Journal of Multivariate Analysis **67**, 35–48 (1998)
Article No. MV981749

Improved Estimation in Measurement Error Models Through Stein Rule Procedure

Shalabh

University of Jammu, Jammu, India

Received April 16, 1996; revised July 18, 1997

This paper examines the role of Stein estimation in a linear ultrastructural form of the measurement errors model. It is demonstrated that the application of Stein rule estimation to the matrix of true values of regressors leads to the overcoming of the inconsistency of the least squares procedure and yields consistent estimators of regression coefficients. A further application may improve the efficiency properties of the estimators of regression coefficients. It is observed that the proposed family of estimators under some constraint on the characterizing scalar dominates the conventional consistent estimator with respect to the criterion of asymptotic risk under a specific quadratic loss function. Then the problem of prediction of the values of the study variable within the sample is considered, and it is found that the predictors based on the proposed family of estimators are always more efficient than the predictors based on the conventional estimator according to asymptotic predictive mean squared error criterion, although both are biased. © 1998 Academic Press

AMS 1991 subject classifications: 62J07, 62F11.

Key words and phrases: Measurement errors; ultrastructural model; Stein rule estimators; predictions; mean squared error matrix criterion.

1. INTRODUCTION

Failure of the least squares procedure to yield even consistent estimators of the regression parameters in linear measurement error models has led to the search for suitable additional information and ways to incorporate such information. Each of the approaches has its own advantages and difficulties. One such popular approach involves the specification of the variance covariance matrix of measurement errors of the regressors. Utilizing it to overcome the problem of inconsistency in the context of functional variant of the model, Schneeweiss (1976) has presented the adjusted or corrected least squares estimator. He has demonstrated that the estimator is consistent. Further, he has derived a formula for its asymptotic variance covariance matrix. In this paper, we consider the ultrastructural model, which encompasses functional, structural, and classical regression models as particular cases, and demonstrate that the adjusted or corrected

0047-259X/98 $25.00

estimator can be interpreted as the least squares estimator in a model in which the matrix of the true values of the regressors is replaced by a biased estimator stemming from Stein estimation. It is interesting to note that Stein estimation eliminates the inconsistency of the traditional least squares estimator; see also Srivastava and Shalabh (1997), Stanley (1986. 88), Van Hoa (1986)) and Whittemore (1989). This has prompted us to investigate the impact of a repeat application of Stein estimation.

The plan of this paper is as follows. We present the specification of the ultrastructural form of the linear measurement error model in Section 2. In Section 3, we state the adjusted or corrected least squares estimator for the coefficient vector and furnish an interpretation from the viewpoint of Stein estimation. Then a family of improved estimators arising from another application of Stein estimation of the regression coefficient vector is presented. Based on these, the predictors for the vectors of the values of the study variable are also presented. Section 4 analyses the asymptotic properties of the coefficient estimators. All the estimators are found to be consistent but biased. Choosing a specific loss matrix of the quadratic loss function and taking the performance criterion as asymptotic risk under the specified loss function, the dominance of the presented family of estimators over the adjusted or corrected least squares estimator is examined. Section 5 studies the asymptotic properties of the predictors. It is seen that all the predictors are biased but the predictor based on the proposed family of estimators is always more efficient than the prediction based on the adjusted or corrected least squares estimator according to the criterion of asymptotic predictive mean squared error. Finally, an Appendix is given in which the expressions for certain expectations needed in the study of asymptotic properties are derived.

2. MODEL SPECIFICATION

Consider the measurement error model

$$\begin{aligned} y_{\text{obs}} &= y + u = X\beta + u \\ X_{\text{obs}} &= X + V, \end{aligned} \tag{2.1}$$

where y is a $n \times 1$ vector of the unobserved study variable and y_{obs} is its observed counterpart, X is a $n \times p$ matrix of the values of p unobservable regressors and X_{obs} is its observed counterpart, β is a $p \times 1$ vector of regression coefficients in the relationship connecting the unobservable variables in the model, u is a $n \times 1$ vector of measurement errors in the study variable, and V is a $n \times p$ matrix of measurement errors in regressors.

The conventional disturbance term in the regression relationship is assumed to be subsumed in the measurement error term in the study variable without disturbing any salient feature of the model.

Let M be the matrix of means of regressors so that we can write

$$X = M + W, \tag{2.2}$$

where W is a random matrix of order $n \times p$.

