1976

Double k-Class Estimators of Coefficients in Linear Regression

Aman Ullah

Shobha Ullah

20208

Not to be quoted.

RESEARCH REPORT 7617

DOUBLE k-CLASS ESTIMATORS OF COEFFICIENTS IN LINEAR REGRESSION

Aman Ullah
University of Western Ontario

Shobha Ullah*
Delhi School of Economics

September, 1976

*The first draft of this paper was written when the first author visited Australian National University. The authors are grateful to David F. Hendry and A. L. Nagar for many stimulating discussions on the subject matter of this paper. They are also thankful to R. A. L. Carter for useful suggestions which improved the exposition of this paper. They are, however, responsible for any error that may still remain.

DOUBLE k-CLASS ESTIMATORS OF COEFFICIENTS IN LINEAR REGRESSION

1. INTRODUCTION

It is well known that under certain assumptions the least squares estimators of the parameters of the general linear regression model are best unbiased in the family of linear unbiased estimators. If, however, we come out of the family of linear unbiased estimators then it is possible to obtain a family of biased estimators which is a nonlinear function of observations on the dependent variable and has a smaller mean squared error. In fact, Stein (1956), and James and Stein (1961) suggested a biased estimator for the orthonormal linear statistical model which dominates the least squares estimator in the sense that the sum of its component wise mean squared errors is smaller than that of the former, provided at least three parameters are to be estimated. Various interpretations and modifications of the above estimator, in the orthogonal regression context have recently appeared in the works of Baranchik (1964), Scolve, et al. (1972), Efron and Morris (1973) and Zellner and Vandaele (1975) among others. More recently, Bock (1975) extended the James and Stein estimator in the nonorthogonal regression context.

In this paper we consider the estimation of the parameters of the general linear regression model with the usual nonorthogonal regressors. We develop the h-class and double k-class families of biased estimators by using an operational variant of the minimum mean square error estimator [See Theil (1971, p. 125)] which depends on unknown parameters. The procedure of developing these families is simple and straightforward. We note that the James and Stein estimator in the regression context is a member of the double k-class family of estimators. We present the model and the families of h-class and double k-class estimators in the Section 2. The h in h-class and k_1 and k_2 in double k-class are taken as arbitrary scalars which could be stochastic or nonstochastic. For $k_2 = 1$ we obtain the Stein rules estimator as a member of the double k-class. In Section 3 we analyze the exact and approximate bias, moment matrix and the risk function of the double k-class estimators. For $0 \leq k_2 \leq 1$, the range of the values of k_1 for which the sum of the component wise mean squared errors of the double k-class estimators (to the order of approximation considered) dominate the least squares estimator is established in corollary 1 of Section 3.2. We also obtain the range of k_1 for which the double k-class estimator with $k_2 = 1$ dominates the estimators for $0 \leq k_2 < 1$. The double k-class estimators have no moments for $k_2 > 1$. Finally, in the Section 4 we give the proof of some of the theorems stated in Section 3. Some of the expectations and useful definitions required for the proof are presented in the Appendix.

2. THE MODEL AND ESTIMATORS

2.1 A Biased Estimator

Let us consider the regression model

(2.1) $$y = X\beta + u$$

where y is a Tx1 vector of observations on the dependent variable, X is a TxK matrix of known values with rank $K < T$, β is a Kx1 parameter vector and u is a Tx1 random vector such that

(2.2) $$Eu = 0 \quad \text{and} \quad Euu' = \sigma^2 I$$

The ordinary least squares (OLS) estimator of β in (2.1) and the residual variance estimator of σ^2 in (2.2) are written as

(2.3) $b = (X'X)^{-1}X'y$

$$s^2 = \frac{1}{n}\hat{u}'\hat{u} = \frac{1}{n}y'My$$

respectively, where

(2.4) $n = T - K,\ \hat{u} = My$ and

$$M = I - X(X'X)^{-1}X'.$$

M is a TxT idempotent matrix of rank T-K.

It is well known that the estimators in (2.3) are consistent and also

(2.5) $Eb = \beta\ ,\quad E(b-\beta)(b-\beta)' = \sigma^2(X'X)^{-1}$

$$Es^2 = \sigma^2.$$

Consider now a class of linear estimators

(2.6) $\beta^* = Ay$

where A is an arbitrary KxT matrix. The moment matrix of β^* can be written as

(2.7) $E(\beta^*-\beta)(\beta^*-\beta)' = \sigma^2AA' + (AX-I)\beta\beta'(AX-I)'.$

The matrix A for which (2.7) is a minimum is

(2.8) $A = \beta\beta'X'(X\beta\beta'X' + \sigma^2I)^{-1}.$

Thus, using (2.6) we obtain the minimum mean square error (MMSE) estimator

(2.9) $\beta^* = \beta\beta'X'(X\beta\beta'X' + \sigma^2I)^{-1}y$

as given by Theil (1971, p. 125). Further, β^* can be written as[1]

(2.10) $$\beta^* = \frac{(y-u)'y}{\sigma^2 + (y-u)'(y-u)}\beta$$

where

(2.11) $y-u = X\beta$

according to (2.1).

We note that the estimator β^* depends on unknown values on β and σ^2. Thus, we propose an operational variant of (2.10) as

$$\tilde{b} = \frac{(y-\hat{u})'y}{\frac{1}{n}\hat{u}'\hat{u} + (y-\hat{u})'(y-\hat{u})} b \tag{2.12}$$

where b, and $\hat{u}$ and n are as given in (2.3) and (2.4), respectively. The proposed estimator is a nonlinear function of y and it may not necessarily be MMSE estimator. We can write $\tilde{b}$ in an alternative form as

$$\tilde{b} = \left[1 - \frac{\hat{u}'\hat{u}/n}{y'y-\hat{u}'\hat{u}(1-\frac{1}{n})} \right] b \tag{2.13}$$

where we note that

$$0 \leq \frac{\hat{u}'\hat{u}/n}{y'y-\hat{u}'\hat{u}(1-\frac{1}{n})} \leq 1 \quad . \tag{2.14}$$

The estimator $\tilde{b}$ is consistent and its small sample properties are given in the following sections.

