ABSTRACT

Best linear unbiased estimation, or equivalently maximum likelihood under normality, is the method most frequently used in animal breeding for estimation of fixed effects. James and Stein found that maximum likelihood is inadmissible under the mean squared error criterion and proposed a nonlinear, biased estimator that has smaller mean squared error than maximum likelihood throughout the parameter space in normal linear models with more than two uniquely estimable fixed effects. In this paper, several biased estimators of fixed effects in the mixed linear model are considered. Dispersion parameters are assumed to be known. An estimator that minimizes mean squared error in a certain class is derived. Because this estimator regresses the maximum likelihood estimates toward zero, an alternative estimator that "shrinks" these estimates to their mean value is also considered. The two estimators require knowledge of the true values of the fixed effects, so approximations to these are presented, including an extension of the James and Stein estimator to the mixed model. These estimators were compared with maximum likelihood in a simulation study involving a balanced group (fixed) plus sire (random) model. The James and Stein and "minimum mean squared error" estimators gave estimates of group effects with slightly smaller mean squared error than maximum likelihood. Improvement was minimal when the true group effects were far from zero. However, shrinkage of group effects toward their mean value substantially reduced mean squared error of group estimates. Genetic trend (the regression of group effects on time) was considerably underestimated using this type of shrinkage. Care must be exercised when these "improved" statistics are used for estimating some functions of fixed effects.

(Key words: mixed models, fixed effects, biased estimation)

INTRODUCTION

Best linear unbiased estimation, which is also maximum likelihood (ML) under normality, is the method most frequently used in animal breeding for estimating fixed effects for several reasons. Maximum likelihood has well-known asymptotic properties, and BLUE gives minimum mean squared error (MSE) of estimation in the class of linear, unbiased estimators (9), provided that dispersion parameters are known. Although unbiasedness is desirable, restricting attention to this class may preclude consideration of other useful estimators. A more global criterion is MSE, the expected squared difference between the estimate and the parameter being estimated; MSE is the sum of the variance of the estimator and its squared bias. Certain biased estimators have been shown to have smaller MSE than BLUE (2, 3, 9, 14). Henderson (9, page 100), however, stated

... we never know what are truly minimum mean squared error estimators and predictors since we do not know some of the parameters required for deriving them.
James and Stein (10) developed a nonlinear estimator with smaller MSE than ML throughout the parameter space when more than two uniquely estimable fixed effects are estimated in a normal linear model. Hence, ML is inadmissible under squared error loss. Inadmissibility means that at least one other estimator exists that gives estimates with MSE smaller than or equal to ML throughout the parameter space, the inequality holding at some point. Their estimator has the form of a scalar times the vector of ML estimates. This and other estimators, which "regress" the ML estimates toward some value, are generally referred to as Stein-type or shrinkage-type statistics. Estimators of this type have received attention in the statistical literature for more than two decades. For example, Efron and Morris (3, 4) and Casella (1) presented simple examples of the "Stein effect" and gave estimators of the mean vector of a multivariate normal distribution that have uniformly smaller MSE than the ML estimator. However, biased estimation of fixed effects has not been studied in animal breeding, except for the work of Goffinet (7), Henderson (9), and Gianola (6). Some theoretical results for the mixed linear model are presented in Gotway and Cressie (8).

The objective of this paper is to adapt the James-Stein (JS) estimator and related estimators to the general mixed linear model. Properties of the estimates such as bias, variance, and MSE were evaluated in a simulation study. Estimation of individual fixed effects (genetic groups) and of a function of fixed effects (genetic trend) was considered in a balanced group plus sire (random) model. In this study, dispersion parameters such as variance components in the univariate case are assumed known.