It is assumed that the elements of u are independently and identically distributed following a normal distribution with mean 0 and variance σ_u^2. Likewise, the row vectors of V are assumed to be independently and identically distributed following a multivariate normal distribution with mean vector 0 and variance covariance matrix Σ_V. Similarly, the row vectors of W are assumed to be independently and identically distributed following a multivariate normal distribution with mean vector 0 and variance covariance matrix Σ_W. Further, we assume that u, V, and W are stochastically independent. These specifications can be relaxed at the cost of slight algebraic complexity but without any conceptual difficulty; see, e.g., Schneeweiss (1976). Thus, we have

$$\begin{aligned}
E(u) &= 0, \qquad E(uu') = \sigma_u^2 I_n, \\
E(V) &= 0, \qquad \frac{1}{n}E(V'V) = \Sigma_V, \\
E(W) &= 0, \qquad \frac{1}{n}E(W'W) = \Sigma_W, \\
E(u'V) &= E(u'W) = 0, \\
E(V'W) &= 0.
\end{aligned} \tag{2.3}$$

Finally, it is assumed that the limiting form of the matrix $n^{-1}(M'M)$ as n tends to infinity is a finite and nonsingular matrix.

When all the row vectors of M are assumed to be identical, implying that rows of X are random and independent, having some multivariate distribution, we get the specification of a structural model. When W is taken identically equal to a null matrix implying that $\Sigma_W = 0$ and consequently that the matrix X is fixed but is measured with error, we obtain the specification of a functional model. When both V and W are identically equal to a null matrix, implying that $\Sigma_V = \Sigma_W = 0$ and consequently that X is fixed and is measured without any measurement error, we get the classical regression model. Thus the ultrastructural model provides a general framework for the study of three interesting models in a unified manner.

3. COEFFICIENT ESTIMATORS AND PREDICTORS

If we apply least squares method for the estimation of β, we get the estimator

$$\hat{\beta} = (X'_{\text{obs}} X_{\text{obs}})^{-1} X'_{\text{obs}} y_{\text{obs}}. \tag{3.1}$$

Similarly, the Stein rule estimator of β is given by

$$\hat{\beta}_s = \left[1 - \left(\frac{k}{n-p+2} \right) \frac{(y_{\text{obs}} - X_{\text{obs}}\hat{\beta})' (y_{\text{obs}} - X_{\text{obs}}\hat{\beta})}{\hat{\beta}' X'_{\text{obs}} X_{\text{obs}} \hat{\beta}} \right] \hat{\beta}, \tag{3.2}$$

where k is the characterizing scalar assumed to be positive and non-stochastic; see, e.g., Judge and Bock (1978) and Stanley (1989).

Let us assume for a moment that measurement errors in explanatory variables are absent so that $X = X_{\text{obs}}$ is fixed. Under this specification, it is well known that $\hat{\beta}$ is the best estimator in the class of all linear and unbiased estimators while $\hat{\beta}_S$ is a nonlinear and biased estimator.

If we define

$$\Delta^* = E(\hat{\beta} - \beta)(\hat{\beta} - \beta)' - E(\hat{\beta}_S - \beta)(\hat{\beta}_S - \beta)'$$

then Δ^*, up to order $O(n^{-2})$, is given by

$$\Delta^* = \frac{2k\sigma_u^4}{\beta' S \beta} \left[S^{-1} - \left(\frac{4+k}{2\beta' S \beta} \right) \beta\beta' \right], \tag{3.3}$$

where $S = n^{-1} X'_{\text{obs}} X_{\text{obs}}$; see Vinod and Srivastava (1995, p. 83).

Applying Lemma 1 of the Appendix, we observe that Δ^* is not a positive definite matrix for all positive values of k, which means that Stein rule estimator $\hat{\beta}_S$ is not superior to the least squares estimator $\hat{\beta}$ with respect to the criterion of mean squared error matrix to the order of our approximation. Similarly, using Lemma 2 of the Appendix, it is seen that the matrix $(-\Delta^*)$ cannot be non-negative definite except in the trivial case of $p = 1$. Thus we observe that neither of the two estimators $\hat{\beta}$ and $\hat{\beta}_S$ is superior to the other according to the mean squared error matrix criterion.

If we compare $\hat{\beta}$ and $\hat{\beta}_S$ with respect to the criterion of weighted mean squared error to the order $O(n^{-2})$, then trace of matrix $(\Delta^* S)$ is positive, implying the superiority of $\hat{\beta}_S$ over $\hat{\beta}$ when k is less than $2(p-2)$ provided that p exceeds 2. The same condition is found, it is interesting to note, when exact expressions of weighted mean squared error are employed; see, e.g., Judge and Bock (1978).

