1986

Unbiased Estimation of the MSE Matrix of Stein-Rule Estimators, Confidence Ellipsoid and Hypothesis Testing

Aman Ullah

V.K. Srivastava

R.A.L. Carter

M.S. Srivastava

THE CENTRE FOR DECISION SCIENCES AND ECONOMETRICS

Unbiased Estimation Of The MSE Matrix of Stein-Rule Estimators, Confidence Ellipsoid and Hypothesis Testing

A. Ullah, V.K. Srivastava, R.A.L. Carter and M.S. Srivastava

TECHNICAL REPORT NO. 12
JULY 1986

Centre For Decision Sciences And Econometrics
Social Science Centre
The University of Western Ontario
London, Ontario N6A 5C2

2812

ISSN 0831-4454
ISBN 0-7714-0760-2

UNBIASED ESTIMATION OF THE MSE MATRIX OF STEIN-RULE ESTIMATORS, CONFIDENCE ELLIPSOID AND HYPOTHESIS TESTING

by

A. Ullah, V.K. Srivastava, R.A.L. Carter and M. S. Srivastava*

University of Western Ontario, Lucknow University,
University of Western Ontario and University of Toronto

1. INTRODUCTION

During the recent past, several families of improved estimators for the coefficient vector in a linear regression model have been proposed and their small-sample properties have been analyzed. Among them, an interesting family is that of the Stein-rule estimators, the properties of which have been extensively studied, see e.g., Judge and Bock (1978) and Vinod and Ullah (1981). However, none of the work so far, has dealt with the estimation of the bias vector and the mean squared error (MSE) matrix which is of paramount importance to users, for it is widely recognized that merely a point estimate without some measure of variability is inadequate. This is what has prompted this article in which, in Section 2, we have derived the unbiased estimators of the exact bias vector and the exact MSE matrix of Stein-rule estimators. We then consider, in Section 3, the use of estimated MSE matrices in the construction of confidence ellipsoids. Finally, in Section 4, we consider the problem of testing linear restrictions using an F-ratio based on the Stein-rule estimator.

2. UNBIASED ESTIMATION OF BIAS AND MSE

Consider a linear regression model:

$$y = X\beta + u \tag{2.1}$$

where y is a TX1 vector of T observations on the variable to be explained, X is a TXp full-column-rank matrix of T observations on p explanatory variables, β is a column vector of regression coefficients and u is a TX1 vector of disturbances assumed to follow a multivariate normal distribution with mean vector 0 and variance covariance matrix $\sigma^2 I_T$, σ^2 being unknown.

The least squares (LS) estimator of β is given by

$$(2.2) \qquad b = (X'X)^{-1}X'y$$

which is unbiased with variance covariance matrix $\sigma^2(X'X)^{-1}$. An unbiased estimate of this variance covariance matrix is given by

$$(2.3) \qquad \hat{V}(b) = s^2(X'X)^{-1}$$

where $s^2 = (y-Xb)'(y-Xb)/n$ and $n = (T-p)$.

The Stein-rule (SR) estimator of β is given by

$$(2.4) \qquad \hat{\beta} = [1- k \frac{(y-Xb)'(y-Xb)}{b'X'Xb}] b$$

where k is the scalar characterizing the estimator.

The estimator $\hat{\beta}$ is biased. Further, if $M(\hat{\beta})$ denotes the mean squared error matrix of $\hat{\beta}$, V(b) is the variance-covariance matrix of b, and k is nonstochastic and positive, the Stein-rule estimator dominates the least squares estimator in the sense that

$$(2.5) \quad tr[X'X\, M(\hat{\beta})] < tr[X'X\, V(b)] \text{ or } E[(\hat{\beta}-\beta)'X'X(\hat{\beta}-\beta)] < E[(b-\beta)'X'X(b-\beta)]$$

when

$$(2.6) \quad 0 < k < (\frac{2(p-2)}{n+2}); \; p \geq 3.$$

The exact expressions for the bias vector and the mean squared error matrix of $\hat{\beta}$ are quite involved (see Judge and Bock (1978) and Ullah and Ullah (1978)) and as such do not provide any clues as to their estimation. Thus, these expressions are of little use to practitioners since they are more concerned with how to estimate the exact root-mean-squared-errors (RMSE) associated with the Stein-rule estimates of the coefficients in any given application. With this in view, we present the unbiased estimators of the exact bias vector and the exact mean squared error matrix of $\hat{\beta}$ assuming k to be nonstochastic and positive.

