A New Adjusted Residual Likelihood Method for the Fay-Herriot Small Area Model

Masayo Yoshimori * and Partha Lahiri †

Abstract
In the context of the Fay-Herriot model, a mixed regression model routinely used to combine information from various sources in small area estimation, certain adjustments to a standard likelihood (e.g., profile, residual, etc.) have been recently proposed in order to produce strictly positive and consistent model variance estimators. These adjustments protect the resulting empirical best linear unbiased prediction (EBLUP) estimator of a small area mean from possible over-shrinking to the regression estimator. However, the existing adjusted likelihood methods can lead to high bias in the estimation of both model variance and the associated shrinkage factors and can produce a negative second-order unbiased mean square error (MSE) estimate of an EBLUP. In this paper, we propose a new adjustment factor that rectifies the above-mentioned problems associated with the existing adjusted likelihood methods. In particular, we show that our proposed adjusted residual maximum likelihood estimators of the model variance and the shrinkage factors enjoy the same higher-order asymptotic bias properties of the corresponding residual maximum likelihood estimators. We compare performances of the proposed method with the existing methods using Monte Carlo simulations.

Key Words: Empirical Bayes; Linear mixed model; Profile likelihood; Residual likelihood; Shrinkage.

1. Introduction

For the last few years, there has been an increasing demand to produce reliable estimates for small geographic areas, commonly referred to as small areas, since such estimates are routinely used for fund allocation and regional planning. The primary data, usually a survey data, are usually too sparse to produce reliable direct small area estimates that use data from the small area under consideration. To improve upon direct estimates, different small area estimation techniques that use multi-level models to combine information from relevant auxiliary data have been proposed in the literature. The readers are referred to Rao (2003) for a comprehensive review of small area estimation.

The following two-level model, commonly referred to as the Fay Herriot model (see Fay and Herriot, 1979), has been extensively used in different small area applications (see, e.g., Carter and Rolph 1974; Efron and Morris 1975; Fay and Herriot 1979, etc.) For $i = 1, \cdots, m$,

$$\text{Level 1 (sampling model): } y_i|\theta_i \overset{\text{ind}}{\sim} N(\theta_i, D_i);$$

$$\text{Level 2 (linking model): } \theta_i \overset{\text{ind}}{\sim} N(x_i'\beta, A).$$

In the above model, level 1 is used to account for the sampling distribution of the direct estimates y_i, which are weighted averages or sums of observations from small area i. Level 2 links the true small area means θ_i to a vector of p known auxiliary

*Graduate school of Engineering Science, Osaka University, Osaka, Japan
†Joint Program of survey methodology, University of Maryland, College Park, USA

variables $x_i = (x_{i1}, \cdots, x_{ip})'$, often obtained from various administrative records. The parameters β and A of the linking model are generally unknown and are estimated from the available data. The assumption of known sampling variances D_i often follows from the asymptotic variances of transformed direct estimates (Efron and Morris 1975; Carter and Rolph 1974) and/or from empirical variance modeling (Fay and Herriot 1979).

The Fay-Herriot two-level model can be viewed as the following simple linear mixed model:

$$y_i = \theta_i + e_i = x_i'\beta + v_i + e_i, \quad i = 1, \cdots, m,$$

where the v_i's and e_i's are independent with $v_i \overset{\text{iid}}{\sim} N(0, A)$ and $e_i \overset{\text{ind}}{\sim} N(0, D_i)$; see Prasad and Rao (1990). Fay and Herriot (1979) called the model a Bayesian model where level 1 and level 2 are the sampling and prior distributions respectively. We define the mean squared error (MSE) of an estimator $\hat{\theta}_i$ of θ_i as $E(\hat{\theta}_i - \theta_i)^2$, where the expectation is with respect to the joint distribution of y and θ under the Fay-Herriot model. The best prediction (BP) estimator of θ_i, which minimizes the MSE, is given by:

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

where $B_i = \frac{D_i}{D_i + A}$ $(i = 1, \cdots, m)$. The superscript 'B' in $\hat{\theta}_i^B$ is a natural notation to indicate that $\hat{\theta}_i^{\text{B}}$ is also the Bayes estimator of θ_i under the squared error loss function.

