$ 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) A second-order unbiased MSE estimator for untransformed data, based on the $\hat{A}_{YL}$ 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) A first-order unbiased MSE estimator $\hat{M}_{1i}^{RE} \equiv M_{1i}(\hat{\lambda})$ using $\hat{A}_{RE}$, where $M_{1i}(\lambda)$ is given in Theorem 1 (denoted by "M1.RE");

4) A 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");

Table 1. Percentage (%) of occurrence of zero estimates when using $\hat{M}_{1i}^{RE}$ in each combination 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

5) A second-order unbiased MSE estimator $\hat{M}_i^{RE} \equiv \hat{M}_i(\hat{\lambda})$ using $\hat{A}_{RE}$, defined in (3.3) (denoted by "Ms.RE");

6) A second-order unbiased MSE estimator $\hat{M}_i^{YL} \equiv \hat{M}_i(\hat{\lambda})$ using $\hat{A}_{YL}$, defined in (3.3) (denoted by "Ms.YL");

To evaluate these MSE estimators, we calculate the percentage of relative bias (PRB) and the percentage 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 rth 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 to compare the MSE estimators (M1.RE,M1.YL,Ms.RE,Ms.YL). Two untransformed MSE estimators (NBT.RE,NBT.YL) are omitted, owing to their large values (over 300) in all situations. From these results, in terms of the relative bias and the relative RMSE, our proposed MSE estimators $\hat{M}_i(\hat{\lambda})$ outperform their competitors.

In this simulation, there were no negative MSE estimates, and only $\hat{M}_{1i}^{RE}$ produced zero estimates. Both of our second-order unbiased MSE estimates always used $\hat{M}_i^0$ in (3.3). Finally, Table 1 shows for the percentage of occurrence of zero estimates in using $\hat{M}_{1i}^{RE}$. The result are the same as the simulated probability of the REML being zero. Thus, the estimator $\hat{M}_{1i}^{RE}$ may yield unrealistic estimates of the MSE.

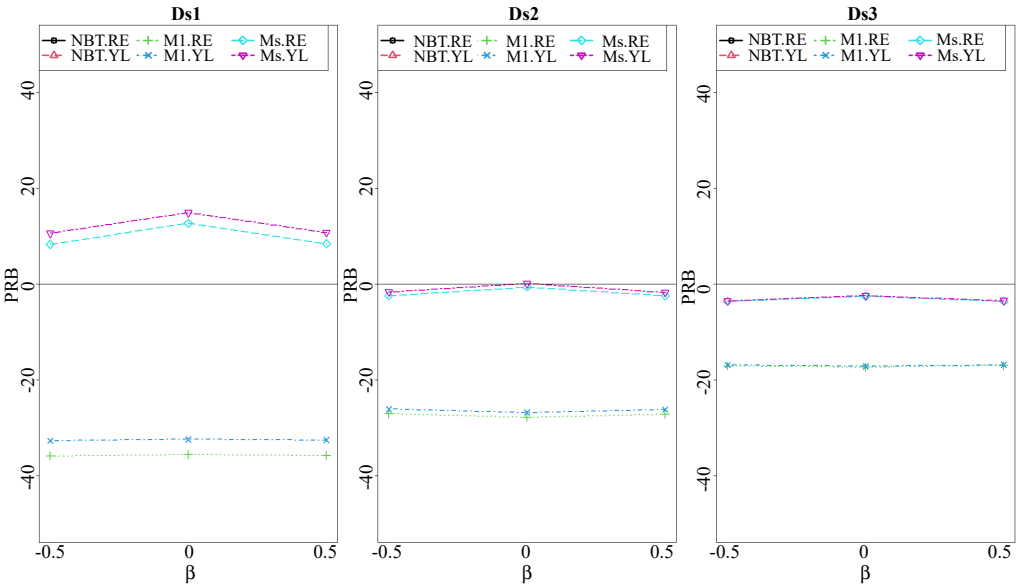

Figure 5. Percentage of relative bias (PRB) of six MSE estimators for MSE of EB estimator $\hat{p}_i^{EB.YL}$ in the case $m = 15$; each sub-figure shows the results in the three sampling variance patterns $Ds1$(Left), $Ds2$(Center), and $Ds3$(Right), with each x-axis indicating $\beta \in \{-0.5, 0,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