An unbiased estimator of the bias vector of $\hat{\beta}$ is

$$(2.7)\qquad \hat{B}(\hat{\beta}) = -k \frac{ns^2}{b'X'Xb} b$$

while an unbiased estimator of the mean squared error matrix of $\hat{\beta}$ is

$$(2.8)\qquad \hat{M}(\hat{\beta}) = s^2[1 - 2k \frac{n^2 s^2}{(n+2)b'X'Xb}] (X'X)^{-1} + k(k + \frac{4}{n+2}) (\frac{ns^2}{b'X'Xb})^2 bb'.$$

The results (2.7) and (2.8) are derived in Section 2.2.

The result in (2.8) is simple, and useful for calculating the MSE of an individual component of $\hat{\beta}$. We note here that Stein (1981) and Manjoge (1984) provide unbiased estimators of the tr $M(\hat{\beta})/\sigma^2$ in the multivariate normal mean and regression models, respectively. Their results, however, do not provide the unbiased MSE of an individual coefficient which is usually of interest in econometrics.

2.1 A Numerical Illustration

To illustrate the empirical usefulness of these results we employed them in the context of the example provided by Aigner and Judge (1977). The equation considered is their "full model" under the assumption that the nine regressors and y are all measured as deviations from sample means. This eliminates the constant, so that p=9, but leaves the slope coefficients and LS residuals unchanged. Then the F statistic can easily be transformed to $ns^2/b'X'Xb$. Of course, the degrees of freedom must still account for the estimated mean of y and so $n=T-p-1=48$.

Aigner and Judge used two modified SR coefficient estimators and approximated their standard errors by the LS standard errors. We used the SR estimator (2.4), with $k=(p-2)/(n+2)=.14$, and estimated the bias and MSE of these point estimates by (2.7) and (2.8). The estimated RMSE for a coefficient is the square root of the appropriate main diagonal element from $M(\hat{\beta})$. The results are displayed in Table 1. (A complete description of the variables can be found in Aigner and Judge (1977.)

Since the SR dominates LS only with respect to the criterion (2.5) and not necessarily with respect to the MSE of individual coefficients, we would expect some of the figures in the last column of Table 1 to be larger than some of those in the second column. This is, indeed, the case. What is interesting is that when the estimated RMSE($\hat{\beta}$) exceeds the SE(b), as in the first and last two rows, the difference is rather small but when the SE(b) exceeds the estimated RMSE($\hat{\beta}$) the difference is, proportionately, quite large; see rows three, five, six and seven. It seems, therefore, that the estimated RMSE are superior to the SE(b) as measures of the precision of the $\hat{\beta}$ elements.

2.2. Derivation of the Results:

In order to obtain the unbiased estimators of the bias vector and mean squared error matrix of $\hat{\beta}$, we observe from (2.4) that

$$(2.9) \qquad (\hat{\beta}-\beta) = (b-\beta) - k \frac{ns^2}{b'X'Xb} b.$$

Thus the bias vector is

$$(2.10) \qquad B(\hat{\beta}) = -\, k\, E[\frac{ns^2}{b'X'Xb}\, b]$$

whence it follows that an unbiased estimator of $B(\hat{\beta})$ is

$$(2.11) \qquad \hat{B}(\hat{\beta}) = -\, k\, \frac{ns^2}{b'X'Xb}\, b$$

which is the result (2.7)