Define $y = (y_1, \cdots, y_m)'$; $X' = (x_1, \cdots, x_m)$, $V = \text{diag}(A + D_1, \cdots, A + D_m)$. If A is known, β can be estimated by the standard weighted least squares estimator:

$$\hat{\beta}(A) = (X'V^{-1}X)^{-1}X'V^{-1}y.$$

Replacing β by $\hat{\beta}(A)$ one gets the following best linear unbiased prediction (BLUP) estimator of θ_i:

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

In the more realistic case when both β and A are unknown, an empirical best linear unbiased prediction (EBLUP) estimator of θ_i is given by

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

where $\hat{B}_i = D_i/(\hat{A} + D_i), i = 1, \cdots, m$, and $\hat{\beta} = \hat{\beta}(\hat{A})$, $\hat{A}$ is a consistent estimator of A, for large m. The superscript 'EB' in $\hat{\theta}_i^B$ is a natural notation to indicate that $\hat{\theta}_i^{EB}$ is also an empirical Bayes estimator of θ_i under the squared error loss function.

Rao (2003) and Jiang and Lahiri (2006) list several consistent estimators of A. They include different method-of-moments estimators (see Fay and Herriot, 1979; Prasad and Rao, 1990) and likelihood based methods such as residual maximum likelihood (REML) and profile maximum likelihood (see, e.g., Datta et al., 1999; Das et al., 2004). In terms of the MSE up to the order $O(m^{-1})$, the REML and ML estimators of A are equivalent and are better than those of the method-of-moments estimators (Datta et al., 2005). It is also known that REML is superior to the ML method in terms of the higher-order asymptotic bias; for example, the bias of REML is $o(m^{-1})$ while that of ML is $O(m^{-1})$, under certain regularity conditions, for large m.

Under certain regularity conditions, $\text{MSE}[\hat{\theta}_i^{\text{EB}}] = g_{1i}(A) + g_{2i}(A) + g_{3i}(A) + o(m^{-1})$, where $g_{1i}(A) = \frac{AD_i}{A + D_i}$, $g_{2i}(A) = \frac{D_i^2}{(A + D_i)^2}\text{Var}(\hat{\beta})$, $g_{3i}(A) = \frac{D_i^2}{(A + D_i)^3}\text{Var}(\hat{A})$, $\text{Var}(\hat{\beta}) = x_i'\left(\sum_{j=1}^{m} \frac{1}{A + D_j} x_j x_j'\right)^{-1} x_i$, and $\text{Var}(\hat{A})$ is the variance of $\hat{A}$ correct up to the order $O(m^{-1})$. The term $g_{1i}(A)$ is the dominating term [of order $O(1)$], capturing the uncertainty of the BP. The additional terms $g_{2i}(A)$ and $g_{3i}(A)$, which are of order $O(m^{-1})$, capture the uncertainty due to estimation of β and A, respectively. It is interesting to note that estimation of A affects the term $g_{3i}(A)$ term through $\text{Var}(\hat{A})$ - the more the variability in the estimator $\hat{A}$ the more the MSE of EBLUP. For example, for both REML and ML, $\text{Var}(\hat{A}) = 2\left\{\sum_{j=1}^{m} \frac{1}{(A+D_j)^2}\right\}^{-1}$, smaller than the asymptotic variance of the Prasad-Rao (PR) and Fay-Herriot (FH) method-of-moments estimators of A; see Datta et al. (2005). It is also interesting to note that the adjustment factor $h(A)$ does not affect the MSE of the EBLUP, under a suitably chosen class of adjustment factors such as the ones corresponding to the adjusted profile likelihood (AM.LL) and the adjusted residual likelihood (AR.LL) estimators given in Li and Lahiri (2011).