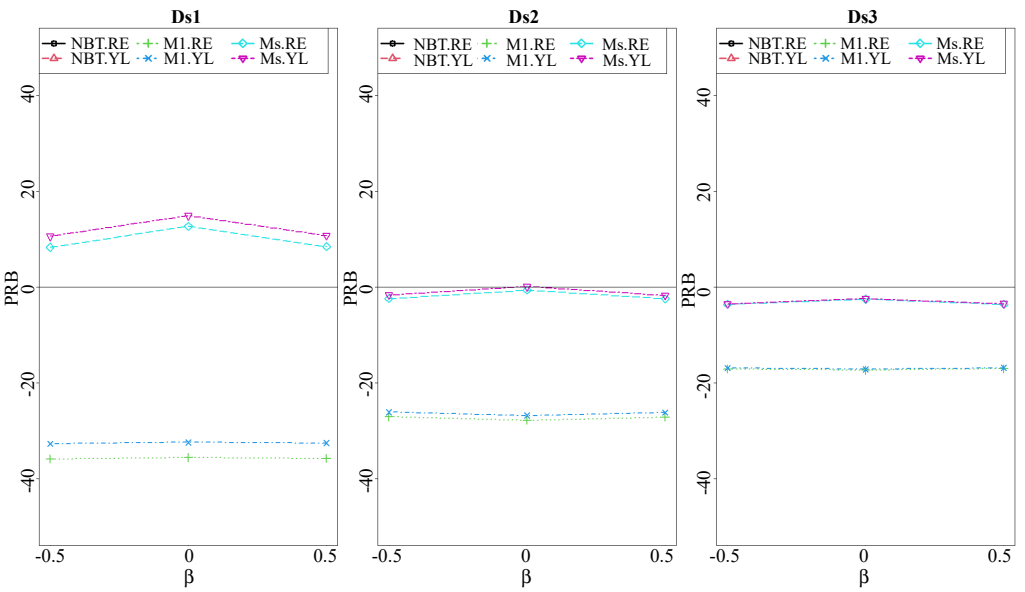

Figure 6. Percentage of relative bias (PRB) of six MSE estimators for MSE of EB estimator $\hat{p}_i^{EB.YL}$ in the case $m = 50$; each sub-figure shows the results in the three sampling variance patterns $Ds1$(Left), $Ds2$(Center), and $Ds3$(Right), with each x-axis indicating $\beta \in \{-0.5, 0,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

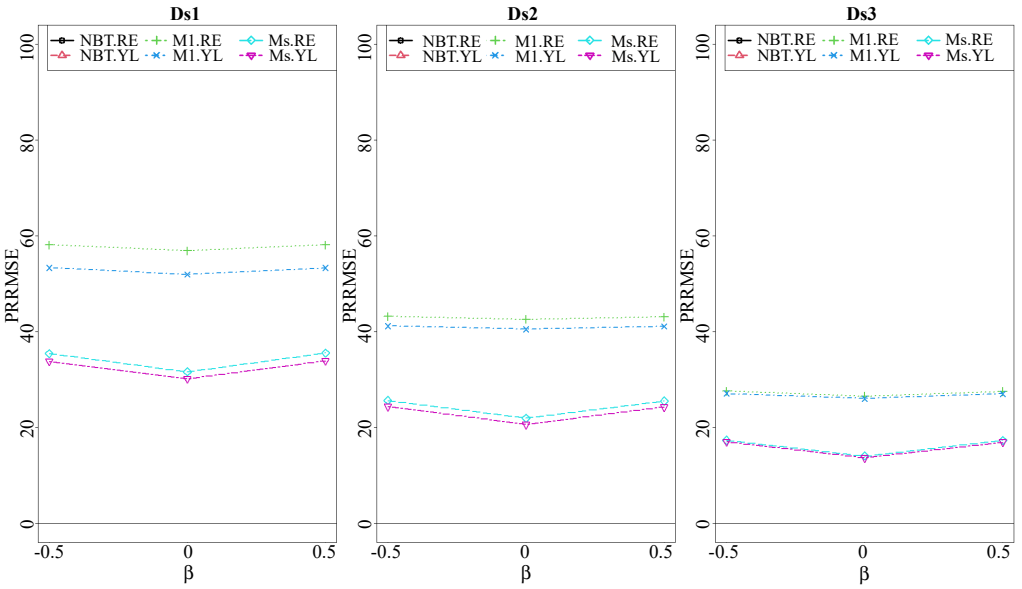

Figure 7. Percentatage of relative RMSE (PRRMSE) of six MSE estimators for MSE of EB estimator $\hat{p}_i^{EB.YL}$ in the case $m = 15$; each sub-figure shows the results in the three sampling variance patterns $Ds1$(Left), $Ds2$(Center), and $Ds3$(Right), with each x-axis indicating $\beta \in \{-0.5,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