2.2 Families of h-Class and Double k-Class Estimators

While obtaining the operational variant of β^* in (2.10) as (2.12) we replace the estimate of $X\beta$ by $y-\hat{u}$ according to (2.11). If, instead, we replace the estimate of $X\beta$ by $y-h\hat{u}$ where h is an arbitrary scalar (stochastic or nonstochastic) we would obtain the operational variant of β^* in (2.9) as

$$\tilde{b}_h = \frac{(y-h\hat{u})'y}{h^2\hat{u}'\hat{u}/n + (y-h\hat{u})'(y-h\hat{u})} b \tag{2.15}$$

$$= \left[1 - \frac{\hat{u}'\hat{u}h^2/n}{y'y-\hat{u}'\hat{u}h^2(1-\frac{1}{n})} \right] b$$

The family of estimators in (2.15) differ with the estimator in (2.13) only with respect to the coefficients of $\hat{u}'\hat{u}$ on their right hand sides. It is clear that $h = 0$ yields the OLS estimator and $h = 1$ gives $\tilde{b}$.

A more natural generalization of $\bar{b}$ can be found in the following double k-class estimators :

$$\tilde{b}_{k_1,k_2} = \left(1 - \frac{k_1\hat{u}'\hat{u}}{y'y-k_2\hat{u}'\hat{u}}\right) b \tag{2.16}$$

where k_1, k_2 are arbitrary scalars which may be stochastic or nonstochastic. This family of estimators embraces the family of h-class as its members. We have shown in the following section that the moments of the estimator in (2.16) can be obtained for fixed k_2 when $0 \leq k_2 \leq 1$. It is interesting to note that for the value $k_2 = 1$ (2.16) is the Stein-rule estimators (1955, 1961). This can be written as

$$\hat{b}_{k_1,1} = (1 - k_1 \frac{\hat{u}'\hat{u}}{y'y - \hat{u}'\hat{u}})b \tag{2.17}$$

The value of k_1 for which the risk function of the estimator in (2.17) dominates the risk function of the OLS estimator is given by[2]

$$0 \leq k_1 \leq \frac{2(d - 2)}{n+2}, \quad d = \sum_{i=1}^{K} \lambda_i/\lambda_L, \tag{2.18}$$

where $\sum_{i=1}^{K} \lambda_i = tr(X'X)^{-1}$ and λ_L represents the largest characteristic root of $(X'X)^{-1}$. If $(X'X) = I$, then

$$0 \leq k_1 \leq \frac{2(K - 2)}{n+2} \tag{2.19}$$

and (2.17) is the James and Stein estimator for the orthogonal regression case as given in Scolve (1968).[3,4]

For the value $k_2 = 0$ the estimator in (2.16) can be written as

$$(2.20) \qquad \tilde{b}_{k_{1,0}} = (1 - \frac{k_1 \hat{u}'\hat{u}}{y'y})b$$

A family of estimators, which is a member of (2.16), can be written as

$$(2.21) \qquad \tilde{b}_k^+ = \left[1 - \frac{k\hat{u}'\hat{u}/n}{y'y - k\hat{u}'\hat{u}(1-\frac{1}{n})} \right]^+ b$$

where k is any scalar such that $0 \leq k \leq 1$. We note that the coefficient of b on the right hand of (2.21) is positive. Further, for $k = 1$ we obtain $\tilde{b}$ as given in (2.13).

3. THE BIAS AND MOMENT MATRIX OF $\tilde{b}_{k_1,k_2}$

In this Section we shall give exact and approximate formulae for the bias, moment matrix and the risk function of the double k-class estimators for $0 \leq k_2 \leq 1$.[5]

3.1 The Exact Results

Firstly, we write the sampling error of the estimator in (2.16) as

$$(3.1) \qquad (\tilde{b}_{k_1,k_2} - \beta) = (b - \beta) - k_1 cb$$

where

$$(3.2) \qquad c = \frac{y'My}{y'Ny} ,$$

and M and N are both TxT matrices given as

$$(3.3) \qquad M = I - X(X'X)^{-1}X' , \quad N = I - k_2 M$$

Further, M is an idempotent matrix with rank n and N is a non-negative definite matrix provided

(3.4) $$0 \le k_2 \le 1 \ .$$

Secondly, we make an assumption that the disturbance vector u in (2.1) is distributed as multivariate normal with mean vector zero and variance covariance matrix $\sigma^2 I$, i.e.,

(3.5) $$u \sim N(0,\sigma^2 I).$$

Thus

(3.6) $$y \sim N(\bar{y},\sigma^2 I)$$

where

(3.7) $$\bar{y} = X\beta \ .$$

Next we introduce the following notations and functions for the sake of simplicity of exposition :

(3.8) $$g_{\mu,\nu} = G(k_2,\theta;\frac{T}{2}+\mu,\ \frac{n}{2}+\nu) \ ; \quad \mu,\nu = 0,1,\ldots..$$

where

(3.9) $$\theta = \frac{\beta' X' X\beta}{2\sigma^2}$$

is a noncentrality parameter,

(3.10) $$0 \le k_2 \le 1$$

and the function G( ) is as defined in (A.5) of the Appendix A.

We can now state the following theorems :

<u>THEOREM 1</u>. *Under the assumption stated in* (3.5) *the exact bias of the double* k-*class estimator of* β *for* $0 \le k_2 \le 1$ *and* $n \ge 1$, *is given by*

(3.11) $$E(\tilde{b}_{k_1,k_2}-\beta) = -\frac{nk_1}{2}\, g_{2,1}\,\beta$$

where $g_{2,1}$ *is as given in* (3.8) *for* $\mu = 2$ *and* $\nu = 1$.