Next, let us consider the case when measurement errors in explanatory variables are present. Under this specification, both the estimators $\hat{\beta}$ and $\hat{\beta}_S$

are generally inconsistent for β. Further, there is no way to obtain a consistent estimator unless some additional information about the model is available; see, e.g., Fuller (1987). Let us therefore suppose that the variance covariance matrix Σ_V is known, which is one of the popular assumptions in the literature. Then a consistent estimator of β is given by

$$b = (X'_{\text{obs}} X_{\text{obs}} - n\Sigma_v)^{-1} X'_{\text{obs}} y_{\text{obs}}. \tag{3.4}$$

This estimator is known as the adjusted or corrected least squares estimator, for it is derived from the least squares estimator by incorporating an adjustment or correction arising from the expression for its inconsistency; see, e.g., Friedman *et al.* (1991), Moran (1971), and Schneeweiss (1985).

We have observed that the structural and functional (and also classical regression) variants of the model arise from the assumed nature of the matrix X. In view of this, let us consider an analysis conditional upon X for providing an interpretation to the estimators (3.1) and (3.4). For this purpose, we observe from (2.1) of the model that X_{obs} serves as a weakly unbiased estimator of X in the sense that $E(X) = E(X_{\text{obs}})$ by virtue of (2.3).

If we define

$$C = I_p - n(X'_{\text{obs}} X_{\text{obs}})^{-1} \Sigma_v = I_p - S^{-1}\Sigma_v \tag{3.5}$$

we see that $X_{\text{obs}} C$ is a biased estimator in the sense that $E(X_{\text{obs}} C)$ is not equal to $E(X_{\text{obs}})$. The estimator $X_{\text{obs}} C$ is obtained from an application of Stein estimation to (3.4) when Σ_V is known; see, e.g., Stein (1973) and Zheng (1986). It may be noticed that we have taken a slightly different form of the estimator $X_{\text{obs}} C$; the scalar n in it is indeed the limiting value of the scalar used by Stein (1973). This is of little consequence, as we shall be concerned with the asymptotic properties only.

Thus if X in the first equation of (2.1) is replaced by the unbiased estimator X_{obs} and then the least squares procedure is applied, we get an inconsistent estimator (3.1) of β. On the other hand, if we replace X by a biased estimator $X_{\text{obs}} C$, we obtain a consistent estimator (3.4). This is an interesting observation related to an important and perhaps relatively unexplored aspect of Stein estimation, which is well known for reducing variability around true parameter values but not so well known for eliminating inconsistency.

Now, given a consistent estimator b, one can use Stein estimation to construct a family of estimators of β for achieving some gain in efficiency. It is defined by

$$b_S = \left[1 - \left(\frac{k}{n-p+2}\right) \frac{(y_{\text{obs}} - X_{\text{obs}} b)' \, (y_{\text{obs}} - X_{\text{obs}} b)}{b'(X'_{\text{obs}} X_{\text{obs}} - n\Sigma_v) \, b}\right] b, \tag{3.6}$$

where k is any positive and nonstochastic scalar, independent of n, characterizing the estimator.

Next, consider the problem of predicting the values of the study variable within the sample. For this purpose, the two natural predictors arising from consistent estimation of β are

$$T = X_{\text{obs}} b \tag{3.7}$$

$$T_S = X_{\text{obs}} b_S. \tag{3.8}$$

Similar predictions can be constructed when the objective is to predict some future values or values outside the sample for one or more sets of preassigned values of the explanatory variables.

4. PROPERTIES OF COEFFICIENT ESTIMATORS

In order to study the asymptotic properties of the estimates b and b_S, we first observe from (2.1) and (2.2) that

$$\begin{aligned} y_{\text{obs}} &= M\beta + (u + W\beta) \\ X_{\text{obs}} &= M + (V + W). \end{aligned} \tag{4.1}$$

Using these in (3.4), writing

$$\begin{aligned} H = n^{-1/2}[&M'(V+W) + (V+W)'M + V'W + W'V + (V'V - n\Sigma_v) \\ &+ (W'W - n\Sigma_w)] \end{aligned} \tag{4.2}$$

$$h = n^{-1/2}[(M+W)'(u - V\beta) + V'u - (V'V - n\Sigma_v)\beta], \tag{4.3}$$

and observing that the elements of the matrix H and the vector h are of order $O_p(1)$, we find

$$\begin{aligned} (b - \beta) &= n^{-1/2}(I_p + n^{-1/2}\Sigma_X^{-1}H)^{-1}\Sigma_X^{-1}h \\ &= n^{-1/2}\Sigma_X^{-1}h - n^{-1}\Sigma_X^{-1}H\Sigma_X^{-1}h + O_p(n^{-3/2}), \end{aligned} \tag{4.4}$$