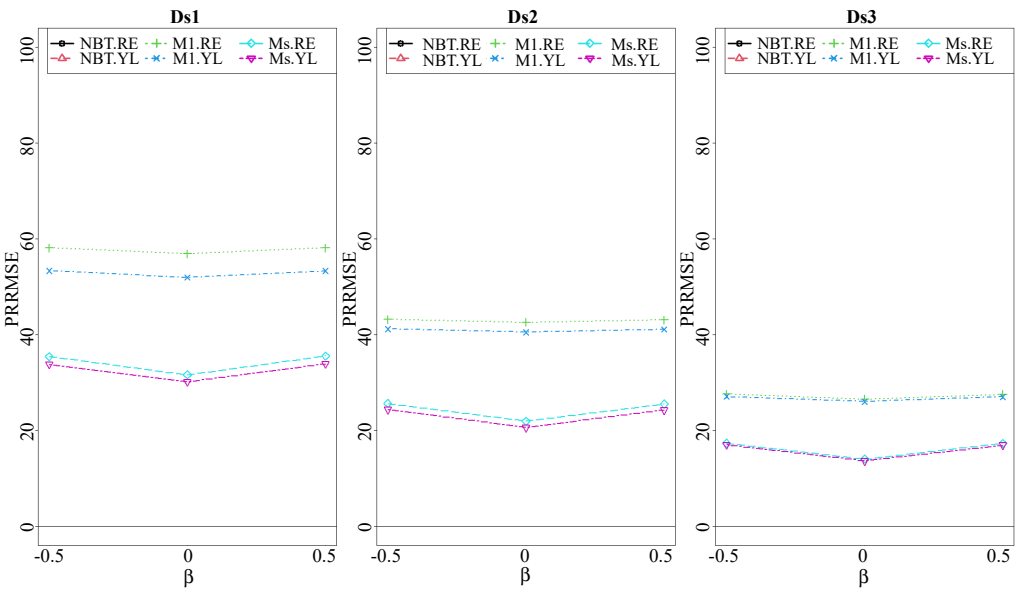

Figure 8. Percentatage of relative RMSE (PRRMSE) of six MSE estimators for MSE of EB estimator $\hat{p}_i^{EB.YL}$ in the case $m = 50$; each sub-figure shows the results in the three sampling variance patterns $Ds1$(Left), $Ds2$(Center), and $Ds3$(Right), with each x-axis indicating $\beta \in \{-0.5,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

5. Data Analysis

COVID-19 has become a global pandemic since 2020. In this study, we illustrate our methodology using real COVID-19 data. The purpose of this study is, as one example, to predict the positive rate in PCR testing for each of the 47 prefectures in Japan.

For this purpose, we use real data on the number of positive cases and the number of people who have taken the PCR test for each prefecture as at date of April 21, 2021. The data are obtained from the website of Japan's Ministry of Health, Labour and Welfare (`https://www.mhlw.go.jp/stf/covid-19/open-data.html`).

We assume model (2.4) and apply our methodology to these data. Now, let y_i and n_i be the direct positive rate of those who took the PCR test and the number of PCR tests conducted for the ith prefecture in Japan, with $i \in \{1, \ldots, 47\}$. In addition, the real auxiliary variable $x_i = (1, x_{2i})'$ is used, where $x_{1i} = 1$ represents a dummy variable for the intercept term, and x_{i2} denotes $N_i \times 10^{-6}$ with population size N_i. We let N_i be the population size, taken from the Japanese Census at 2015, which provides the latest open census data. These data are obtained from the website (`https://www.e-stat.go.jp/`). From the model assumption, the sampling variance D_i is assumed to be $1/n_i$ for the ith prefecture. We call the pattern $D1$, hereafter. Moreover, we tried a hypothetical setting of sample size $n_i^* = \lceil n_i \times 10^{-4} \rceil$ with the real data (y_i, x_{2i}), where $\lceil n \rceil$ indicates the smallest integer greater than or equal to n. Here, let D2: $D_i = 1/n_i^*$ be the hypothetical pattern of the sampling variance for all 47 prefectures in Japan. In this case, the range of n_i^* becomes 1 to 194. We believe this situation is also important for prediction in the early stages of the pandemic.

5.1. Predict positive rate in PCR testing

We first compare the six EB estimates of p_i, introduced in Section 4.1, with the direct proportion estimates (Direct). Figure 9 shows the results for 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 the 47 prefectures in Japan and each prediction, respectively. In addition, the resulting predicted values are arranged in ascending order of the size n_i.