<u>THEOREM 2</u>. *Under the assumption stated in* (3.5) *the exact moment matrix of the double* k-*class estimator of* β *for* $0 \le k_2 \le 1$ *and* $n \ge 3$ *is given by*

$$(3.12)\qquad E(\tilde{b}_{k_1,k_2}-\beta)(\tilde{b}_{k_1,k_2}-\beta)' = \sigma^2(X'X)^{-1}[I-nk_1g_{2,1}-k_1^2\,\frac{n(n+2)}{4}\,(g_{3,2}-g_{2,2})] - \beta\beta'[nk_1(g_{3,1}-g_{2,1})+k_1^2\,\frac{n(n+2)}{4}\,(g_{4,2}-g_{3,2})].$$

THEOREM 3. *Under the assumption stated in* (3.5) *the exact risk function of the double* k*-class estimator of* β *for* $0 \le k_2 \le 1$ *and* $n \ge 3$ *is given by*

$$(3.13)\qquad E(\tilde{b}_{k_1,k_2}-\beta)'(\tilde{b}_{k_1,k_2}-\beta) = \sigma^2 tr(X'X)^{-1}[I-nk_1g_{2,1}-k_1^2\,\frac{n(n+2)}{4}\,(g_{3,2}-g_{2,2})] - \beta'\beta[nk_1(g_{3,1}-g_{2,1}) + k_1^2\,\frac{n(n+2)}{4}\,(g_{4,2}-g_{3,2})]$$

where 'tr' *represents the trace of the matrix and*

$$(3.14)\qquad tr(X'X)^{-1} = \sum_{i=1}^{k} \lambda_i ,$$

λ_i *is the i-th characteristic root of* $(X'X)^{-1}$.

The results corresponding to (3.11), (3.12) and (3.13) for the h-class and estimators in (2.15) and (2.21) can be obtained by substituting $k_1 = h^2/n$, $k_2 = h^2(1 - \frac{1}{n})$ and $k_1 = k/n$, $k_2 = k(1 - \frac{1}{n})$ respectively. The bias and moment matrix for $\tilde{b}_1$ in (2.12) is obtained by substituting $k = 1$. However, if

$$(3.15)\qquad k_2 = 1 \quad\text{or}\quad k_2 = 0$$

then we note from (A.7) and (A.8) of the Appendix that $g_{\mu,\nu}$ in (3.8) will be represented in terms of confluent hypergeometric functions. Thus, using the recurrence relations of the confluent hypergeometric functions given in Slater (1960, p.19) and using the notation

$$(3.16)\qquad f_{\delta,\eta} = e^{-\theta}\,\frac{\Gamma(\frac{K}{2}+\delta)}{\Gamma(\frac{K}{2}+\eta)}\;{}_1f_1\;(\frac{K}{2}+\delta;\,\frac{K}{2}+\eta;\theta)$$

the following corollary can be obtained.

Corollary 1. *The exact moment matrix of the James and Stein estimator in the nonorthogonal regression context, for* $k_2 = 1$ *and* $K \ge 3$, *is given by*[6]

$$(3.17)\qquad E(\tilde{b}_{k_1,1}-\beta)(\tilde{b}_{k_1,1}-\beta)' = \sigma^2(X'X)^{-1}[I - nk_1f_{0,1} + \frac{k_1^2\, n(n+2)}{4} f_{-1,1}] + \beta\beta'[nk_1 + \frac{k_1^2\, n(n+2)}{4}]f_{0,2}$$

and the risk function is given by

$$(3.18)\qquad E(\tilde{b}_{k_1,1}-\beta)'(\tilde{b}_{k_1,1}-\beta) = \sigma^2 tr(X'X)^{-1}[1 - nk_1f_{0,1} + \frac{k_1^2\, n(n+2)}{4} f_{-1,1}] + \beta'\beta[nk_1 + \frac{k_1^2 n(n+2)}{4}]f_{0,2}$$

where $f_{0,1}$ *is obtained from* (3.16) *for* $\delta = 0$ *and* $\eta = 1$, *and so on.*

It has been shown by Bock (1975) and Judge and Bock (1976) that for

$$(3.19)\qquad 0 \le k_1 \le \frac{2}{n+2}(d-2);\qquad d = \sum_{i=1}^{K} \lambda_i/\lambda_L \ge 2$$

$\tilde{b}_{k_1 1}$ dominates OLS estimator in the sense that

$$E(\tilde{b}_{k_1,1}-\beta)'(\tilde{b}_{k_1,1}-\beta) - E(b-\beta)'(b-\beta) < 0 ,$$

where, as given in (3.14),

$$\sum_{i=1}^{k} \lambda_i = tr(X'X)^{-1} .$$

For $X'X = I$, the results in (3.17) and (3.18) give the moment matrix and risk function of the James-Stein estimator in the orthogonal regression context. In this case these results are identical with the expressions obtained by Sclove (1968), Ullah (1970, 1974) and Ullahs (1976), among others.

Next, the corollary below can be easily stated.

<u>Corollary 2</u>. *The following results regarding the exact bias of the double k-class estimator in (3.11) are true for* $0 \leq k_2 \leq 1$ *and* $k_1 \geq 0$.

(a) $\tilde{b}_{k_1,k_2}$ *is unbiased only for* $k_1 = k_2 = 0$ *or* $k_1 = 0$ *and in that case it is OLS estimator.*

(b) *The exact relative bias of an element of* $\tilde{b}_{k_1,k_2}$ *is a decreasing function of both* k_1 *and* k_2.

(c) *The exact relative bias of an element of* $\tilde{b}_{k_1,k_2}$, *for a given sample size, lies in the following range:*

$$-\frac{nk_1}{T+2} \leq E\,\frac{(\tilde{b}_{k_1,k_2}-\beta)_i}{\beta_i} \leq 0\ , \quad i = 1,\ldots\ldots,K \tag{3.20}$$

(d) *The absolute value of the exact relative bias of an element of* $\tilde{b}_{k_1,k_2}$ *is a decreasing function of the noncentrality parameter* θ.