where

$$\Sigma_X = n^{-1}M'M + \Sigma_w. \tag{4.5}$$

Noticing that $E(h)=0$ from (2.3), the bias vector of b is

$$E(b-\beta)=-n^{-1}\Sigma_X^{-1}E(H\Sigma_X^{-1}h)+O(n^{-3/2}). \tag{4.6}$$

Similarly, from (3.6), we have

$$(b_S-\beta)=(b-\beta)-\left(\frac{k}{n-p+2}\right)\frac{(y_{\text{obs}}-X_{\text{obs}}b)'\,(y_{\text{obs}}-X_{\text{obs}}b)}{b'(X'_{\text{obs}}X_{\text{obs}}-n\Sigma_v)\,b}$$
$$\times[\beta+(b-\beta)]. \tag{4.7}$$

Now consider the ratio

$$\begin{aligned}\frac{(y_{\text{obs}}-X_{\text{obs}}b)'\,(y_{\text{obs}}-X_{\text{obs}}b)}{b'(X'_{\text{obs}}X_{\text{obs}}-n\Sigma_v)\,b}&=\frac{(\sigma_u^2+\beta'\Sigma_v\beta)(1+n^{-1/2}\varepsilon)+O_p(n^{-1})}{\beta'\Sigma_X\beta+n^{-1/2}(\beta'H\beta+2\beta'h)+O_p(n^{-1})}\\&=\theta-n^{-1/2}\theta\left(\frac{\beta'H\beta+2\beta'h}{\beta'\Sigma_X\beta}-\varepsilon\right)+O_p(n^{-1}),\end{aligned} \tag{4.8}$$

where

$$\theta=\frac{\sigma_u^2+\beta'\Sigma_v\beta}{\beta'\Sigma_X\beta}$$

$$\varepsilon=\frac{(u-V\beta)'\,(u-V\beta)}{n^{1/2}(\sigma_u^2+\beta'\Sigma_v\beta)}-n^{1/2}.$$

Substituting this along with (4.4) into (4.7), we get

$$(b_S-\beta)=n^{-1/2}\Sigma_X^{-1}h-n^{-1}[\Sigma_X^{-1}H\Sigma_X^{-1}h+k\theta\beta]+O_p(n^{-3/2}), \tag{4.9}$$

whence the bias vector of b_S is

$$E(b_S-\beta)=-n^{-1}E[\Sigma_X^{-1}H\Sigma_X^{-1}h+k\theta\beta]+O(n^{-3/2}). \tag{4.10}$$

Utilizing the expression for $E(H\Sigma_X^{-1}h)$ derived in the Appendix, we obtain from (4.6) and (4.10) the following results.

THEOREM I. *The large sample asymptotic approximations for the bias vectors, to order $O(n^{-1})$, of the estimators b and b_S are given by*

$$B(b)=n^{-1}[(p+\operatorname{tr}\Sigma_X^{-1}\Sigma_v+1)\,\Sigma_X^{-1}\Sigma_v\beta+(\Sigma_X^{-1}\Sigma_v)^2\,\beta] \tag{4.11}$$

$$B(b_S)=n^{-1}[(p+\operatorname{tr}\Sigma_X^{-1}\Sigma_v+1)\,\Sigma_X^{-1}\Sigma_v\beta+(\Sigma_X^{-1}\Sigma_v)^2\,\beta-k\theta\beta]. \tag{4.12}$$

It is thus seen that both the estimators b and b_S are biased, at least to the order of our approximation. It is, however, difficult to draw any clear inference about the relative magnitude of bias.

Next, let us compare the mean squared error matrices of the estimators b and b_S. It is seen from (4.4) and (4.9) that

$$E(b-\beta)(b-\beta)' = E(b_S-\beta)(b_S-\beta)' \\ = n^{-1}\Sigma_X^{-1} E(hh')\,\Sigma_X^{-1} + O(n^{-3/2}), \tag{4.13}$$

whence it follows that both the estimators are asymptotically equivalent in the sense that they are consistent and have the same mean squared error matrix up to order $O(n^{-1})$ or the same asymptotic variance covariance matrix.