The left sub-figure indicates no large differences between the estimates. On the other hand, there appear to be considerable differences between the direct estimates and the other EB estimates when the sample sizes are small, as seen from the right figure. This may be because of the small sample size at the early

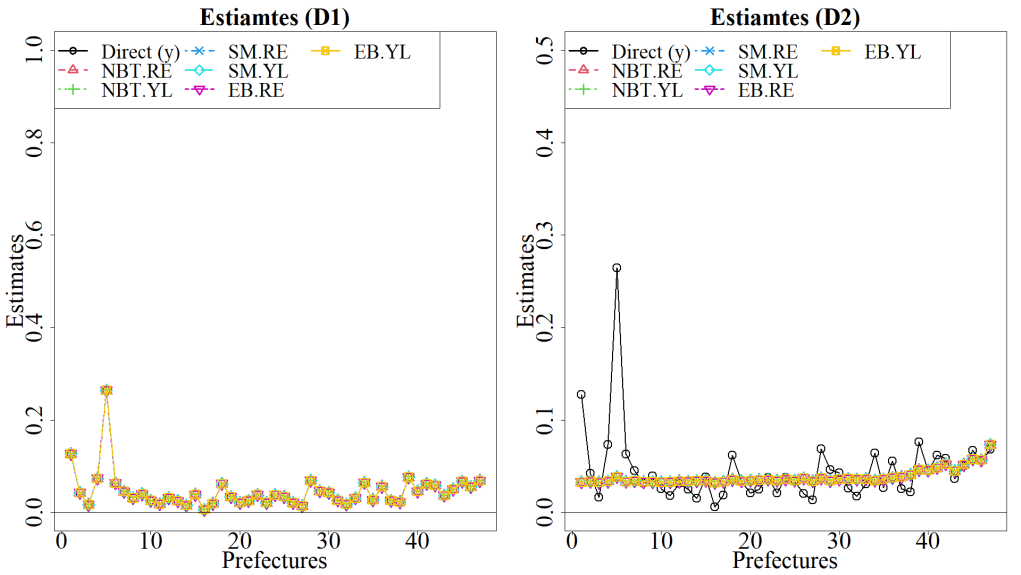

Figure 9. Seven predicted values of positive rates in PCR testing with two sampling variance patterns $D1 : D_i = 1/n_i$ (Left) and $D2 : D_i = 1/n_i^*$ (Right).

stage of the pandemic.

5.2. Estimates of coefficients of variation

Next, in order to more investigate the effect of the MSE estimates based on several EB 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 using the natural back-transformed estimator $\hat{p}_i^{NBT.RE}$ and untransformed MSE estimator $\hat{M}_i^{NBT.RE}$, (denoted by "NBT.RE") ;

2) $CV^{NBT.YL} \equiv \sqrt{\hat{M}_i^{NBT.YL}}/\hat{p}_i^{NBT.YL} \times 100$, constructed using 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 using the bias-adjusted EB 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 using the bias-adjusted EB estimator $\hat{p}_i^{EB.YL}$ and the first-order unbiased MSE estimator $\hat{M}_{1i}^{YL}$, (de-