Note that the second-order approximation involves unknown A and thus cannot be used to assess the uncertainty of EB for a given data set. A MSE estimator, denoted as $\widehat{\text{MSE}}(\hat{\theta}_i^{\text{EB}})$, is called a second-order unbiased (or nearly unbiased) estimator of $\text{MSE}(\hat{\theta}_i^{\text{EB}})$ if $\text{E}[\widehat{\text{MSE}}(\hat{\theta}_i^{\text{EB}})] = \text{MSE}(\hat{\theta}_i^{\text{EB}}) + o(m^{-1})$. The second-order approximation given above is useful in obtaining a second-order unbiased MSE estimator of EB:

$$\widehat{\text{MSE}}(\hat{\theta}_i^{\text{EB}}) = g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + 2g_{3i}(\hat{A}) - \hat{B}_i^2\widehat{\text{Bias}}(\hat{A}),$$

where $\widehat{\text{Bias}}(\hat{A})$ is a second-order unbiased estimator of $\text{Bias}(\hat{A})$. For REML and ML, $\widehat{\text{Bias}}(\hat{A})$ are always non-negative.

One problem with the standard method-of-moments and likelihood-based methods is that estimate of A could be zero, especially when A is small relative to the sampling variances and m is small. This causes an overshrinkage problem in the sense that the EBLUPs reduce to the regression estimates $x_i'\hat{\beta}$ for all small areas and the estimates do not use the direct estimates, even for areas with moderately large sample. Wang and Fuller (2003) [17] suggested a strictly positive method-of-moments estimator for the model variance. On the other hand, Li and Lahiri (2011) [13] introduced an adjusted REML and PML methods, which are strictly positive and consistent. Their methods are equivalent to the standard likelihood methods in terms of higher-order MSE. The order of bias for the Li-Lahiri adjusted maximum residual and profile likelihood methods is $O(m^{-1})$, same as the order of bias for the PML but higher than the REML method. In a Monte Carlo simulation study, Lahiri and Pramanik (2011) observed that the Li-Lahiri adjusted maximum likelihood method can be subject to high bias, especially when m is small and B is close to 1.

In section 2, we propose a new adjustment factor for the adjusted residual and profile maximum likelihood methods. The purpose of introducing this new adjustment factor is to reduce the bias of the adjusted likelihood based methods. In this section, we analytically showed that the new adjusted REML is equivalent to the REML in terms of higher order asymptotic bias and MSE for estimation of both A and B_i. In terms of the estimation of the shrinkage factor, we argue analytically that our new adjusted REML generally overestimates while the Li-Lahiri adjusted REML underestimates. Also, our proposed adjusted REML is better than the Li-Lahiri adjusted REML when B_i is greater than 1/2, but worse than the Li-Lahiri adjusted REML when B_i is less than 1/2. When $B_i = 1/2$, the absolute values of the leading term of the bias of the two adjusted REML methods are exactly the same. Thus, for a small area problem with highly unstable direct estimates, we expect our method to perform well. Another advantage of the new adjusted

REML method is that the second-order unbiased MSE estimator of the EBLUP is always non-negative - this is not necessarily true for the Li-Lahiri adjusted REML. In section 3, we examine the small sample performances of different estimators of A, B, θ_i for the balanced case using a Monte Carlo simulation experiment. In this section, we also study the performances of the second-order MSE estimators of different EBLUPs that use different estimators of A. To save space, we report only the balance case and exclude proofs of all technical results.

2. A New Adjustment for the Adjusted Likelihood Method

Following Lahiri and Li (2009), we define an adjusted likelihood as

$$L_{ad}(A) \propto h(A) \times L(A),$$

where $h(A)$ is an adjustment factor and $L(A)$ is a standard likelihood function (e.g., profile likelihood, residual likelihood, etc). Adjusted maximum likelihood estimator of A is obtained by maximizing $L_{ad}(A)$ with respect to A.