Next, the mean-squared-error matrix of $\hat{\beta}$ is given by

$$(2.12) \qquad M(\hat{\beta}) = E(\hat{\beta}-\beta)(\hat{\beta}-\beta)'$$

$$= E(b-\beta)(b-\beta)' + k^2\, E[(\frac{ns^2}{b'X'Xb})^2\, bb']$$

$$-\, k\, E[(\frac{ns^2}{b'X'Xb})\, \{b(b-\beta)' + (b-\beta)b'\}]$$

$$= \sigma^2 (X'X)^{-1} + k^2\, E[(\frac{ns^2}{b'X'Xb})^2\, bb']$$

$$-\, k\, E[(\frac{ns^2}{b'X'Xb})\, \{b(b-\beta)' + (b-\beta)b'\}].$$

It is easily seen that unbiased estimators of the first two terms on the right hand side are

$$(2.13) \qquad s^2 (X'X)^{-1} + k^2 (\frac{ns^2}{b'X'Xb})^2\, bb' \;.$$

Now let us introduce the following notation

$$(2.14) \qquad z = (X'X)^{\frac{1}{2}} b, \quad \theta = (X'X)^{\frac{1}{2}} \beta,$$

$$w = ns^2$$

so that z has a multivariate normal distribution with mean vector θ and variance covariance matrix $\sigma^2 I_p$ while (w/σ^2) has a χ^2-distribution with n degrees of freedom. Further, z and w are independently distributed.

Next, consider the quantity

$$(2.15)\quad (X'X)^{\frac{1}{2}} E[\frac{ns^2}{b'X'Xb} b(b-\beta)'] (X'X)^{\frac{1}{2}} = E[\frac{w}{z'z} z(z-\theta)']$$

$$= \sigma^2 E[\frac{\partial}{\partial z'} (\frac{w}{z'z} z)]$$

$$= \sigma^2 E[\frac{w}{z'z} (I_p - 2P_z)],$$

$$= \frac{1}{n+2} E[\frac{w^2}{z'z} (I_p - 2P_z)]$$

where $P_z = z(z'z)^{-1} z'$. Then from the last equality it follows that

an unbiased estimator of the left hand side of the matrix expression (2.15) is

$$(2.16)\quad \frac{w^2}{(n+2)\, z'z} (I_p - 2P_z)$$

$$= \frac{n^2 s^4}{(n+2)\, b'X'Xb} [I_p - \frac{2}{b'X'Xb} (X'X)^{\frac{1}{2}} bb'(X'X)^{\frac{1}{2}}].$$

Finally, using the unbiased estimators (2.13) and (2.16) in the last equality of (2.12), we get the result in (2.8).

3. CONFIDENCE ELLIPSOIDS

Following Stein (1981), the estimated MSE matrix, (2.8), can be used in a quadratic form which defines a confidence ellipsoid for the vector β;

(3.1) $Q_s = (\hat{\beta}-\beta)'[\hat{M}(\hat{\beta})]^{-1}(\hat{\beta}-\beta).$

Alternatively, the asymptotic value of $\hat{M}(\hat{\beta})$, as n grows large, can be used to give

(3.2) $Q_a = (\hat{\beta}-\beta)'[s^2(X'X)^{-1}]^{-1}(\hat{\beta}-\beta).$

A third alternative, which provides a standard against which to compare the first two, is

(3.3) $Q_{\ell s} = (b-\beta)'[s^2(X'X)^{-1}]^{-1}(b-\beta).$

We could decide which of these quadratic forms is the most desirable as a confidence ellipsoid by choosing a confidence level, $1-\alpha$, and finding the values of c_s, c_a and $c_{\ell s}$ for which

(3.4) $P(Q_s \leq c_s) = P(Q_a \leq c_a) = P(Q_{\ell s} \leq c_{\ell s}) = 1-\alpha.$

Versions of Q with low c values are preferred to those whith high c values. Alternatively, for a given c, one can obtain $P(Q_s \leq c) = 1-\alpha_s$, $P(Q_a \leq c) = 1-\alpha_a$ and $P(Q_{\ell s} \leq c) = 1-\alpha_{\ell s}$ and prefer the version of Q for which $1-\alpha$ is a maximum.

The sampling distribution of Q_a has been analysed by Ullah et al (1983, 1984). From their corollary 4 the small-σ asymptotic expansion, up to order $O(\sigma^2)$, and the large-T expansion, up to order $O(T^{-1})$, of the distribution of Q_a, are respectively given by