The results in the above corollary follows by noting that $g_{2,1} > 0$ and

$$\frac{\partial}{\partial k_2} g_{2,1} = \frac{\partial}{\partial k_2} G\left(k_2\,,\theta;\ \frac{T}{2}+2,\ \frac{n}{2}+1\right) > 0 \tag{3.21}$$

$$\frac{\partial}{\partial \theta} g_{2,1} = \frac{\partial}{\partial \theta} G\left(k_2\,,\theta;\ \frac{T}{2}+2,\ \frac{n}{2}+1\right) < 0 \tag{3.22}$$

according to (A.11) of the Appendix. Further, for given T

$$\underset{\theta\to\infty}{\text{lt}}\ E\,\frac{(\tilde{b}_{k_1,k_2}-\beta)_i}{\beta_i} = 0 \tag{3.23}$$

and

$$\underset{\theta\to 0}{\text{lt}}\ E\,\frac{(\tilde{b}_{k_1,k_2}-\beta)_i}{\beta_i} = -\frac{nk_1}{T+2} \tag{3.24}$$

by using (3.8) and (A.6) and (A.12) of the Appendix A. This gives the result (3.20).

3.2 Large - θ Asymptotic Expansion

We now present the asymptotic expansions of the bias and moment matrix of the double k-class estimators in terms of the inverse of θ.[7] These results help in analysing the complicated expressions of the exact moment matrix and the risk function given in (3.12) and (3.13), respectively.

The following three theorems can now be stated.

THEOREM 4. *The asymptotic expansion of the bias of the double* k-*class estimator of* β *in* (3.11) *up to order* $1/\theta^2$ *is given by*

$$E(\tilde{b}_{k_1,k_2}-\beta)=-\frac{nk_1}{2}\left[\frac{1}{\theta}+\frac{1}{2}\{(n+2)k_2-T\}\frac{1}{\theta^2}\right]\beta \tag{3.25}$$

where $n \geq 1$ *for* $0 \leq k_2 \leq 1$ *and* $K \geq 1$ *for* $k_2 = 1$.

THEOREM 5. *The asymptotic expansion of the moment matrix of the double* k-*class estimator of* β *in* (3.12) *upto order* $1/\theta^3$ *is given by*

$$\begin{aligned}E(\tilde{b}_{k_1,k_2}-\beta)(\tilde{b}_{k_1,k_2}-\beta)' &= \sigma^2(X'X)^{-1}+\frac{nk_1}{4\theta^2}[\beta\beta'\{4I+k_1(n+2)\}\\ &\quad -2(\beta'X'X\beta)(X'X)^{-1}]-\frac{nk_1}{4\theta^3}[\beta\beta'\{4(T-2k_2)\\ &\quad +k_1(n+2)(T+2-2k_2)\}+(X'X)^{-1}\beta'X'X\beta\\ &\quad \times\{(n+2)k_2-T-k_1\frac{(n+2)}{2}\}]\end{aligned} \tag{3.26}$$

where $n \geq 3$ *for* $0 \leq k_2 < 1$ *and* $K \geq 3$ *for* $k_2 = 1$.

THEOREM 6. *The asymptotic expansion of the risk function of the double* k-*class estimator of* β *in* (3.13) *upto* $1/\theta^3$ *is given by*

$$\begin{aligned}E(\tilde{b}_{k_1,k_2}-\beta)'(\tilde{b}_{k_1,k_2}-\beta) &= \sigma^2\,\mathrm{tr}\,(X'X)^{-1}+\frac{nk_1}{4\theta^2}[\beta'\beta\{4+k_1(n+2)\\ &\quad -2(\beta'X'X\beta)\mathrm{tr}(X'X)^{-1}]-\frac{nk_1}{4\theta^3}[\beta'\beta\{4(T-2k_2)\\ &\quad +k_1(n+2)(T+2-2k_2)\}+\\ &\quad +\beta'X'X\beta\;\mathrm{tr}(X'X)^{-1}\{(n+2)k_2-T-k_1\frac{(n+2)}{2}\}]\end{aligned} \tag{3.27}$$

where $n \geq 3$ *for* $0 \leq k_2 < 1$ *and* $K \geq 3$ *for* $k_2 = 1$.

Proof of Theorems 4 to 6: Using (3.8) and substituting (A.12) in (3.11), (3.12) and (3.13) the results in (3.25), (3.26) and (3.27) can be established easily.

We can now state the following corollaries.

Corollary 1. *The double* k-*class estimator of* β *in* (3.1) *dominates over the ordinary least squares estimator* b in (2.3) *in large-*θ *asymptotics upto the order* $1/\theta^2$, *in the sense that*

(3.28) $$\lim_{\theta\to\infty} \theta^2[E(\tilde{b}_{k_1,k_2}-\beta)'(\tilde{b}_{k_1,k_2}-\beta) - E(b-\beta)'(b-\beta)] < 0 ,$$

for

(3.29) $$d = \sum_{i=1}^{K} \lambda_i/\lambda_L \geq 2 \; ; \quad 0 < k_1 \leq \frac{2}{n+2}(d-2)$$

and for any k_2 *in*

(3.30) $$0 \leq k_2 \leq 1 ,$$

where $\sum_{i=1}^{K} \lambda_i = tr(X'X)^{-1}$, λ_i *is the i-th characteristic root of* $(X'X)^{-1}$ *and* λ_L *is the maximum of* λ_i, $i = 1,\ldots,K$.