We assume that $\log(h(A))$ is four times continuously differentiable with respect with A ($A > 0$) and is free of y. Note that

$$\begin{aligned} l_{ad}(A) &= l(A) + \log(h(A)), \\ l_{ad}^{(1)}(A) &= l^{(1)} + \tilde{l}_{ad}^{(1)}, \\ l_{ad}^{(2)}(A) &= l^{(2)} + \tilde{l}_{ad}^{(2)}, \end{aligned}$$

where $l(A) = \log(L(A))$, $l_{ad}(A) = \log(L_{ad}(A))$, $\tilde{l}_{ad}^{(1)}$, $l^{(j)}$ and $l_{ad}^{(j)}$ denotes the jth derivative of $\log h(A)$, $l(A) = \log L(A)$, and $l_{ad}(A)$ with respect of A ($j \geq 1$), respectively.
The following theorem provides the higher-order bias and MSE properties of the adjusted profile and residual maximum likelihood estimators for a general adjustment factor $h(A)$.

Theorem 1. *Under certain regularity conditions,*

(i) $E[\hat{A}_{g.AR} - A] = \frac{2\tilde{l}_{ad}^{(1)}}{tr[V^{-2}]} + o(m^{-1})$,

(ii) $E[\hat{A}_{g.AM} - A] = \frac{tr[P-V^{-1}]+2\tilde{l}_{ad}^{(1)}}{tr[V^{-2}]} + o(m^{-1})$,

(iii) $E[(\hat{A}_{g.Ad} - A)^2] = \frac{2}{tr[V^{-2}]} + o(m^{-1})$,
where $\hat{A}_{g.Ad} \in \{\hat{A}_{g.AM}, \hat{A}_{g.AR}\}$ *denotes an adjusted likelihood estimator.*

We choose $h(A) = \left(\tan^{-1}\{\text{tr}[I - B]\}\right)^{1/m}$, where $B = \text{diag}(B_1, \cdots, B_m)$. This choice satisfies the conditions of Theorem 1. Let $\hat{A}_{AM.YL}$ and $\hat{A}_{AR.YL}$ denote the adjusted profile and residual likelihood estimator of A, respectively, with this choice of $h(A)$. These adjusted maximum likelihood estimators of A are consistent and strictly positive, under certain regularity conditions. We have the following theorem that provides the higher-order bias and MSE properties of the proposed adjusted profile and residual maximum likelihood estimators.

Theorem 2. *Under certain regularity conditions,*

(i) $E[\hat{A}_{AR.YL}] - A = o(m^{-1})$;

(ii) $E[\hat{A}_{AM.YL}] - A = \frac{tr[P-V^{-1}]}{tr[V^{-2}]} + o(m^{-1})$;

(iii) $E[(\hat{A}_{A.YL} - A)^2] = \frac{2}{tr[V^{-2}]} + o(m^{-1})$,
where $A_{A.YL} \in \{A_{AR.YL},\ A_{AM.YL}\}$.

Note that the variances of $B_i(\hat{A})$ for all likelihood methods considered are equivalent, up to order $O(m^{-1})$. Thus we can compare different estimators using the bias to the variance (BV) ratio. Using Li and Lahiri (2009), we have

$$\frac{Bias(B(\hat{A}))}{Var(B(\hat{A}))} = \frac{A + D_i}{D_i}(1 - \tilde{l}_{ad}^{(1)}(A + D_i)) + o(1) \quad (1)$$

For our proposed choice of the adjustment factor, we obtain the following results:

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

(i) $BV_{RE} = \frac{1}{B_i} + o(1),\ BV_{AR.YL} = \frac{1}{B_i} + o(1),\ BV_{AR.LL} = -\frac{1}{1-B_i} + o(1)$,

(ii) $BV_{ML} = \frac{1}{B_i}[1 + (A + D_i)\frac{H}{2}] + o(1),\ BV_{AM.YL} = \frac{1}{B_i}[1 + (A + D_i)\frac{H}{2}] + o(1)$, $BV_{AM.LL} = -\frac{1}{1-B_i} + \frac{(A+D_i)^2}{D_i}\frac{H}{2} + o(1)$,

where $H = tr[V^{-1}X(X'V^{-1}X)^{-}X'V^{-1}] > 0$.