(3.5) $P[Q_a \leq c] = G(pF_{p,n}) + \sum_{i=-1}^{1} \sum_{j=1}^{4} \eta_{i,j} G\left(\frac{n(p+2i)}{n+2j} F_{p+2i,n+2j}\right)$

and, assuming $k=O(T^{-1})$,

(3.6) $P[Q_a \leq c] = G(\chi^2_p) + \sum_{i=-1}^{1} \mu_i G(\chi^2_{p+2i})$

where $G((nr/w)F_{r,w}) = P[(nr/w)F_{r,w} \leq c]$, for any r and w, represents the distribution of central F with r and w degrees of freedom, $G(\chi_r^2) = P[\chi_r^2 \leq c]$ is the distribution of central χ^2 with r degrees of freedom, and

$$\eta_{-1,1} = 0 = \eta_{0,1} = \eta_{-1,2} = \eta_{1,3} = \eta_{0,4} = \eta_{1,4}$$

$$\eta_{1,1} = -\frac{\sigma^2 kn(n+2)}{\lambda(n+p+2)}(p-2), \quad \eta_{0,2} = \eta-\eta_{1,1}; \quad \eta = \frac{\sigma^2 k^2}{2\lambda}\,\frac{n(n+2)(n+4)}{(n+p+2)}$$

$$(3.7) \quad \eta_{1,2} = \frac{-\eta(n+2)}{(n+p+4)}, \quad \eta_{-1,3} = -\eta, \quad \eta_{0,3} = \frac{2\eta(n+4)}{(n+p+4)}, \quad \eta_{-1,4} = \frac{-\eta(n+6)}{(n+p+4)}$$

$$\mu_0 = \frac{\sigma^2 kT}{\lambda}(p-2+\frac{3}{2}kT), \quad \mu_{-1} = -\frac{\sigma^2 k^2 T^2}{\lambda}, \quad \mu_1 = -\frac{\sigma^2 kT}{\lambda}(p-2+\tfrac{1}{2}kT)$$

such that $\sum_{i=-1}^{1}\sum_{j=1}^{4}\eta_{i,j} = 0 = \sum_{i=-1}^{1}\mu_i$; $\lambda = \beta'X'X\beta$. We note here that the above results hold provided $\beta \neq 0$.

Similarly, the small-σ expansion, up to order $O(\sigma^2)$, and the large-T expansion up to order $O(T^{-1})$, of the the distribution of Q_s are, respectively, given by (provided $\beta \neq 0$).

$$(3.8) \quad P[Q_s \leq c] = G(pF_{p,n}) + \sum_{i=-1}^{1}\sum_{j=1}^{4}\eta^*_{i,j}\, G(\frac{n(p+2i)}{n+2j}F_{p+2i,n+2j})$$

and

$$(3.9) \quad P[Q_s \leq c] = G(\chi_p^2) + \sum_{i=-1}^{1}\mu_i^* G(\chi_{p+2i}^2)$$

where $\eta^*_{i,j}$ and μ^*_i are the same $\eta_{i,j}$ and μ_i given in (3.7) except that

$$(3.10) \quad \eta^*_{0,2} = \eta_{0,2} + \eta_0, \quad \eta^*_{1,1} = \eta_{1,1} - \eta_0; \quad \eta_0 = \frac{k\sigma^2}{2\lambda}\,\frac{n^2(n+2)}{n+p+2}\,(k - \frac{2(p-2)}{n+2})$$

$$\mu^*_0 = \mu_0 + \mu, \quad \mu^*_1 = \mu_1 - \mu; \quad \mu = \frac{k\sigma^2 T^2}{2\lambda}(k - \frac{2(p-2)}{T}).$$

The derivation is given in Section 3.2.

From (3.6) and (3.9) it is clear that

$$(3.11)\quad P[Q_s \le c] - P[Q_a \le c] = \mu[G(\chi^2_p) - G(\chi^2_{p+2})] \le 0,$$

or $P[Q_a \le c] \ge P[Q_s \le c]$, provided

$$(3.12)\quad 0 < k \le \frac{2(p-2)}{T}$$

Further,under (3.12),it can easily be verified that $P[Q_a \le c] \ge P[Q_{\ell s} \le c]$

It is interesting to note that the condition on k in (3.12) is the same as the condition (2.6) for the dominance of $\hat{\beta}$ over $\dot{b}$ under the MSE criterion, for large T.

The above results show that when the true value of β is not zero, the point towards which $\hat{\beta}$ shrinks b, then Q_a performs better than both Q_s and $Q_{\ell s}$. However, as shown below, this may not be true when the true value of β is zero.