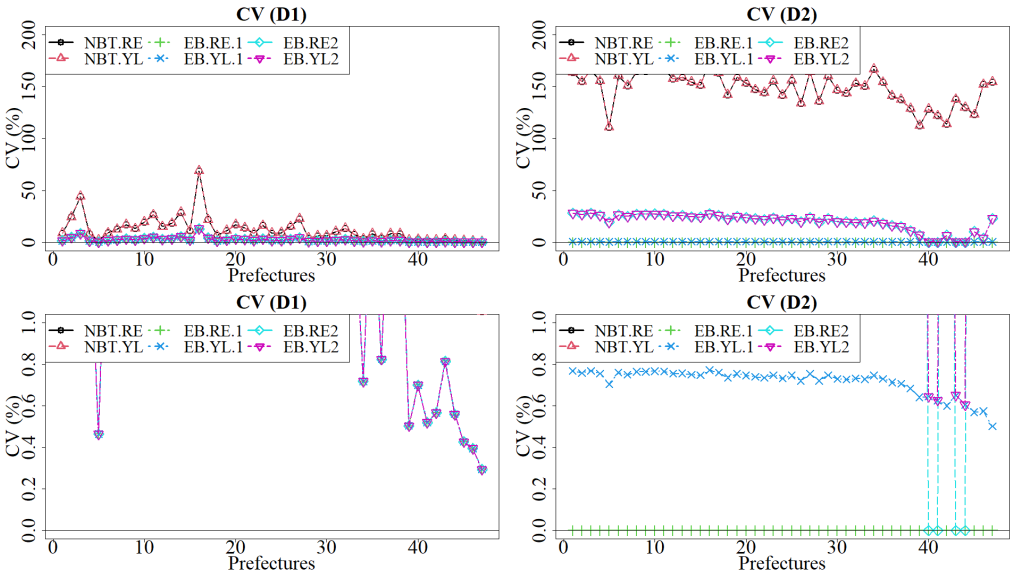

Figure 10. Six estimates of the coefficient of variation (CV) of the positive rate in PCR testing with two sampling variance patterns, $D1 : D_i = 1/n_i$ (Left two figures) and $D2 : D_i = 1/n_i^*$ (Right two figures); Each top and bottom sub-figure are the same, except the scale changes on the y-axis.

noted by "EB.YL.1");

5) $CV^{EB.RE.2} \equiv \sqrt{\hat{M}_i^{RE}}/\hat{p}_i^{EB.RE} \times 100$, constructed using the bias-adjusted EB 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 using the bias-adjusted EB 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 $CV^{EB.YL.2}$, note that $\hat{M}_{1i}^{RE}$ and $\hat{M}_{1i}^{YL}$, respectively, are used as $\hat{M}_i^*$ in (3.3).

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

The top left sub-figure in the sampling variance pattern D1 shows that the overall differences of the CV estimates 3)–6) from the two naive CV estimates 1)–2) decrease as the sample size n_i increases. The CV estimates 3)–6) do not show large differences in the sub-figure.

We also report the result for an alternative situation using the pattern D2 in the top right sub-figure in Figure 10. The figure shows that the two naive CV estimates, $CV^{NBT.RE}$ and $CV^{NBT.YL}$, are much larger than the others are. In contrast, the estimates $CV^{EB.RE.1}$ and $CV^{EB.YL.1}$ provide much smaller estimates for all prefectures. Nevertheless, these CV estimates need to be examined carefully, owing to the small sample size. There are two reasons: first, the MSE estimators in the four CV estimators 1)–4) do 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 the right sub-figure. The latter unrealistic phenomenon is caused by the REML estimate being zero. Furthermore, the bottom of the right sub-figure shows that the $CV^{EB.RE.2}$ also provides zero estimates of CV for a few prefectures, although it is constructed using the second-order unbiased estimator of the MSE. Note that it is not based on our final suggested MSE estimator in (3.3) because $M_i^* = M_{1i}^{RE}$ may not be strictly positive.

On the other hand, there are no 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, we used $\hat{M}_i^0$ as given in (3.3).

From the perspective of both the theoretical and the simulation results, we believe that $CV^{EB.YL.2}$ provides relatively precise and realistic CV estimates.

6. Conclusion

In this study, we focus on arc-sin transformations for binomial sample proportions in the context of small area estimation. This specific variance stabilizing transformation avoids the important assumption of known sample variances, as in the Fay–Herriot small area model. We find an explicit EB estimator for such transformed data, and then evaluate the asymptotic order of its bias. We also 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 that maintains strict positivity. Simulations show the superiority of our proposed method over competing 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 of the 47 prefectures in Japan. Although this is just one example of using our methodology, it may contribute to quick studies at an early stage, because the model does not require personal information in the micro-data or the estimates of the sampling variances, owing to its aggregated model and variance stabilization, respectively.

Nevertheless, we may also consider a more general model to treat complex COVID-19 data. In future, we intend to study EB estimation under more general transformation models.

Supplementary Material

All technical proofs are provided in the online Supplementary Material.

Acknowledgments

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. *Mathematical Proceedings of the Cambridge Philosophical Society* **48**, 600–607.

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 M. Pratesi), 379–403. John Wiley & Sons, Chichester.

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. *Statistica 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*. John Wiley & Sons, Hoboken.

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

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.

Masayo Y. Hirose

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

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

Malay Ghosh

Department of Statistics, University of Florida, 223 Griffin-Floyd Hall, Gainesville, Florida, USA.

E-mail: ghoshm@ufl.edu

Tamal Ghosh

Citibank, N.A., 3800 Citigroup Center Drive, Tampa, Florida, USA.

E-mail: tamal.ghosh@citi.com

(Received October 2020; accepted August 2021)