3. A Monte Carlo Simulation Study

In this section, we design a Monte Carlo simulation study to compare small sample performances of different estimators of the model variance A, the shrinkage parameter B, different EBLUP estimators corresponding to different estimators of A and the associated MSE estimators for the balanced case of the Fay-Herriot model with common mean $x_i'\beta = 0$ and equal sampling variance $D_i = D$ $(i = 1, \cdots, m)$. To be explicit, we generate $R = 10^4$ independent replicates $\{Y_i,\ v_i,\ i = 1, \ldots, 15\}$ using the following Fay Herriot model;

$$Y_i = v_i + e_i,$$

where v_i and e_i are mutually independent with $v_i \overset{iid}{\sim} N(0, A)$, $e_i \overset{ind}{\sim} N(0, D)$, $i = 1, \cdots, m$.

In order to examine effects of m and A/D on the performances of different estimators, we consider m =15, 45 and $D \in \{0.05, 0.1, 1, 10, 20\}$, for fixed $A = 1$. We have also investigated performances of different estimators when we vary $A \in (0.05, 0.1, 1, 10, 20)$, keeping $D = 1$ fixed. But our findings about the relative performances of different estimators are similar and so to save space we do not report the results for this case.

3.1 Comparison of Different Estimators of A, B and θ_i

We compare following estimators of A: residual likelihood (RE), maximum likelihood (ML), Li-Lahiri adjusted residual maximum likelihood (AR.LL), Li-Lahiri adjusted maximum likelihood (AM.LL), Wang-Fuller method-of-moments (WF), proposed adjusted residual maximum likelihood (AR.YL), and proposed adjusted maximum likelihood (AM.YL). Note that in the balanced case, the residual maximum likelihood and the Prasad-Rao ([15]) estimators of A are identical. We also compare the corresponding estimators of $B = D/(A + D)$ and EBLUP of θ_i.

Table 1 displays simulated probabilities of obtaining zero estimates of A by different methods. Note that only the RE and ML could yield zero estimates and the probability of getting zero estimates is high when A/D is small for both methods, although RE is relatively less prone to zero estimate than the ML. The performances of both RE and ML improve as m increases.

We define the percent relative bias (RB) of a given estimator, say $\hat{A}$, of A by $\frac{1}{R}\sum_{r=1}^{R}(\hat{A}^{(r)} - A)/A \times 100$, where $\hat{A}^{(r)}$ denotes an estimator of A for the rth replication, $r = 1, \ldots, R$. The relative bias of the corresponding estimator of B is defined in a similar way. Table 2 displays the percent relative bias (RB) of different estimators of A. When $A/D = 1$, REML is the best. For higher values of A/D, performances of RE, WF and AR.YL are almost identical and are generally better AR.LL and AM.LL. For small values of A/D, it is interesting to note that ML performs the best, followed by AM.YL and they are both better than RE. In this case, both AM.YL and AR.YL perform better than WF and substantially better than AM.LL and AR.LL. Overall, in terms of relative bias, AR.YL is tracking RE well supporting our asymptotic theory. As m increases, the performances of all the estimators improve.

We define the simulated MSE of an estimator $\hat{B}$ of B as: $MSE(\hat{B}) = \frac{1}{R}\sum_{r=1}^{R}(\hat{B}^{(r)} - B)^2$, where $\hat{B}^{(r)}$ denotes an estimator of B for the rth replication, $r = 1, \ldots, R$. The MSE of an EBLUP is defined in a similar way. We express the simulated MSEs in percentages.