In order to discriminate and to examine the dominance of one estimator over the other, let us consider the mean squared error matrix to order $O(n^{-2})$. For this purpose, let us consider the difference

$$\begin{aligned}\Delta &= E(b-\beta)(b-\beta)' - E(b_S-\beta)(b_S-\beta)' \\ &= \left(\frac{k}{n-p+2}\right) E\left[\frac{(y_{\text{obs}}-X_{\text{obs}}b)'\,(y_{\text{obs}}-X_{\text{obs}}b)}{b'(X'_{\text{obs}}X_{\text{obs}}-n\Sigma_v)\,b}\{b(b-\beta)' + (b-\beta)\,b'\}\right] \\ &\quad - \left(\frac{k}{n-p+2}\right)^2 E\left[\left\{\frac{(y_{\text{obs}}-X_{\text{obs}}b)'\,(y_{\text{obs}}-X_{\text{obs}}b)}{b'(X'_{\text{obs}}X_{\text{obs}}-n\Sigma_v)\,b}\right\}^2 bb'\right].\end{aligned}$$

Using (4.4), we can express

$$\Delta = n^{-3/2}k\theta E(\Sigma_X^{-1}h\beta' + \beta h'\Sigma_X^{-1}) + n^{-2}k\theta[E(D) - k\theta\beta\beta'] + O(n^{-5/2}), \tag{4.14}$$

where

$$D = 2\Sigma_X^{-1}hh'\Sigma_X^{-1} - \Sigma_X^{-1}H\Sigma_X^{-1}h\beta' - \beta h'\Sigma_X^{-1}H\Sigma_X^{-1} \\ + \left(\varepsilon - \frac{\beta' H\beta + 2\beta' h}{\beta'\Sigma_X\beta}\right)(\Sigma_X^{-1}h\beta' + \beta h'\Sigma_X^{-1}). \tag{4.15}$$

Substituting the required expectations from the Appendix into (4.13) and (4.14), we obtain the results as follows.

THEOREM II. *The asymptotic variance covariance matrix, i.e., the mean squared error matrix to order $O(n^{-1})$, of the estimators b and b_S is given by*

$$n^{-1}[(\sigma_u^2+\beta'\Sigma_v\beta)\,\Sigma_X^{-1}+\sigma_u^2\Sigma_X^{-1}\Sigma_v\Sigma_X^{-1}+\Sigma_X^{-1}\Sigma_v\beta\beta'\Sigma_v\Sigma_X^{-1}]. \qquad (4.16)$$

Further, the difference in the mean squared error matrices to order $O(n^{-2})$ is

$$\begin{aligned}\Delta &= E(b-\beta)(b-\beta)' - E(b_S-\beta)(b_S-\beta)'\\ &= n^{-2}\theta kF, \end{aligned} \qquad (4.17)$$

where

$$\begin{aligned}F = {} & 2[(\sigma_u^2+\beta'\Sigma_v\beta)\,\Sigma_X^{-1}+\sigma_u^2\Sigma_X^{-1}\Sigma_v\Sigma_X^{-1}+\Sigma_X^{-1}\Sigma_v\beta\beta'\Sigma_v\Sigma_X^{-1}]\\ & +(\Sigma_X^{-1}\Sigma_v)^2\,\beta\beta'+\beta\beta'(\Sigma_v\Sigma_X^{-1})^2\\ & +\left[p+\operatorname{tr}\Sigma_X^{-1}\Sigma_v-1-\frac{2\sigma_u^2}{\beta'\Sigma_X\beta}\right](\beta\beta'\Sigma_v\Sigma_X^{-1}+\Sigma_X^{-1}\Sigma_v\beta\beta')\\ & -\left[\frac{(4+k)\,\sigma_u^2+k\beta'\Sigma_v\beta}{\beta'\Sigma_X\beta}\right]\beta\beta'. \end{aligned} \qquad (4.18)$$

The expression (4.16) in the special case of functional model has been obtained by Schneeweiss (1976).

It is difficult to examine whether F is positive definite or not, and consequently there is no clear indication whether b_S dominates b with respect to the criterion of mean squared error matrix.

Taking Σ_X as the loss matrix, let us now compare the risk functions of the estimators under quadratic loss measure. Premultiplying Δ by Σ_X and then taking trace, we observe that the estimator b_S dominates the estimator b with respect to the criterion of risk function to order $O(n^{-2})$ when

$$\begin{aligned}k<2\Bigg[& p-2+\left(\frac{\sigma_u^2}{\sigma_u^2+\beta'\Sigma_v\beta}\right)\operatorname{tr}\Sigma_X^{-1}\Sigma_v+2\left(\frac{\beta'\Sigma_v\Sigma_X^{-1}\Sigma_v\beta}{\sigma_u^2+\beta'\Sigma_v\beta}\right)\\ & -2\left(\frac{\sigma_u^2}{\sigma_u^2+\beta'\Sigma_v\beta}\right)\left(\frac{\beta'\Sigma_v\beta}{\beta'\Sigma_X\beta}\right)+\frac{\beta'\Sigma_v\beta}{\sigma_u^2+\beta'\Sigma_v\beta}(p+1+\operatorname{tr}\Sigma_X^{-1}\Sigma_v)\Bigg] \end{aligned} \qquad (4.19)$$

provided that the expression on the right hand side of the inequality is positive.