The result in the above corollary follows by looking into the condition under which the coefficient of $\frac{1}{\theta^2}$ in the second term on the right hand side of (3.27) will be negative and also noting that k_2 is not involved up to order $\frac{1}{\theta^2}$.[8]

We note that the condition (3.29) is the same as the condition (3.19) for $k_2 = 1$ under which the exact risk function of $\tilde{b}_{k_1,1}$ (an extended form of James and Stein estimator) dominates over the ordinary least squares estimator.[9]

Corollary 2. *The double* k-*class estimator of* β *in* (3.1) *for* $k_2 = 1$ *dominates over the estimators for* $0 \leq k_2 < 1$ *in large-*θ *asymptotics upto the order* $\frac{1}{\theta^3}$, *in the sense that*

$$\text{(3.31)} \qquad \lim_{\theta\to\infty} \theta^3[E(\tilde{b}_{k_1,1}-\beta)'(\tilde{b}_{k_1,1}-\beta) - E(\tilde{b}_{k_1,k_2}-\beta)'(\tilde{b}_{k_1,k_2}-\beta)] < 0 ,$$

if

$$\text{(3.32)} \qquad d \geq 4 \text{ and } 0 < k_1 \leq \frac{1}{n+4}(d-4)$$

where **d** ***is as given in*** **(3.29).**

3.3 Conclusion.

It is interesting to note from (3.27) that the double k-class estimators, for given k_1, have the same risk functions (up to the order $1/\theta^2$) for any k_2 in $0 \leq k_2 \leq 1$. Further, we observe from (3.28) that for

$$0 < k_1 \leq \frac{2}{n+2}(d-2)$$

the risk function of the double k-class estimatros, up to the order $1/\theta^2$, dominates the risk function of the OLS estimator. Finally, the result in Corollary 2 indicates that if we consider the range of k_1 as

$$0 < k_1 \leq \frac{1}{n+4}(d-4), \quad d \geq 4$$

which is smaller than the range of k_1 in (3.29), then the double k-class estimator $\tilde{b}_{k_1,1}$ for $k_2 = 1$ will have smaller risk, up to order $1/\theta^3$, than the double k-class estimators from $0 \leq k_2 < 1$. If, however, $d \leq 4$ then for any positive k_1 estimators $\tilde{b}_{k_1,k_2}$ for $0 \leq k_2 < 1$ will dominate the estimator $\tilde{b}_{k_1,1}$. This is because for $d \leq 4$ and $k_1 > 0$ the inequality in (3.31) will be reversed.

4. PROOF OF THEOREMS 1 TO 3

In this section we shall give the proofs of theorems 1 to 3 stated in the section 3.1.

4.1 Proof of Theorem 1.

Let us take the expectation on both sides of (3.1) and write

$$\text{(4.1)} \qquad E(\tilde{b}_{k_1,k_2} - \beta) = - k_1 Ecb$$

where

$$c = \frac{y'My}{y'Ny}$$

as given in (3.2) and use has been made of (2.5). To obtain the expectation of cb on the right of (4.1) we write

(4.2) $$Ecb = \sigma(X'X)^{-1}X'Ezc$$

where

(4.3) $$z = \frac{y}{\sigma} \sim N(\bar{z}, I) \ , \quad \bar{z} = \frac{X\beta}{\sigma}$$

by using (3.6) and (3.7), and

(4.4) $$c = \frac{y'My}{y'Ny} = \frac{z'Mz}{z'Nz} \ .$$

Now we note that[10]

(4.5) $$\begin{aligned} Ezc &= E(z-\bar{z})c + \bar{z}Ec \\ &= \frac{1}{(2\pi)^{T/2}} \int_z (z-\bar{z})c \exp - \frac{1}{2} \{(z-\bar{z})'(z-\bar{z})\}dz + \bar{z}Ec \\ &= \frac{\partial}{\partial \bar{z}} Ec + \bar{z}Ec \end{aligned}$$

where, for $n \geq 1$ and $0 \leq k_2 \leq 1$,

(4.6) $$Ec = \frac{n}{2} G(k_2, \theta; \frac{T}{2} + 1, \frac{n}{2} + 1) \ ,$$

has been obtained from (B.15) of the Appendix B by substituting $k = k_2$ and x by

(4.7) $$\theta = \frac{\bar{z}'\bar{z}}{2}$$

Next, noting the fact that

(4.8) $$\frac{\partial}{\partial \bar{z}} Ec = \left(\frac{\partial}{\partial \theta} Ec \right) \frac{\partial \theta}{\partial \bar{z}}$$

and using (A.9) of the Appendix, (4.5) can be obtained as

(4.9) $$Ezc = \frac{n}{2} \bar{z} \ G(k_2, \theta; \frac{T}{2} + 2, \frac{n}{2} + 1) \ .$$

Further, substituting (4.9) in (4.2) we get

(4.10) $$Ecb = \frac{n}{2} \beta \ G(k_2, \theta; \frac{T}{2} + 2, \frac{n}{2} + 1) \ .$$

Finally, using (4.10) in (4.1) we obtain the result stated in the Theorem 1.

Proof of Theorems 2 and 3.

Using (3.1) we write the moment matrix of the double k-class estimators as

$$E(\tilde{b}_{k_1,k_2}-\beta)(\tilde{b}_{k_1,k_2}-\beta)' = E(b-\beta)(b-\beta)' + k_1(E\beta b'c + Ecb\beta') - 2k_1 Ecbb' + k_1^2\, Ec^2bb' \quad (4.11)$$

The first term on the right hand of (4.11) is $\sigma^2(X'X)^{-1}$ as given in (2.5). Next, considering the second term we note from (4.10) that

$$\begin{aligned} E\beta b'c &= \frac{n}{2}\,\beta\beta' G(k_2,\theta;\ \frac{T}{2}+2,\ \frac{n}{2}+1) \\ &= Ecb\beta' . \end{aligned} \quad (4.12)$$