We report the relative biases and MSEs of different estimators of B in Table 3. For small A/D, AR.LL and AM.LL are subject to severe underestimation and large MSE, even when $m = 45$. The proposed adjustment factor cuts down this underestimation and MSE substantially, but suffers from an overestimation problem for large A/D, which diminishes when $m = 45$. It appears that the new adjustment factor works better than the one proposed by Li and Lahiri (2011) when A/D is small. However, the opposite is true when A/D is large. The performances of AR.YL and RE are almost identical when $m = 45$.

Table 4 displays the simulated MSE of EBLUPs using different estimators of A. For large A/D, all the methods provides similar results. The new adjustment factor performs better than the Li-Lahiri adjustment factor for small A/D, even when $m = 45$. The new adjustment factor provides results similar to the standard likelihood method. Overall, ML and AM.YL seem to perform the best.

3.2 Comparison of Different MSE Estimators

For a given area i, the relative bias and relative root mean square error (RRMSE) of a MSE estimator, say $\widehat{MSE}_i$, are defined as

$$RB_{i,\widehat{MSE}_i} = \frac{\sum_{r=1}^{R}\widehat{MSE}_i^{(r)}/R - MSE_i}{MSE_i} \times 100,$$

$$RRMSE_{i,\widehat{MSE}_i} = \frac{\sqrt{MSE(\widehat{MSE}_i)}}{MSE_i} \times 100,$$

where $\widehat{MSE}_i^{(r)}$ denotes an estimator for MSE_i of $\hat{\theta}_i^{EB}$ in the rth replication ($r = 1, \cdots, R$).. In table 5, we report the average relative bias and average relative root mean squared error (RRMSE) of MSE estimators of EBLUPs, where the average is taken over the small areas. For this particular simulation, we increase the number of replication to $R = 10^5$ to reduce the Monte Carlo errors. Our new adjusted likelihood methods perform similar to the standard likelihood methods. However, for small A/D, the Li-Lahiri adjustment works better than the new adjustment.

4. Acknowledgment

The first author's research was supported by JSPS KAKENHI Grant Number A2427420. She completed this paper while visiting the University of Maryland, College Park, USA as a visiting Ph.D. student. The second author's research was supported in part by the National Science Foundation SES-085100

References

[1] W. Bell, Accounting for uncertainty about variances in small area estimation. Bulletin of the International statistical Institute, 52nd session, Helsinki, 1999, available at http//www.census.gov/hhes/www/saipe under "Publications.".

[2] A.Chatterjee, P. Lahiri, H. Li, Parametric bootstrap approximation to the distribution of EBLUP, and related prediction intervals in linear mixed models, The Annals of Statistics 36 (2008) 1221-1245.

[3] S. Chen and P. Lahiri, On Mean Squared Prediction Error Estimation in Small Area Estimation Problems, Communications in Statistics–Theory and Methods 37 (2008) 1792-1798

[4] G.M. Carter and J.F. Rolph, Empirical Bayes methods applied to estimating fire alarm probabilities, Journal of the American Statistical Association 69 (1974) 880-885.

[5] D.R. Cox, Prediction intervals and empirical Bayes confidence intervals, in:J.Gani(Ed.), Perspectives in Probability and Statistics, Papers in Honor of M.S. Bartlett, Academic Press, 1975, pp.47-55.

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

[7] G.S. Datta, P.Lahiri, A unified measure of uncertainty of estimated best linear unbiased predictions in small area estimation problems, Statistica Sinica 10 (2002) 613-627

[8] G.S. Datta, J.N.K. Rao, D.D. Smith, On measuring the variability of small area estimators under a basic area level model, Biometrika 92 (2005) 183-196.

[9] B. Efron and C. Morris, Data analysis using Stein's estimator and its generalizations, J. Amer. Stat. Assoc 70 (1975) 311-319.

[10] R.E. Fay, R.A. Herriot, Estimates of income for small places: an application of James-Stein procedures to census data, Journal of the American Statistical Association 74 (1979) 269-277

[11] J. Jiang and P. Lahiri, Mixed model prediction and small area estimation, Editor's invited discussion paper, Test, 15 (2006) 1-96.