When β=0, the expansions in (3.5) and (3.8) do not hold. To obtain the results in this case, we first substitute β=0 in Q_s, Q_a and $Q_{\ell s}$ and rewrite them as

$$Q_s = nf(1 - \frac{k}{f})^2 [1 + \frac{nk^2}{f} + \frac{2nk}{f(n+2)}]^{-1} = Q_s(k)$$

$$(3.13)\quad Q_a = n(f-2k+\frac{k^2}{f}) = Q_a(k)$$

$$Q_{\ell s} = nf.$$

where f=b'X'Xb/(y-Xb)'(y-Xb) and use has been made of (2.8). Now observe from (2.6) and (3.12) that, for large T, k will usually be a small positive number. Also for k=0, $Q_s=Q_a=Q_{\ell s}$. Thus, using a Taylor series expansion around k=0 we can approximate $Q_s(k)$ and $Q_a(k)$, as

$$Q_s \simeq nf-4kn\left(\frac{n+1}{n+2}\right)$$

(3.14)
$$Q_a \simeq nf-2kn.$$

It is clear from(3.14) that

(3.15) $P[Q_s \leq m] \geq P[Q_a \leq m] \geq P[Q_{\ell s} \leq m].$

That is, Q_s outperforms both Q_a and $Q_{\ell s}$.

These results suggest that when $\hat{\beta}$ shrinks b towards β, Q_s is better than both Q_a and $Q_{\ell s}$. However, when $\hat{\beta}$ shrinks b towards a vector different from β, Q_a is better than both Q_s and $Q_{\ell s}$.

3.1 <u>Monte Carlo Analysis</u>:

The approximate analytical results in (3.11) and (3.15) were derived on the basis of small-σ, large-T or small-k assumptions. Although exact analytical results would be most desirable, they would be quite involved, especially for Q_s, and so we employed Monte Carlo simulation to produce some further evidence on the desirability of Q_s and Q_a vis-a-vis $Q_{\ell s}$ under the criterion (3.4).[2] For this purpose we employed a fixed X matrix with ten rows and five columns measured in deviations from their averages. For each of the 2500 replications uniform pseudo-random numbers were generated using a procedure by McLeod (1982) and converted by the Box-Muller (1958) transformation to normal errors with mean zero and σ=1.25. The resulting u vectors were used in (2.1) to produce y vectors which were then converted to deviations from their sample averages. Equations (2.1), (2.4), (3.1), (3.2) and (3.3) were then evaluated and raw estimates of c_s, c_a and $c_{\ell s}$ were obtained from the empirical distributions of Q_s, Q_a and $Q_{\ell s}$. The difference between the estimated $c_{\ell s}$ values and the known, true $c_{\ell s}$ values provided estimates of the Monte Carlo sampling errors which were subtracted from the raw c_s and c_a estimates. Two experiments of this type were performed. In the first β was set equal to a vector of zeros, the vector towards which SR shrinks b. In the second experiment β was set equal to a vector of ones. The results are shown in Table 2.

These results suggest that when $\hat{\beta}$ shrinks b towards β, Q_s offers a substantial improvement over Q_a which is in turn considerably better than $Q_{\ell s}$. However, when $\hat{\beta}$ shrinks b towards a vector different from β, as in experiment 2, Q_a is preferred to both Q_s and $Q_{\ell s}$ which seem tied, the difference between them being, most credibly, residual Monte Carlo sampling error. Overall the results confirm our analytical findings in (3.11) and (3.15).

3.2 Derivation of the Results

Here we derive the results in (3.8) and (3.9). For this purpose let us rewrite the model (2.1) as

(3.16) $y=X\beta+\sigma v \qquad [u=\sigma v]$

so that v follows a multivariate normal distribution $N(0,I_T)$ and as σ approaches 0, the disturbance term grows small. Now we note that

$$(y-Xb)'(y-Xb)=\sigma^2 w;\ w=v'\bar{P}_X v \text{ and } \bar{P}_X=I-X(X'X)^{-1}X'.$$

Now, for sufficiently small σ, as in Kadane (1971), we can write $\hat{\beta}$ in (2.4) as

(3.17) $\hat{\beta}-\beta=\sigma e_1+\sigma^2 e_2+\sigma^3 e_3$

where

(3.18) $e_1=(X'X)^{-1/2}e,\ e_2=-\frac{kw}{\lambda}\beta,\ e_3=-\frac{kw}{\lambda}Ae;\ e=(X'X)^{-1/2}X'v\sim N(0,I_p),$

$A=(I_p-2\beta\beta'X'X/\lambda)(X'X)^{-1/2}$ and $\lambda=\beta'X'X\beta$ as in (3.7).