Now, taking the third term on the right hand of (4.11) we write

$$Ecbb' = \sigma^2(X'X)^{-1}X'(Ezz'c)\,X\,(X'X)^{-1} \quad (4.13)$$

where z and c are as defined in (4.3) and (4.4), respectively. Using the procedure in (4.5), we note that

$$\begin{aligned} Ezz'c &= E\{(z-\bar{z})(z-\bar{z})' + (z-\bar{z})\bar{z}' + \bar{z}(z-\bar{z})' + \bar{z}\bar{z}'\}c \\ &= \frac{\partial^2}{\partial\bar{z}\partial\bar{z}'}Ec + 2\,\bar{z}\,\frac{\partial}{\partial\bar{z}'}Ec + (\bar{z}\bar{z}' + I)Ec \\ &= \bar{z}\bar{z}'\,\frac{\partial^2}{\partial\theta^2}Ec + (2\bar{z}\bar{z}' + I)\,\frac{\partial}{\partial\theta}Ec + (\bar{z}\bar{z}' + I)Ec \end{aligned} \quad (4.14)$$

where Ec is as given in (4.6). Further, using (A.9) and (A.10) of the Appendix A, (4.14) can be simplified as

$$Ezz'c = \frac{n}{2}\,[\bar{z}\bar{z}'G(k_2,\theta;\ \frac{T}{2}+3,\ \frac{n}{2}+1) + G(k_2,\theta;\ \frac{T}{2}+2,\ \frac{n}{2}+1)] \quad (4.15)$$

and (4.13) can be written as

$$Ecbb' = \frac{n}{2}\,[\beta\beta' G(k_2,\theta;\ \frac{T}{2}+3,\ \frac{n}{2}+1) + \sigma^2(X'X)^{-1} \times G(k_2,\theta;\ \frac{T}{2}+2,\ \frac{n}{2}+1)]. \quad (4.16)$$

Similarly, considering the fourth term on the right hand side of (4.11) we first note that

$$(4.17) \qquad Ezz'c^2 = \bar{z}\bar{z}' \frac{\partial^2}{\partial\theta^2} Ec^2 + (2\bar{z}\bar{z}' + I) \frac{\partial}{\partial\theta} Ec^2 + (\bar{z}\bar{z}' + I)Ec^2$$

where, for $n \geq 3$,

$$(4.18) \qquad Ec^2 = \frac{n(n+2)}{4} \left[G(k_2,\theta; \tfrac{T}{2} + 1, \tfrac{n}{2} + 2) - G(k_2,\theta; \tfrac{T}{2} + 2, \tfrac{n}{2} + 2)\right]$$

has been obtained from (B.16) of the Appendix B by replacing k by k_2 and x by θ. Now using the partial derivatives of G given in (A.9) and in (A.10) of the Appendix A we can obtain (4.17) and hence Ec^2bb' as

$$(4.19) \qquad Ec^2bb' = \sigma^2(X'X)^{-1}X'(Ezz'c^2)X(X'X)^{-1} = -\frac{n(n+2)}{4} \Big[\beta\beta'\{G(k_2,\theta; \tfrac{T}{2} + 4, \tfrac{n}{2} + 2)$$

$$- G(k_2,\theta; \tfrac{T}{2} + 3, \tfrac{n}{2} + 2)\} + \sigma^2(X'X)^{-1}\{G(k_2,\theta; \tfrac{T}{2} + 3, \tfrac{n}{2} + 2)$$

$$- G(k_2,\theta; \tfrac{T}{2} + 2, \tfrac{n}{2} + 2)\}\Big]$$

Finally, substituting (2.5), (4.12), (4.16) and (4.19) in (4.11) the result stated in Theorem 2 follows. The Theorem 3 follows by taking the trace on both sides of (3.12).

APPENDIX

A. Hypergeometric and G Functions

The hypergeometric functions ${}_1F_1$ and ${}_2F_1$ have the following power series representations [Slater (1960)],

$$(A.1)\qquad {}_1F_1(a;c;x) = \frac{\Gamma c}{\Gamma a} \sum_{m=0}^{\infty} \frac{\Gamma(a+m)}{\Gamma(c+m)} \frac{x^m}{m!}, \quad c > 0, \quad |x| < \infty$$

$$(A.2)\qquad {}_2F_1(a,b;c;x) = \frac{\Gamma c}{\Gamma a \Gamma b} \sum_{m=0}^{\infty} \frac{\Gamma(a+m)\Gamma(b+m)}{\Gamma(c+m)} \frac{x^m}{m!}, \quad c > 0, \quad |x| < 1$$

Similarly, the power series representation of the function G,

$$(A.3)\qquad G(k,x;a,c) = \int_{-\infty}^{0} h(t;k,x,a,c)dt$$

where $x > 0$, $a \geq c \geq 1$, $0 \leq k \leq 1$ and

$$(A.4)\qquad h(t;k,x,a,c) = \frac{2 \exp\,[2xt/1-2t]}{[1-2t]^{a-c}[1-2(1-k)t]^c},$$

is given by [see Sawa (1972, p. 659)]

$$(A.5)\qquad G = e^{-x} \frac{\Gamma(a-1)}{\Gamma(c)} \sum_{h=0}^{\infty} (k)^h \frac{\Gamma(c+h)}{\Gamma(a+h)}\, {}_1F_1(a-1;a+h,x), \quad a > 1.$$

An alternative representation is

$$(A.6)\qquad G = e^{-x} \sum_{h=0}^{\infty} \frac{x^h}{h!} \frac{\Gamma(a-1+h)}{\Gamma(a+h)}\, {}_2F_1(1,c;a+h;k), \quad a > 1.$$

For k=1, (A.5) and (A.6) reduce to

$$(A.7)\qquad G = e^{-x} \frac{\Gamma(a-c-1)}{\Gamma(a-c)}\, {}_1F_1(a-c-1;a-c;x), \quad a-c > 1.$$

Further, for k=0

$$G = e^{-x} \frac{\Gamma(a-1)}{\Gamma(a)} {}_1F_1(a-1;a;x). \tag{A.8}$$

The partial derivatives of G with respect to x can be written as

$$\frac{\partial}{\partial x} G(k,x;a,c) = G(k,x;a+1,c) - G(k,x;a,c) \tag{A.9}$$

$$\frac{\partial^2}{\partial x^2} G(k,x;a,c) = G(k,x;a+2,c) - 2G(k,x;a+1,c) + G(k,x;a,c) \tag{A.10}$$

and so on. Also we note from (A.3) and (A.5) that

$$\frac{\partial}{\partial k} G(k,x;a+1,c) > 0 \quad \text{and} \quad \frac{\partial}{\partial x} G(k,x;a,c) < 0. \tag{A.11}$$