MATERIALS AND METHODS

Theory

Minimum Mean Squared Error Estimation. Given a data vector y, we want to estimate a vector of parameters, S. For any estimator 0, define a loss function L(, S) such that L(, 0) = 0; i.e., the loss function is null when the estimates are equal to the true parameters. The loss function may involve absolute error, squared error, or some other function describing the consequences of basing a decision on 0 when 0 contains the true parameters (12). The estimator 0 is chosen to minimize the risk, or expected loss

\[ R(0, *) = E_{y}[L(0, *)] = \int_{\mathcal{R}_y} L(0, *) p(y|0) \, dy. \]  

[1]

In Equation [1], p(y|0) is the joint density of the observations (or likelihood), and \( \mathcal{R}_y \) is the sample space. For example, consider estimating fixed effects in the general mixed linear model

\[ y = X\beta + Zu + e \]  

[2]

where \( y \sim N(X\beta, V) \), \( u \sim N(0, G) \), and \( e \sim N(0, R) \). The vectors \( \beta, u, \) and \( e \) are fixed, random, and residual effects, respectively; \( X \) and \( Z \) are incidence matrices, and \( V, G, \) and \( R \) are covariance matrices. It is assumed that \( X \) has full column rank. The quadratic loss and risk functions are

\[ L(\hat{\beta}, \beta) = (\hat{\beta} - \beta)'(\hat{\beta} - \beta) \]  

[3]

\[ R(\hat{\beta}, \beta) = \text{tr} \{ \text{Var}(\hat{\beta}) + E(\hat{\beta} - \beta)(\hat{\beta} - \beta)' \} \]  

[4]

where \( \hat{\beta} \) is an estimator of \( \beta \). If the vector 0 is used as an estimator of \( \beta \), \( R(0, \beta) = \beta'\beta \). Thus, 0 may be a very good estimator when the elements of \( \beta \) are "small" in value, but a very poor one when they are "large". If ML(\( \beta \)) = BLUE(\( \beta \)) is used as an estimator, the bias term in Equation [4] is null and

\[ R(\hat{\beta}, \beta) = \text{tr} \{ \text{Var}(\hat{\beta}) \} = \text{tr}(X'V^{-1}X)^{-1}. \]  

[5]

Note that when the elements of \( \beta \) are small, the biased estimator 0 can have smaller risk than ML(\( \beta \)). Also, the risk of \( \text{ML}(\beta) \) is constant throughout the parameter space. The ML estimator has minimum risk in the class of linear, unbiased estimators under squared error loss (9). However, certain nonlinear estimators, though biased, may have risk smaller than ML. We will consider estimators of the form \( \beta^* = c\hat{\beta} \) where \( \hat{\beta} \) is the ML estimator and \( c \) is a
scalar constant to be determined. If $c > 1$, then $\hat{\beta}^*$ is an “expansion-type” estimator, and the absolute value of elements of $\hat{\beta}^*$ will be larger than those of $\hat{\beta}$. If $c < 1$, then $\hat{\beta}^*$ is said to be a “shrinkage-type” estimator, and the magnitude of the estimates will be reduced, relative to ML. Replacing $\hat{\beta}$ by $\hat{\beta}^*$ in Equation [4] and taking expectations

$$R(c\hat{\beta}, \hat{\beta}) = c^2 \text{tr}(X'V^{-1}X)^{-1} + (c - 1)^2 \beta'\beta$$

so the risk depends on $c$ and on $\beta$. If $c = 1$, Equation [6] gives the risk of the ML estimator; for $c > 1$ the risk will be larger than that of ML throughout the parameter space. Hence, expansion-type estimators are of no interest from this perspective. By taking derivatives, it can be shown (5) that the value of $c$ that minimizes $R(c\hat{\beta}, \hat{\beta})$ is

$$c = \frac{\beta'\beta}{\beta'\beta + \text{tr}(X'V^{-1}X)^{-1}}$$

Because $0 \leq c \leq 1$, $\hat{\beta}^*$ is a shrinkage-type estimator, and we will refer to it as the minimum MSE (MMSE) estimator. Thus,

$$\beta_{\text{MMSE}} = c\hat{\beta} = \left[ \frac{\beta'\beta}{\beta'\beta + \text{tr}(X'V^{-1}X)^{-1}} \right] \hat{\beta}$$

Observe that when the likelihood is highly informative about $\beta$, $\text{tr}(X'V^{-1}X)^{-1} \to 0$, and $\beta_{\text{MMSE}} \to \hat{\beta}$. Conversely, when the data contain little information about $\beta$, $\beta_{\text{MMSE}} \to 0$. Otherwise, $c$ depends on the true value of $\beta$. Because $\beta$ is unknown, one can replace $\beta$ in Equation [8] by $\hat{\beta}$ to obtain an estimated minimum MSE estimator, $\hat{\beta}_{\text{MMSE}}$. Although $\beta_{\text{MMSE}}$ is closer, in the MSE sense, to $\hat{\beta}$ than $\hat{\beta}$, this may not be so for $\hat{\beta}_{\text{MMSE}}$. The last estimator is nonlinear, so evaluating its risk is difficult by analytical means.