Similarly we can write $M(\hat{\hat{\beta}})$ in (2.8), up to order $O(\sigma^4)$, as

(3.19) $$M(\hat{\hat{\beta}})=\sigma^2\frac{w}{n}(X'X)^{-1}+\sigma^4w^2B;$$

where

(3.20) $$B=[k(k+\frac{4}{n+2})\frac{\beta\beta'}{\lambda}-\frac{2k}{n+2}(X'X)^{-1}]\frac{1}{\lambda}.$$

Substituting (3.17) and (3.19) in (3.1) and simplifying we can write Q_s, up to order $O(\sigma^2)$, as

(3.21) $$\frac{Q_s}{n}=e'C_1e+\sigma\delta'e+\sigma^2w\delta_0$$

where

(3.22) $$C_1=w^{-1}I_p+\sigma^2C;\ C=\frac{-2k}{\lambda}(I_p-\frac{2}{\lambda}(X'X)^{1/2}\beta\beta'(X'X)^{1/2})-n(X'X)^{1/2}B(X'X)^{1/2}$$

$$\delta'=\frac{-2k}{\lambda}\beta'(X'X)^{1/2},\ \delta_0=\frac{k^2}{\lambda}.$$

Now the characteristic function of $Q=Q_s/n$ is

(3.23) $$\phi(t)=E[\exp\{itQ\}]=E_w[\exp(it\sigma^2w\delta_0)E_e\exp\{it(e'C_1e+\sigma\delta'e)|_w\}]$$

because w and e are independent.

Using (3.22), it is easy to see that up to order $O(\sigma^2)$

(3.24) $$E_e\exp\{it(e'C_1e+\sigma\delta'e)|_w\}=|J|^{-1/2}[1+it\sigma^2\operatorname{tr}J^{-1}C-\frac{t^2}{2}\sigma^2\delta'J^{-1}\delta],$$

where $J=(1-2itw^{-1})I_p$. Also up to order $O(\sigma^2)$, $\exp\{it\sigma^2w\delta_0\}=1+it\sigma^2w\delta_0$. Thus, substituting this and (3.24) in (3.23) we get, up to $O(\sigma^2)$.

$$(3.25)\quad \phi(t) = E_w|J|^{-\frac{1}{2}} + it\sigma^2 \operatorname{tr}CE_w(|J|^{-\frac{1}{2}}J^{-1}) - \frac{t^2}{2}\sigma^2\delta' E_w(|J|^{-\frac{1}{2}}J^{-1})\delta$$

$$+ it\sigma^2\delta_0 E_w(w|J|^{-\frac{1}{2}}).$$

We note that $E_w|J|^{-\frac{1}{2}} = E_w(1-2itw^{-1})^{-p/2} = \phi(F^0_{p,n},t)$, that is the characteristic function of a central $pF/n=F^0$ ratio with the p and n degrees freedom, see Phillips (1982, p. 262). Similarly, it can easily be verified that $E_w(|J|^{-\frac{1}{2}}J^{-1}) = \phi(F^0_{p+2,n},t)$ and $E_w(\frac{w}{n}|J|^{\frac{1}{2}}) = \phi(F^0_{p,n+2},t)$. Using these results we can write

$$(3.26)\quad \phi(t) = \phi(F^0_{p,n},t) + it\sigma^2(\operatorname{tr}C)\phi(F^0_{p+2,n},t) - \frac{t^2}{2}\sigma^2\delta'\delta\phi(F^0_{p+2,n},t)$$

$$+ it\sigma^2 n\delta_o\phi(F^0_{p,n+2},t).$$

Now, using the Inversion Theorem, the density function of Q can be written after simplification as