If $r_1 \leqslant r_2 \leqslant \cdots \leqslant r_p$ denote the characteristic roots of the matrix $\Sigma_X^{-1}\Sigma_V$, then from Rao (1965, p. 59) we have

$$0 \leqslant \left(\frac{\sigma_u^2}{\sigma_u^2 + \beta'\Sigma_v\beta}\right) \leqslant 1$$

$$0 \leqslant \left(\frac{\beta'\Sigma_v\beta}{\sigma_u^2 + \beta'\Sigma_v\beta}\right) \leqslant 1$$

$$0 \leqslant \left(\frac{\beta'\Sigma_v\Sigma_X^{-1}\Sigma_v\beta}{\sigma_u^2 + \beta'\Sigma_v\beta}\right) < \left(\frac{\beta'\Sigma_v\Sigma_X^{-1}\Sigma_v\beta}{\beta'\Sigma_v\beta}\right) \leqslant r_p$$

$$r_1 \leqslant \left(\frac{\beta'\Sigma_v\beta}{\beta'\Sigma_X\beta}\right) \leqslant r_p.$$

Utilizing these results, we observe that the right hand side of the inequality (4.19) lies between $2(p-2-2r_p)$ and $4(p-\frac{1}{2}+r_p+\Sigma r_i)$. Thus we obtain a sufficient condition for the dominance of b_S over b as

$$k < 2(p-2-2r_p); \qquad p > 2(1+r_p), \tag{4.20}$$

which does not involve any unknown quantity.

When the explanatory variables contain no measurement errors so that Σ_V is a null matrix and consequently r_p is zero, the condition (4.20) reduces to

$$k < 2(p-2); \qquad p > 2, \tag{4.21}$$

which is a well known condition for the dominance of Stein rule estimators over the least squares estimators in classical regression models.

Conditions like (4.20) for some other choice of loss matrix can be deduced in a similar manner.

5. PROPERTIES OF PREDICTORS

When the aim is to predict the observed value y_{obs} of the study variable, we observe from (3.7) that

$$\begin{aligned}(T - y_{\text{obs}}) &= X_{\text{obs}}(b-\beta) - (u - V\beta)\\ &= (M + W + V)(b-\beta) - (u - V\beta),\end{aligned} \tag{5.1}$$

whence it is obvious that

$$E(T - y_{\text{obs}}) \neq 0. \tag{5.2}$$

Similarly, we have

$$E(T_S - y_{\text{obs}}) \neq 0. \tag{5.3}$$

Thus both the predictors T and T_S are biased in the sense of (5.2) and (5.3).

Next, let us compare their predictive mean squared errors. Using (4.4) and (4.9), it is seen that

$$\begin{aligned} &E(T - y_{\text{obs}})'(T - y_{\text{obs}}) - E(T_S - y_{\text{obs}})'(T_S - y_{\text{obs}}) \\ &\quad = 2(\beta'\Sigma_X\beta)^{-1}(\sigma_u^2 + \beta'\Sigma_v\beta)(\beta'\Sigma_v\beta)\,k + O(n^{-1/2}), \end{aligned} \tag{5.4}$$

which is positive, implying the superiority of T_S over T.

Similarly, if T and T_S are used for predicting y, the true values of the study variable, it can be easily verified that both are weekly biased. Further, the difference in their predictive mean squared errors is

$$\begin{aligned} &E(T - y)'(T - y) - E(T_S - y)'(T_S - y) \\ &\quad = 2(\beta'\Sigma_X\beta)^{-1}(\sigma_u^2 + \beta'\Sigma_v\beta)(\beta'\Sigma_v\beta)\,k + O(n^{-1/2}), \end{aligned} \tag{5.5}$$

whence it follows that T_S is better than T.

Finally, when the aim is to predict $M\beta$, the average values of study variables, we observe that both T and T_S are weakly biased predictors of $M\beta$. Further, we have

$$\begin{aligned} &E(T - M\beta)'(T - M\beta) - E(T_S - M\beta)'(T_S - M\beta) \\ &\quad = 2(\beta'\Sigma_X\beta)^{-1}(\sigma_u^2 + \beta'\Sigma_v\beta)(\beta'\Sigma_v\beta + \beta'\Sigma_w\beta)\,k + O(n^{-1/2}), \end{aligned} \tag{5.6}$$

which again implies the superiority of T_S over T.