The James-Stein Estimator in the Mixed Linear Model. James and Stein (10) derived a shrinkage-type estimator for an orthonormal linear model that does not require known $\beta$. We extend their developments to the mixed linear model expressed in its equivalent form:

$$y = X\beta + \epsilon \quad \epsilon \sim N(0, V)$$

Premultiply Equation [9] by $(X'V^{-1}X)^{-1} X'V^{-1}$ to obtain

$$\hat{\beta} = \beta + (X'V^{-1}X)^{-1} X'V^{-1}\epsilon$$

where $(X'V^{-1}X)^{-1} X'V^{-1}\epsilon$ is the ML estimator of $\beta$. Because $X'V^{-1}X$ is positive definite, this matrix allows the decomposition

$$X'V^{-1}X = S^{1/2}S^{1/2}$$

where $S^{1/2}$ is symmetric and nonsingular. Then, from Equation [10]

$$S^{1/2}\hat{\beta} = S^{1/2}\beta + S^{1/2}(X'V^{-1}X)^{-1}X'V^{-1}\epsilon$$

Defining $\alpha = S^{1/2}\beta$, $\hat{\alpha} = S^{1/2}\hat{\beta}$, and $\epsilon^* = S^{1/2}(X'V^{-1}X)^{-1}X'V^{-1}\epsilon$, Equation [11] is expressible as

$$\hat{\alpha} = \alpha + \epsilon^*.$$  

It is easy to verify that $\hat{\alpha} \sim N(\alpha, I)$, and Equation [12] is often referred to as an orthonormal model (12). From Equation [5], $R(\hat{\alpha}, \alpha) = p$, where $p$ is the order of $\alpha$ (or $\beta$).

Let the risk be

$$E((\hat{\alpha} - \alpha)'(\hat{\alpha} - \alpha)) = E((\hat{\beta} - \beta)'X'V^{-1}X(\hat{\beta} - \beta))$$

where $\hat{\alpha}$ is an estimator of $\alpha$.

The JS estimator (10), which is nonlinear, has the form

$$\alpha_{JS} = \left[ 1 - \frac{a_{JS}}{\hat{a}_{\text{JS}}} \right] \hat{\alpha}$$

where $a_{JS}$ is a positive scalar to be determined. Now

$$E(\alpha_{JS}) = \alpha - a_{JS} E\left( \frac{\hat{\alpha}}{\hat{a}_{\text{JS}}} \right),$$

so the JS estimator is biased. Further, Judge et al. (12) give its risk to be
IMPROVING ESTIMATES OF FIXED EFFECTS

\[ R(\alpha^{JS}, \alpha) = p - a^{JS} \left[ 2(p - 2) - a^{JS} \right] E \left[ \frac{1}{\chi^2_p, \frac{2}{2}} \right] \]  

where \( \chi^2_p, \frac{2}{2} \) is a noncentral chi-square random variable with \( p \) degrees of freedom. Because the expected value of its reciprocal must be positive, \( R(\alpha^{JS}, \alpha) \leq R(\delta, \alpha) \) for all \( \alpha \) if and only if

\[ a^{JS}[2(p - 2) - a^{JS}] \geq 0 \]

so that \( 0 \leq a^{JS} \leq 2(p - 2) \). In order to find the value of \( a^{JS} \) that minimizes Equation [15], the condition

\[ \frac{\partial R(\alpha^{JS}, \alpha)}{\partial a^{JS}} = 0 \]

must be met. Solving Equation [16] gives \( a^{JS} = p - 2 \). Hence, the minimum risk JS estimator is

\[ a^{JS} = \left[ 1 - \frac{\alpha}{\delta a} \right] \delta. \]  

Recalling that \( \alpha = S^{1/2}\beta \) and that \( S^{1/2} \) is non-singular, the JS estimator in terms of the original parameterization becomes

\[ \beta^{JS} = \left[ 1 - \frac{p - 2}{\beta X V^{-1} X \beta} \right] \beta \]

\[ = \left[ 1 - \frac{p - 2}{\beta X V^{-1} y} \right] \beta \]  