[12] P. Lahiri, S. Pramanik, Discussion of "Estimating Random Effects via Adjustment for Density Maximization" by C. Morris and R. Tang, Statistical Science 26 (2011) 291-295

[13] H. Li, P. Lahiri, An adjusted maximum likelihood method for solving small area estimation problems, Journal of Multivariate Analysis, 101 (2011) 882-892

[14] P.Lahiri and H. Li, Generalized Maximum Likelihood Method in Linear Mixed Models with an Application in Small Area Estimation, In Proceedings of the Federal Committee on Statistical Methodology Research Conference, (2009), available at http//www.fcsm.gov/events/papers2009.html

[15] N.G.N. Prasad, J.N.K. Rao, The estimation of the mean squared error of small area estimators, Journal of the American Statistical Association 85 (1990) 163-171

[16] J.N.K. Rao, Small Area Estimation, Wiley, New York, 2003.

[17] J. Wang and W. A. Fuller, The Mean Squared Error of Small Area Predictors Constructed with Estimated Area Variances, Journal of the American Statistical Association 98 (2003) 716-723

Table 1: Percentage of zero estimates

m	A/D	RE	ML	AR.LL	AM.LL	AR.YL	AM.YL	WF
15	0.05	49.65	56.36	0	0	0	0	0
	0.1	45.13	52.42	0	0	0	0	0
	1	6.48	8.52	0	0	0	0	0
	10	0	0	0	0	0	0	0
	20	0	0	0	0	0	0	0
45	0.05	44.22	48.44	0	0	0	0	0
	0.1	35.79	39.61	0	0	0	0	0
	1	0.22	0.38	0	0	0	0	0
	10	0	0	0	0	0	0	0
	20	0	0	0	0	0	0	0

Table 2: RB of $\hat{A} \times 100$

m	A/D	RE	ML	AR.LL	AM.LL	AR.YL	AM.YL	WF
15	0.05	269.29	181.93	1290.13	1110.01	322.91	235.88	416.71
	0.1	112.07	63.71	632.42	537.77	138.14	90.37	182.55
	1	1.62	-11.32	63.88	45.33	2.37	-10.47	3.02
	10	-0.48	-7.78	19.61	9.65	-0.47	-7.77	-0.48
	20	-0.22	-7.2	18.12	8.65	-0.22	-7.2	-0.22
45	0.05	126.17	97.27	553.31	514.16	137.58	108.99	223.78
	0.1	48.03	30.88	262.46	241.06	53.02	36.11	88.85
	1	0.03	-4.41	18.98	14.1	0.04	-4.39	0.08
	10	-0.09	-2.53	5.67	2.98	-0.09	-2.53	-0.09
	20	0.09	-2.24	5.35	2.78	0.09	-2.24	0.09