It is thus seen that both the predictors T and T_S are weakly biased but T_S is more efficient than T according to the criterion of the asymptotic predictive mean squared error.

APPENDIX

LEMMA 1. *If A is any $p \times p$ positive definite matrix and a is any $p \times 1$ vector, a necessary and sufficient condition for $(A^{-1} - aa')$ to be positive definite is that $a'Aa$ is less than* 1.

Proof. See Yancy *et al.* (1978).

Lemma 2. *If A is any $p\times p$ positive definite matrix and a is any $p\times 1$ vector, then the matrix $(aa'-A^{-1})$ cannot be non-negative definite for p exceeding* 1.

Proof. See Guilky and Price (1981).

Lemma 3. *We show that*

$$E(H\Sigma_X^{-1}h) = -(p+\operatorname{tr}\Sigma_X^{-1}\Sigma_v+1)\,\Sigma_v\beta - \Sigma_v\Sigma_X^{-1}\Sigma_v\beta \tag{1}$$

$$E(hh') = (\sigma_u^2+\beta'\Sigma_v\beta)\,\Sigma_X + \sigma_u^2\Sigma_v + \Sigma_v\beta\beta'\Sigma_v \tag{2}$$

$$\begin{aligned} E(D) = {} & 2[(\sigma_u^2+\beta'\Sigma_v\beta)\,\Sigma_X^{-1} + \sigma_u^2\Sigma_X^{-1}\Sigma_v\Sigma_X^{-1} + \Sigma_X^{-1}\Sigma_v\beta\beta'\Sigma_v\Sigma_X^{-1}] \\ & + (\Sigma_X^{-1}\Sigma_v)^2\,\beta\beta' + \beta\beta'(\Sigma_v\Sigma_X^{-1})^2 \\ & + \left[p+\operatorname{tr}\Sigma_X^{-1}\Sigma_v - 1 - \frac{2\sigma_u^2}{\beta'\Sigma_X\beta}\right] \\ & \times(\beta\beta'\Sigma_v\Sigma_X^{-1} + \Sigma_X^{-1}\Sigma_v\beta\beta') - \frac{4\sigma_u^2}{\beta'\Sigma_X\beta}\,\beta\beta' \end{aligned} \tag{3}$$

Proof. Let us first state the following results which can be easily deduced, for example, from Srivastava and Tiwari (1976),

$$\begin{aligned} E(V'BV) &= (\operatorname{tr} B)\,\Sigma_v \\ E(VBV') &= (\operatorname{tr} B\Sigma_v)\,I_n \\ E(V'VBV'V) &= n(n+1)\,\Sigma_v B\Sigma_v + n(\operatorname{tr} B\Sigma_v)\,\Sigma_v \\ E(VCVBV'V) &= nC'\Sigma_v B\Sigma_v + (\operatorname{tr} B\Sigma_v)\,C'\Sigma_v + C'\Sigma_v B'\Sigma_v \\ E(VCV'BVV') &= n(\operatorname{tr} C\Sigma_v)\,B\Sigma_v + B\Sigma_v(C+C')\,\Sigma_v \\ E(VCV'VBV) &= nB'\Sigma_v C'\Sigma_v + (\operatorname{tr} C\Sigma_v)\,B'\Sigma_v + B'\Sigma_v C\Sigma_v, \end{aligned} \tag{4}$$

where B and C are nonstochastic matrices of appropriate order in each case.

Similar results can be stated for the vector u and matrix W.

First consider the result (1). Using (2.3) and dropping the terms with expectation equal to a null vector, we observe from (4.2), (4.3), and (4.5) that

$$\begin{aligned} E(H\Sigma_X^{-1}h) = {} & -n^{-1}E[(M'V+V'M)\,\Sigma_X^{-1}M'V\beta + (V'W+W'V)\,\Sigma_X^{-1}W'V\beta \\ & + (V'V-n\Sigma_v)\,\Sigma_X^{-1}(V'V-n\Sigma_v)\,\beta]. \end{aligned}$$

Now employing the results (4), we find (1), after a little algebraic simplification.

Similarly, dropping the terms having expectation as a null matrix, it is easy to see that

$$E(hh') = n^{-1}E[M'(u - V\beta)(u - V\beta)' M + W'(u - V\beta)(u - V\beta)' W + V'uu'V + (V'V - n\Sigma_v)\beta\beta'(V'V - n\Sigma_v)] = (\sigma_u^2 + \beta'\Sigma_v\beta)\Sigma_X + \sigma_u^2\Sigma_v + \Sigma_v\beta\beta'\Sigma_v,$$

which proves the result (2).