In the mixed model equations, we have the relationship

\[ \text{Var}(\hat{\beta}) = (X'V^{-1}X)^{-1} = C_{\beta\beta} \]  

where \( C_{\beta\beta} \) is the portion of the inverse of the coefficient matrix corresponding to the fixed effects. Using this in Equation [18],

\[ \beta^{JS} = \left[ 1 - \frac{p - 2}{\beta(C_{\beta\beta})^{-1} \beta} \right] \beta. \]  

Note that \( \beta^{JS} = \beta \) when \( p = 2 \). It is instructive to note the similarity between the JS shrinker (the scalar quantity premultiplying \( \beta \) in Equation [20]) and its counterpart in Equation [8]. It can be shown (5) that the JS shrinker can be viewed as an unbiased estimator of the optimal Bayes’ shrinker, which is the value that minimizes Bayes’ risk assuming a normal prior distribution for \( \beta \). In this case, the optimal Bayes’ estimator, the mean of the posterior distribution of \( \beta \), is a matrix weighted average of the mean of the prior distribution of \( \beta \) (a null vector in the JS situation) and of the ML estimator. The matrix weights are the inverse of the variance-covariance matrix of the prior distribution of \( \beta \), and \( X'V^{-1}X \), the inverse of the variance-covariance matrix of \( \beta \). This suggests that unless the elements of \( \beta \) are small, it may be desirable to shrink the ML estimates towards a nonnull vector.

The Shrinkage to the Mean Estimator. As noted earlier, both the MSE estimator in Equation [8] and the JS estimator in Equation [20] shrink the ML estimates toward zero. Judge et al. (12) observed that a more general formulation should permit shrinkage toward a value other than the null vector. It is reasonable to conjecture that this should reduce MSE if the average of the elements of \( \beta \) is far from zero. Consider the estimator suggested by Lindley (13):

\[ \tilde{\beta} = \beta_0 + \left[ \frac{1}{\beta_0} \frac{\alpha (X'V^{-1}X)^{-1}}{\beta - \beta_0 + \alpha (X'V^{-1}X)^{-1}} \right] \beta - \beta_0 \]  

where \( \beta_0 \) is an arbitrarily specified vector, when \( \beta_0 = 0 \), \( \beta \) reduces to \( \bar{\beta} \) in Equation [8]. In this paper, we consider the specification \( \beta_0 = 1\delta \), where \( I \) is a \( p \times 1 \) vector of ones and

\[ \delta = \frac{\sum_{i=1}^{p} \beta_i}{p}, \] a scalar parameter, which is the average of the elements of \( \beta \). For this specification, we refer to \( \tilde{\beta} \) in Equation [21] as the “shrinkage to the mean” (STM) estimator.

Replacing \( \delta \) by \( \delta = \frac{\sum_{i=1}^{p} \beta_i}{p} \), the average of the ML estimates of fixed effects, we obtain an estimated shrinkage to the mean (ESTM) estimator.

Evaluation of Estimators

Simulation of Data. Properties of the estimators defined previously were studied by
TABLE 1. Structure of the simulated data.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of groups (G)</td>
<td>20</td>
<td>20</td>
<td>15</td>
</tr>
<tr>
<td>Number of sires per group (S)</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Number of progeny per sire (n)</td>
<td>5</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>Heritability (h²)</td>
<td>0.01</td>
<td>0.10</td>
<td>0.25</td>
</tr>
<tr>
<td>α = σ²/s²</td>
<td>399</td>
<td>39</td>
<td>15</td>
</tr>
<tr>
<td>Residual variance (σ²ε)</td>
<td>10</td>
<td>5</td>
<td>250,000</td>
</tr>
<tr>
<td>“Base” group effect (G₁)</td>
<td>50</td>
<td>8</td>
<td>4500</td>
</tr>
</tbody>
</table>