Finally, for large x, the asymptotion of the function G in (A.5), up to order $\frac{1}{x^3}$, is given by [see Sawa (1972, p. 667)].

$$G(k,x;a,c) = \frac{1}{x} + (c\ k-a+2)\frac{1}{x^2} + [c(c+1)k^2 - 2c(a-2)k + (a-2)(a-3)]\frac{1}{x^3}\ . \tag{A.12}$$

B. Evaluation of Some Expectations Required in Section 4.

Let z by a Tx1 normally distributed random vector such that

$$Ez = \bar{z} \quad \text{and} \quad E(z-\bar{z})(z-\bar{z})' = I_T. \tag{B.1}$$

Further, consider M as a TxT idempotent matrix with rank $n < T$. Then, it can be verified that the matrix

$$N = I - kM \tag{B.2}$$

will be non-negative definite for $0 \leq k \leq 1$.

The joint moment generating function $M(t_1,t_2)$ of the quadratic forms z'Nz and z'Mz can be written as

$$\begin{aligned} M(t_1,t_2) &= E \exp\,[t_1 z'Nz + t_2 z'Mz] \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\,[t_1 z'Nz + t_2 z'Mz]\, f(z)dz \end{aligned} \tag{B.3}$$

where f(z) represents the multivariate normal density of z with mean vector $\bar{z}$ and covariance matrix I.

Since M is an idempotent matrix of rank $n < T$ we can always obtain an orthogonal matrix P such that the orthogonal transformation of the matrix

$$Q = I - 2t_1 N - 2t_2 M \tag{B.4}$$

can be written as

$$P'QP = \begin{bmatrix} I_n - (2t_1 k^* + 2t_2) I_n & 0 \\ 0 & (1-2t_1) I_{T-n} \end{bmatrix} = \Lambda \tag{B.5}$$

where $k^* = 1-k$ and Λ is a TxT diagonal matrix. Thus, if we restrict the domains of t_1, t_2 and k as

(B.6) $\quad 2t_1 < 1,\ 2t_1 + 2t_2 < 1$ and $0 \le k \le 1$

it follows that Q is a non-negative definite matrix.

We can now simplify (B.3) as

(B.7) $$M(t_1,t_2) = \exp \frac{1}{2}\ [\bar{z}'Q^{-1}\bar{z}-\bar{z}'\bar{z}]\ |Q^{-1}|^{\frac{1}{2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(z)dz$$

where g(z) is multivariate normal density of z with mean vector $Q^{-1}\bar{z}$ and variance covariance matrix Q^{-1}. Finally, using (B.5) and noting that the integral value on the right of (B.7) is unity we obtain

(B.8) $$M(t_1,t_2) = \exp \frac{1}{2}\ [\bar{z}'Q^{-1}\bar{z}-\bar{z}'\bar{z}] \Big/ (1-2t_1)^{\frac{T-n}{2}}\ (1-2t_1k^*-2t_2)^{n/2}$$

In the case where the matrix M is such that

(B.9) $\quad M\bar{z} = 0$

we note

(B.10) $\quad Q\bar{z} = (1-2t_1)\bar{z}.$

In this case (B.8) is simplified as

(B.11) $$M(t_1,t_2) = \exp \left[\frac{t_1\bar{z}'\bar{z}}{1-2t_1}\right] \Big/ (1-2t_1)^{\frac{T-n}{2}}\ (1-2t_1k^*-2t_2)^{n/2}$$

where $k^* = 1-k$. The following derivatives of (B.11) can then be easily verified.

(B.12) $$\left.\frac{\partial M(t_1,t_2)}{\partial t_2}\right|_{t_2=0} = \frac{n}{2}\, h(t_1;k,x,\ \frac{T}{2}+1,\ \frac{n}{2}+1)$$

(B.13) $$\left.\frac{\partial^2 M(t_1,t_2)}{\partial t_2^2}\right|_{t_2=0} = \frac{n(n+2)}{2}\, h(t_1;k,x,\ \frac{T}{2}+2,\ \frac{n}{2}+2)$$

where

(B.14) $$x = \frac{\bar{z}'\bar{z}}{2}$$

and h( ) is as defined in (A.4).

Finally, using (B.12), (B.13) and (A.3) we obtain the following expectations for $0 \leq k \leq 1$:[11]

(B.15) $$E\left(\frac{z'Mz}{z'Nz}\right) = \int_{-\infty}^{0} \left[\frac{\partial M(t_1,t_2)}{\partial t_2}\right]_{t_2=0} dt_1$$

$$= \frac{n}{2}\, G(k,x;\frac{T}{2}+1,\ \frac{n}{2}+1)\ , \qquad T \geq n \geq 1$$

(B.16) $$E\left(\frac{z'Mz}{z'Nz}\right)^2 = \int_{-\infty}^{0} - t_1 \left[\frac{\partial^2 M(t_1,t_2)}{\partial t_2^2}\right]_{t_2=0} dt_1$$

$$= \frac{n(n+2)}{4}\ [G(k,x;\ \frac{T}{2}+1,\ \frac{n}{2}+2) -$$

$$- G(k,x;\ \frac{T}{2}+2,\ \frac{n}{2}+2)]\ , \qquad T \geq n \geq 3$$

where use has been made of

(B.17) $$2t_1 h(t_1;k,x,\ \frac{T}{2}+2,\ \frac{n}{2}+2) = \frac{\partial}{\partial x}\, h(t_1;k,x,\ \frac{T}{2}+1,\ \frac{n}{2}+2)$$

and (A.9).