Table 3: RB and MSE for $\hat{B}_i$

Relative Bias × 100 of $B_i(\hat{A})$								
m	A/D	RE	ML	AR.LL	AM.LL	AR.YL	AM.YL	WF
15	0.05	-7.5	-4.9	-35.8	-32.6	-10.1	-7.5	-14.2
	0.1	-4.8	-1.8	-34.1	-30.6	-7.4	-4.5	-11.5
	1	12.7	19.7	-17.3	-11.4	11.6	18.3	10.3
	10	17.2	25.5	-1.6	6.5	17.2	25.5	17.2
	20	16.9	25.2	-0.8	7.4	16.9	25.2	16.9
45	0.05	-4.1	-3	-19.8	-18.7	-4.7	-3.6	-8.7
	0.1	-2.2	-0.9	-18.1	-16.8	-2.7	-1.4	-6.2
	1	4.9	7.2	-5.1	-2.9	4.8	7.2	4.8
	10	4.8	7.2	-0.5	1.9	4.8	7.2	4.8
	20	4.8	7.1	-0.2	2.1	4.8	7.1	4.8
MSE × 100 of $B_i(\hat{A})$								
m	A/D	RE	ML	AR.LL	AM.LL	AR.YL	AM.YL	WF
15	0.05	2.97	2.2	12.79	10.7	2.96	2.17	3.21
	0.1	2.9	2.27	10.78	8.87	2.7	2.04	2.63
	1	4.5	5.21	2.2	1.88	4.04	4.66	3.52
	10	0.25	0.31	0.15	0.17	0.25	0.31	0.25
	20	0.07	0.09	0.04	0.05	0.07	0.09	0.07
45	0.05	1.32	1.14	4.24	3.82	1.29	1.1	1.31
	0.1	1.45	1.3	3.45	3.06	1.38	1.23	1.13
	1	1.45	1.58	0.96	0.94	1.44	1.57	1.41
	10	0.05	0.05	0.04	0.04	0.05	0.05	0.05
	20	0.01	0.02	0.01	0.01	0.01	0.02	0.01

Table 4: Simulated value of MSE of EBLUP

m	A/D	RE	ML	AR.LL	AM.LL	AR.YL	AM.YL	WF
15	0.05	3.17	2.95	5.34	4.87	3.18	2.95	3.22
	0.1	1.97	1.87	2.9	2.67	1.96	1.85	1.95
	1	0.59	0.6	0.59	0.58	0.59	0.59	0.59
	10	0.09	0.09	0.09	0.09	0.09	0.09	0.09
	20	0.05	0.05	0.05	0.05	0.05	0.05	0.05
45	0.05	1.74	1.69	2.41	2.31	1.73	1.68	1.74
	0.1	1.3	1.28	1.55	1.5	1.29	1.27	1.27
	1	0.53	0.54	0.53	0.53	0.53	0.54	0.53
	10	0.09	0.09	0.09	0.09	0.09	0.09	0.09
	20	0.05	0.05	0.05	0.05	0.05	0.05	0.05

Table 5: Comparison of different estimators for MSE

Average RB								
m	A/D	RE	ML	AR.LL	AM.LL	AR.YL	AM.YL	WF
15	0.05	162.86	199.61	85.42	101.72	172.79	210.85	186.02
	0.1	114.76	139.18	70.51	83.20	124.11	149.88	137.25
	1	5.09	10.55	-7.66	-7.42	6.47	11.92	7.92
	10	-0.07	-0.12	0.10	0.08	-0.07	-0.12	-0.07
	20	0.03	0.12	-0.11	-0.04	0.03	0.12	0.03
45	0.05	156.88	178.37	90.57	96.11	161.85	184.39	205.41
	0.1	61.03	73.01	2.71	3.10	65.08	77.20	90.35
	1	-0.12	-0.60	3.81	3.66	-0.09	-0.56	0.02
	10	0.04	0.05	0.03	0.04	0.04	0.05	0.04
	20	-0.01	-0.04	0.05	0.02	-0.01	-0.04	-0.01
Average RRMSE								
m	A/D	RE	ML	AR.LL	AM.LL	AR.YL	AM.YL	WF
15	0.05	66.3	64.0	26.9	28.6	60.3	58.6	48.7
	0.1	56.0	53.7	25.2	26.6	51.4	49.6	42.7
	1	23.0	23.1	13.9	14.6	22.0	22.1	20.8
	10	3.4	3.6	2.8	3.0	3.4	3.6	3.4
	20	1.7	1.9	1.5	1.6	1.7	1.9	1.7
45	0.05	129.9	126.8	72.4	73.3	126.4	123.0	93.6
	0.1	81.3	79.3	49.6	50.1	79.1	77.1	62.8
	1	19.3	19.7	15.6	15.9	19.3	19.6	19.1
	10	2.1	2.1	2.0	2.0	2.1	2.1	2.1
	20	1.0	1.1	1.0	1.0	1.0	1.1	1.0