Next consider the expression (4.15) for D whence we have

$$E(D) = 2\Sigma_X^{-1}E(hh')\Sigma_X^{-1} - \Sigma_X^{-1}H\Sigma_X^{-1}h\beta' - \beta h'\Sigma_X^{-1}H\Sigma_X^{-1} + E\left[\frac{(u - V\beta)'(u - V\beta)}{n^{1/2}(\sigma_u^2 + \beta'\Sigma_v\beta)}(\Sigma_X^{-1}h\beta' + \beta h'\Sigma_X^{-1})\right] - \frac{1}{\beta'\Sigma_X\beta} \times [\Sigma_X^{-1}E(h\beta'H)\beta\beta' + \beta\beta'E(H\beta h')\Sigma_X^{-1} + 2\Sigma_X^{-1}E(hh')\beta\beta' + 2\beta\beta'E(hh')\Sigma_X^{-1}]. \quad (5)$$

Using the distributional properties of u, V and W along with the results (2) and (4), it can be verified that

$$E\left[\frac{(u - V\beta)'(u - V\beta)}{n^{1/2}(\sigma_u^2 + \beta'\Sigma_v\beta)}h\right] = -2\Sigma_v\beta$$

$$E(h\beta'H) = -(\Sigma_X + \Sigma_v)\beta\beta'\Sigma_v - (\beta'\Sigma_v\beta)(\Sigma_X + \Sigma_v)$$

which substituted into (5) leads to (3).

REFERENCES

R. Friedmann, H.-J. Mittag, and A. Brandstäter, An application of approximate sample results to parameter estimation in linear errors-in-variables model, *Statist. Neerlandica* **45** (1991), 93–106.

W. A. Fuller, "Measurement Error Models," Wiley, New York, 1987.

D. K. Guilkey and J. M. Price, On comparing restricted least squares estimators, *J. Econometrics* **15** (1981), 397–404.

G. G. Judge and M. E. Bock, "The Statistical Implications of Pre-test and Stein-Rule Estimators in Econometrics," North-Holland, Amsterdam, 1978.

P. A. P. Moran, Estimating structural and functional relationship, *J. Multivariate Anal.* **1** (1971), 232–255.

C. R. Rao, "Linear Statistical Inference and Its Applications," Wiley, New York, 1965.

C. A. Robertson, Large-sample theory for the linear structural model, *Biometrika* **61** (1974), 353–359.

H. Schneeweiss, Consistent estimation of a regression with errors in the variables, *Metrika* **23** (1976), 101–115.

H. Schneeweiss, Estimating linear relations with errors-in-variables: The merging of two approaches, *in* “Contributions To Econometrics And Statistics Today” (H. Schneeweiss and H. Strecker, Eds.), Springer-Verlag, Berlin, 1985.

A. K. Srivastava and Shalabh, A new property of Stein procedure in measurement error model, *Statist. Probab. Lett.* **32** (1997), 231–234.

V. K. Srivastava and R. Tiwari, Evaluation of expectation of product of stochastic matrices, *Scand. J. Statist.* **3** (1976), 135–138.

T. D. Stanley, Stein-rule least squares estimation, a heuristic for falliable data, *Econ. Lett.* **20** (1986), 147–150.

T. D. Stanley, Forecasting from falliable data: Correcting prediction bias with Stein-rule least squares, *J. Forecasting* **7** (1988), 103–113.

T. D. Stanley, Galton, Stein and Empirical Bayes: Artifacts and connections, *Statistica* **XLIX** (1989), 499–511.

C. Stein, Estimation of mean of a multivariate normal distribution, *in* “Prague Symposium Asymptotic Statistics,” pp. 345–381, 1973.

T. Van Hoa, Improved estimators in some linear errors-in-variables models in finite samples, *Econ. Lett.* **20** (1986), 355–358.

H. D. Vinod and V. K. Srivastava, Large sample asymptotic properties of double *k*-class estimators in linear regression models, *Econometric Rev.* **14** (1995), 75–100.

A. S. Whittemore, Errors-in-variables regression using Stein estimation, *Amer. Statist.* **43** (1989), 226–228.

T. A. Yancy, G. G. Judge, and M. E. Bock, A mean squared error test when stochastic restrictions are used in regression, *Commun. Statist.* **3** (1974), 755–768.

Z. Zheng, On estimation of matrix of normal mean, *J. Multivariate Anal.* **18** (1986), 70–82.