FOOTNOTES

[1]In obtaining (2.10) from (2.9) we note that

$$(X\beta\beta'X' + \sigma^2 I)^{-1} = \frac{1}{\sigma^2}\,[I - X\beta(\sigma^2 + \beta'X'X\beta)^{-1}\beta'X']$$

[2]Also, see Bock (1975) and Judge and Bock (1976).

[3]For the variants of Stein-rule estimators, see, for example, Baranchik (1964), Scolve et al. (1972), Bock (1973), Efron and Morris (1973), Zellner and Vandaele (1975).

[4]If we take Q = X'X in the weighted loss function considered by Zellner-Vandaele, then for $k_1 = K/n+2$ we obtain Zellner-Vandaele (1975) estimator.

[5]In general, k_1 and k_2 could be stochastic. However, in this paper, we analyze the results only for the fixed k_1 and k_2. It can be easily shown that for $k_2 > 1$ the moments do not exist.

[6]We note that $K \geq 3$ implies $n = T-K \geq 3$.

[7]The results are valid for sufficiently large θ, which according to (3.9) means sufficiently small σ. The behavior of estimators, when the disturbance term is small, has been analyzed by Kadane (1970, 1971). Also, for large θ asymptotic expansion, see the workds by Sawa (1972), Basmann (1963) and Mariano (1973), among others.

[8]We also note by using a result in Rao (1965, p. 59) that the minimum of

$$\frac{\beta'X'X\beta}{\beta'\beta} = 1/\lambda_L \text{ where } \lambda_L \text{ is the maximum characteristic root of } (X'X)^{-1}.$$

[9]It would be more interesting to obtain k_1 and k_2 for which the difference between exact risk functions, viz.,

$$E(\tilde{b}_{k_1,k_2}-\beta)'(\tilde{b}_{k_1,k_2}-\beta) - E(b-\beta)'(b-\beta) < 0$$

In this paper, however, we have considered the above difference up to order $1/\theta^2$. Also see the work by Sawa (1972, p. 669) in this respect.

[10]A similar technique of obtaining expectations, in a different context, was originated in the work by Ullah and Nagar (1974).

[11]This technique of obtaining expectations was originated by Williams (1941) and later used by Sawa (1972) in different contexts.

REFERENCES

Baranchik, A. J., 1964, "Multiple Regression and the Estimation of the Mean of a Multivariate Normal Distribution," Technical Report No. 51, Department of Statistics, Stanford University.

Basmann, R. L., 1963, "Remarks Concerning the Application of Exact Finite Sample Distribution Functions of GCL Estimators in Econometric Statistical Inference," Journal of the American Statistical Association, Vol. 58, pp. 943-76.

Bock, M. E., 1975, "Minimax Estimators of the Mean of a Multivariate Distribution," The Annals of Statistics, Vol. 3, pp. 209-18.

__________, 1972, "A Comparison of the Risk Functions for Preliminary Test and Positive Part Estimators," Research Paper, University of Illinois--Urbana.

Efron, B., and C. Morris, 1973, "Stein's Estimation Rule and Its Competitors--An Empirical Bayes Approach," Journal of the American Statistical Association, Vol. 68, pp. 117-30.

James, W., and C. Stein, 1961, "Estimation with Quadratic Loss," Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, pp. 361-379.

Judge, G. G. and M. E. Bock, 1976, "A Comparison of Traditional and Stein-Rule Estimators under Weighted Squared Error Loss," International Economic Review, Vol. 17, No. 1.

Kadane, J. B., 1970, "Testing Overidentifying Restrictions when the Disturbances are Small," Journal of the American Statistical Association, Vol. 65, pp. 182-85.

__________, 1971, "Comparison of k-class Estimators when the Disturbances are Small," Econometrica, Vol. 39, pp. 723-738.

Mariano, R. S., 1973, "Approximations to the Distribution Functions of Theil's k-class Estimators," Econometrica, Vol. 41, pp.

Rao, C. R., 1965, Linear Statistical Inference and Its Applications, John Wiley and Sons, Inc.

Sawa, T., 1972, "Finite-Sample Properties of the k-class Estimators," Econometrica, Vol. 40, pp. 653-80.

Sclove, S. L., 1968, "Improved Estimators for Coefficients in Linear Regression," Journal of the American Statistical Association, Vol. 63, pp. 599-606.

Sclove, S. L., C. Morris, and R. Radhakrishnan, 1972, "Non Optimality of Preliminary-Test Estimators for the Multinormal Mean," Annals of Mathematical Statistics, Vol. 43, pp. 1481-90.

Stein, C., 1956, "Inadmissibility of the Usual Estimator for the Man of a Multivariate Normal Distribution," Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1, pp. 197-206.

Theil, H., 1971, Principles of Econometrics (North-Holland, Amsterdam).

Ullah, A., 1970, "Statistical Estimation of Economic Relations in the Presence of Errors in Equations and in Variables," Ph.D. Dissertation (Economics Department, Delhi University, Delhi).

________, 1974, "On the Sampling Distribution of Improvied Estimators for Coefficients in Linear Regression," Journal of Econometrics, Vol. 2, pp. 143-50.

Ullah, A. and A. L. Nagar, 1974, "The Exact Mean of the Two Stage Least Squares Estimator of the Structural Parameters in an Equation having three Endogenous Variables," Econometrica.

Ullah, A. and S. Ullah, 1976, "The Estimation of Regression Coefficients and the Residual Variance in the Regression Model Using the Stein Estimator," Working Paper, University of Western Ontario.

Williams, J. D., 1941, "Moments of the Ratio of the Mean Square Successive to the Mean Square Difference in Samples from a Normal Universe," _Annals of Mathematical Statistics_, Vol. 12, pp. 239-241.

Zellner, A., and W. Vandaele, 1975, "Bayes-Stein Estimators for k-Means, Regression, and Simultaneous Equations Models," in _Studies in Bayesian Econometrics and Statistics_ (S. E. Fienberg and A. Zellner, eds.), North-Holland Publishing Co., Amsterdam, pp. 627-53.