$$(3.27)\quad g_o(Q) = g(F^0_{p,n}) + \eta[g(F^0_{p,n+4}) - g(F^0_{p-2,n+6})] + \eta_{1,1}[g(F^0_{p+2,n+2})$$

$$- g(F^0_{p,n+4})] + \frac{2\eta_{1,2}}{(n+2)}[(\frac{n}{2}+1)g(F^0_{p+2,n+4}) + (\frac{n}{2}+3)g(F^0_{p-2,n+8})$$

$$- 2(\frac{n}{2}+2)g(F^o_{p,n+6})]$$

where η, $\eta_{1,1}$ and $\eta_{1,2}$ are as given in (3.7). The result in (3.8) then follows. Further the results in (3.9) is obtained by noting that $(p+2i)F_{p+2i;n+2j}$ tends to χ^2_{p+2j} as $T \to \infty$.

4. HYPOTHESIS TESTING[3]

Suppose we want to test the hypothesis,

(4.1) $H_0: \beta=0$ against $H_1: \beta\neq 0$.

for the linear model $y=X\beta+u$ in (2.1). For this purpose the uniformly most powerful invariant test statistic is available, which is

$$(4.2) \quad F = \frac{1}{p} b'[\hat{Asy}\cdot V(b)]^{-1} b = \frac{1}{p} b'[s^2 (X'X)^{-1}]^{-1} b = \frac{b'X'Xb}{ps^2}$$

where the LS estimator b and $\hat{Asy}\cdot V(b)$ are given in (2.2) and (2.3), respectively.

The distribution of the F-ratio in (4.2) is an F-distribution with p and n degrees of freedom and

(4.3) $\lambda_0 = \beta'X'X\beta/2\sigma^2$

as the noncentrality parameter.

Consider now the F-ratio based on the SR estimator $\hat{\beta}$ in (2.4) as

$$(4.4) \quad F^* = \frac{1}{p}\hat{\beta}'[\hat{Asy}\cdot V(\hat{\beta})]^{-1}\hat{\beta} = \frac{1}{p}\hat{\beta}'[s^2(X'X)^{-1}]^{-1}\hat{\beta} = F[1 - \frac{nk}{pF}]^2,$$

where F is as in (4.2).

We note from (4.4) that since F* (the SR estimator based F-ratio) is a function of F, it is invariant. Thus, from the uniformly most powerful invariance property of F, it follows that the power of F will be uniformly higher than that of F*. This result is also true if $\hat{Asy}\cdot V(\hat{\beta})$ in F* is replaced by the unbiased estimator $\hat{M}(\hat{\beta})$ given in (2.8). This because F* based on $\hat{M}(\hat{\beta})$ is also a function of F and hence invariant. The main conclusion then is that while the SR estimator is known to dominate the LS estimator in the estimation context, it is uniformly dominated by the LS in the testing context.

The results in Table 3 below give an illustration of the powers of F* in (4.4) and F in (4.2). The exact power of F* is calculated by noting that

$$P[F^* > c] = P[\delta_1 F^2 + \delta_2 F + \delta_3 > 0]$$

where $\delta_1 = p^2$, $\delta_2 = -(2nkp + p^2 c)$, $\delta_3 = n^2 k^2$ and c is the critical value of F* at a 5% level of significance. The value of k was set to (p-2)/(n+2) for our calculations. The results in Table 3 show that the difference between the powers is most pronounced for low λ_0 values and converges to zero as λ_0 grows large.

TABLE 1: POINT ESTIMATES AND ESTIMATED RMSE FOR AIGNER AND JUDGE FULL MODEL

LS ESTIMATES[a]		SR ESTIMATES		
b	SE(b)	$\hat{\beta}$	bias	RMSE
41.18	18.37	29.75	-11.43	19.05
-42.9	18.5	-30.99	11.91	19.6
9050	19027	6538	-2512	13376
8778	4633	6341	2437	4399
282.5	770.2	204.1	-78.41	535.5
376.5	580.4	-272.0	104.5	417.7
253.6	282.7	183.2	-70.39	212.4
13050	6076	-9428	3622	6153
12040	5169	-8698	3342	5480

[a] These two columns are taken directly from Aigner and Judge (1977)