Simulation. Data were simulated using the balanced mixed linear model

\[
y_{ijk} = G_i + s_{ij} + e_{ijk}
\]

where

\[
i = 1, 2, \ldots, G; \\
j = 1, 2, \ldots, S \text{ for all } i; \text{ and} \\
k = 1, 2, \ldots, n \text{ for all } (i, j).
\]

\[E(y_{ijk}) = G_i \]

\[s_{ij} \sim \text{NIID } (0, \sigma^2_s)\]

\[e_{ijk} \sim \text{NIID } (0, \sigma^2_e)\]

with \(s_{ij}, e_{ijk}\) being mutually independent for all such pairs. In Equation [22], \(y_{ijk}\) is the record of progeny \(k\) of sire \(j\) (random) in genetic group \(i\) (fixed). In the balanced model, the estimators to be evaluated can be written in scalar form. Three scenarios were simulated (Table 1). The first scenario may be thought of as describing a fitness or reproductive trait, the second could be viewed as a model for litter size in swine, and the third as corresponding to milk yield in dairy cattle. Parameters used in the simulation are displayed in Table 1. Group effects were equally spaced, starting from the “base group”, and the difference between the largest and smallest group effects was two genetic standard deviations.

**Forms of the Estimators of Fixed Effects.** Letting \(y_{ij}\) be the average of the progeny of sire \(j\) in group \(i\), the six estimators of group effects can be expressed as

\[
g_{\text{ML}} = \frac{\sum_{j=1}^{S} \frac{y_{ij}}{S}}{S}
\]

\[
g_{\text{JS}} = \left[ 1 - \frac{(G - 2)V_ε}{S \sum_{i=1}^{G} \frac{\hat{g}_i^2}{S} + V_ε} \right] \hat{g}_i
\]

\[
g_{\text{MMSE}} = \left[ 1 - \frac{GV_ε}{S \sum_{i=1}^{G} \hat{g}_i^2 + GV_ε} \right] \hat{g}_i
\]

\[
g_{\text{EMMSE}} = \left[ 1 - \frac{GV_ε}{S \sum_{i=1}^{G} (\hat{g}_i - \bar{g})^2 + GV_ε} \right] \hat{g}_i
\]

\[
g_{\text{STM}} = \bar{g} + \left[ 1 - \frac{GV_ε}{S \sum_{i=1}^{G} (\hat{g}_i - \bar{g})^2 + GV_ε} \right] (\hat{g}_i - \bar{g})
\]
IMPROVING ESTIMATES OF FIXED EFFECTS

6) ESTM:

\[ b_{estm} = \bar{g} + \left[ 1 - \frac{gV_e}{\sum_{i=1}^{S} (g_i - \bar{g})^2 + V_e} (\bar{g}_i - \bar{g}) \right] \]

In these equations, \( S \) is the number of sires per group, \( G \) is the number of groups, \( g_i \) is the true effect of group \( i \), \( \bar{g} \) is the mean of the true group effects, and \( \bar{g} \) is the mean of the ML estimates of group effects.

Two types of problems were considered: estimation of individual fixed effects (groups) and estimation of a function of fixed effects, which we refer to as genetic trend. Groups were viewed as the mean of breeding values in a population evolving over a time scale; because group effects were equally spaced, true genetic trend is simply the difference between two successive group effects. Genetic trend was assessed for each estimator as the regression of estimated group effect on group number. The MSE of estimates of group effects and of genetic trend were evaluated from 100 replications of the appropriate scenario.

RESULTS AND DISCUSSION

Sampling properties of the MMSE estimator were similar to those of EMMSE. The same was true for the STM and ESTM estimators. Therefore, because MMSE and STM require knowing \( \beta \), only EMMSE and ESTM will be discussed further.

The relative performance of the estimators was virtually identical in the three simulation scenarios, so only results for the second scenario will be presented. This suggests that the results are representative of what can be expected from a wider range of parameter values, at least for this balanced group plus sire model.

As shown in Figure 1, the mean empirical MSE of group estimates was similar for ML, JS, and EMMSE but was almost 50% lower for the ESTM estimator. Improvement with the JS estimator was not as large as has been suggested (4) because the mean of the group effects in our simulation was far from zero. The ESTM estimator seems superior for estimation of individual group effects. However, the ESTM estimator is biased, as depicted in Figure 2, especially for group effects that are far from their mean value. This illustrates the well-known result that it is possible to improve upon an unbiased estimator in the MSE sense by allowing bias. The extent to which the bias is tolerable depends on the problem in question. The empirical MSE of estimated genetic trend was similar for ML, JS, and EMMSE but was substantially larger for the ESTM estimator (Figure 3). This was due to its larger downward bias, as shown in Figure 4, in spite of its smaller Monte Carlo variance (Figure 5). The reduction in variance was insufficient to compensate for the larger bias because of the bias in estimation of individual group effects, as discussed and illustrated in Figure 2.

With ESTM, individual fixed effects are regressed toward the overall mean of the fixed effects, so (early) groups with a “small” true effect are estimated with an upward bias, toward the mean, whereas (later) groups with “large” true effects are estimated with a downward bias. The end result of this is a “flatten-
Among the estimators with better MSE properties than ML, the MMSE estimator has smallest risk in a certain class but requires known \( \beta \). The properties of the estimator obtained when \( \beta \) is replaced by \( \hat{\beta} = \text{ML}(\beta) \) are difficult to evaluate analytically, so one must resort to simulation, as in the present study. The JS estimator gives smaller risk than ML without requiring known \( \beta \). However, the improvement obtained from the JS, MMSE, and EMMSE estimators, which regress ML estimates toward zero, may be small when the true fixed effects are far from zero. This occurs because these estimators are not invariant to changes in the location parameters (15).

This led to the development of the STM estimator, a modification of the MMSE estimator in which the elements of \( \hat{\beta} \) are regressed toward the mean of the elements of \( \beta \), rather than toward zero (13). Because this estimator

... of the fitted line of genetic trend (Figure 6). The JS and EMMSE estimators, which shrink all the group effects toward zero, resulted in a small negative bias in each group estimate but little bias in estimated genetic trend.

The inadmissibility of ML led Lindley [cited in (11, page 85)] to state the following:

The result of Stein undermines the most important practical technique in statistics. . . . The next time you do an analysis of variance or fit a regression surface (a line is all right!) remember you are, for sure, using an unsound procedure. Unsound, that is, by the criteria you have presumably been using to justify the procedure. . . .

Stein's result potentially undermines the most widely used procedure, BLUE, for estimation of fixed effects in animal breeding.
requires knowledge of the true value of $\beta$, we suggest replacing this by its ML estimate. If there is more than one set of fixed effects in the model, different shrinkage coefficients should probably be used (5), so that differing sets of estimates are regressed to their corresponding means such that the degree of shrinkage depends on the amount of information for each set of effects.

The simulation results suggest that all estimators proposed here give estimates of fixed effects with smaller MSE than ML. Perhaps BLUP, a function of BLUE, can also be improved in a MSE context, at least when the predictand is a location variant (e.g., group plus sire). However, when ESTM was used to estimate group effects, at least for the equally spaced structure considered, the estimates of low merit (early) genetic groups were biased upward, whereas estimates of high merit (later) groups were biased downward. One may then speculate that predictions of older animals would also be biased upward and those of younger animals biased downward (L. D. Van Vleck, personal communication). This seems undesirable and could have a detrimental effect on genetic progress if the predictions were to be used as selection criteria.

Perhaps this can be remedied by introducing an additional model for the fixed effects. Suppose that we put

$$\hat{\beta} = K\lambda + \left[ 1 - \frac{1}{(\lambda - K\hat{\lambda})/(\hat{\beta} - K\lambda) + u (X'X)^{-1}} \right] \hat{\beta} - K\lambda$$

and the ESTM statistic would be obtained by replacing $\beta$ and $\lambda$ by $\hat{\beta}$ and $\hat{\lambda}$, respectively.

CONCLUSIONS

Shrinkage-type estimators have been widely used in statistical applications for many years. Such estimators could prove useful in analysis of agricultural experiments for estimation of effects such as diet or treatment, especially when information is limited and the number of levels is large. Unless appropriately amended, these biased estimators may be poor when estimating a "structural" function of fixed effects such as genetic trend. In these cases, additional theoretical developments are needed, and some possibilities were suggested. Further study, including the effects of unbalancedness, nonnormality, selection, and unknown dispersion parameters, is needed before it can be recommended that biased estimators should replace BLUE in animal breeding applications. In an unbalanced model, the differences between estimators may be larger than those reported here.

REFERENCES