TABLE 2: MONTE CARLO ESTIMATES OF c_s AND c_a

		EXPERIMENT 1		EXPERIMENT 2	
		$\beta'=[00000]$		$\beta'=[11111]$	
$1-\alpha$	$c_{\ell s}$	c_s	c_a	c_s	c_a
.50	5.18370	.81584	2.20450	5.25852	4.70864
.75	10.36150	4.17788	7.12842	10.48702	9.64516
.90	20.25290	12.46008	16.88394	20.24716	19.09978
.95	31.28028	22.63224	28.05418	30.21879	30.03253
.99	77.60929	67.82404	73.66024	77.74524	77.24054

TABLE 3: POWERS OF F AND F^*

	$p = 3, n = 6$		$p = 6, n = 40$		$p = 10, n = 24$	
λ_0	F	F^*	F	F^*	F	F^*
1.00	.08556	.08346	.08696	.07442	.07245	.06596
1.25	.09503	.09245	.09733	.08193	.07883	.07066
1.50	.10469	.10167	.10812	.08994	.08503	.07562
1.75	.11454	.11109	.11930	.09843	.09165	.08084
2.00	.12455	.12069	.13085	.10739	.09847	.08630
4.00	.20907	.20244	.23369	.19275	.15992	.13808
8.00	.38422	.37418	.46013	.40131	.30816	.27240
16.00	.67361	.66271	.80331	.75785	.61098	.56821
32.00	.93278	.92803	.98824	.98231	.92944	.91207
64.00	.99839	.99814	.99999	.99998	.99936	.99900

Footnotes

[*]The first author gratefully acknowledges research support from NSERC. The authors are thankful to M. King, P.C.B. Phillips and N.E. Savin for useful discussions on the subject matter of this paper.

[1]This condition was not mentioned in Ullah et al (1983, 1984).

[2]These results could be developed using the fractional calculus technique of Phillips (1984). This will be the subject of a future study.

[3]The material in this section provides an answer to questions raised by M. King during the presentation of an earlier version of the paper at the Finite Sample Econometrics Conference at the University of Western Ontario, August 25-27, 1985.

REFERENCES

Aigner, D.J. and G.J. Judge, (1977), "Application of Pre-Test and Stein Estimators to Economic Data," Econometrica, Vol. 45, pp. 1279-1288.

Box, G.E.P. and M.E. Muller, (1958), "A Note on the Generation of Normal Deviates" Annals of Mathematical Statistics, Vol. 28, pp. 610-611.

Judge, G.J. and M.E. Bock, (1978), The Statistical Implication of Pre-Test and Stein-Rule Estimators in Econometrics, North-Holland, New York.

Kadane, J.B., (1971), "Comparison of k-Class Estimators when the Disturbances are Small", Econometrica, Vol. 39, pp. 723-737.

McLeod, A.I., (1982), "Efficient Fortran Coding of a Random Number Generator," Department of Statistics and Actuarial Sciences, University of Western Ontario, TR-82-088.

Manjoge, S.S., (1984), "On Double k-Class Estimators of Coefficients in Linear Regressions," Economics Letters, Vol. 15, pp. 295-300.

Phillips, P.C.B., C 1982), "The True Characteristic Function of the F Distribution", Biometrika, Vol. 69, pp. 261-264..

Phillips, P.C.B., (1984), "The Exact Distribution of the Stein-Rule Estimator," Journal of Econometrics, Vol. 25, pp. 1-9.

Stein, C.M., (1981), "Estimation of the Mean of a Multivariate Normal Distribution," The Annals of Statistics, Vol. 9, pp. 1135-1167.

Ullah, A., R.A.L. Carter, and V.K. Srivastava, (1983), "The Sampling Distribution of Shrinkage Estimators and Their F-Ratios in the Regression Model," Research Report 8324, University of Western Ontario.

Ullah, A., R.A.L. Carter and V.K. Srivastava, (1984), "The Sampling Distribution of Shrinkage Estimators and Their F-Ratios in the Regression Model," Journal of Econometrics, Vol. 25, pp. 109-122.

Ullah, A. and S. Ullah, (1978), "Double k-Class Estimators of Coefficients in Linear Regression," Econometrica, Vol. 46, pp. 705-722.

Vinod, H.D. and A. Ullah, (1981), Recent Advances in Regression Methods, Marcel Dekker.