Unbiased Estimation of the MSE Matrices of Improved Estimators in Linear Regression

Alan T.K.Wan
Department of Management Sciences, City University of Hong Kong, Hong Kong

Anoop Chaturvedi
Department of Mathematics and Statistics, University of Allahabad, India

Guohua Zou
Institute of System Sciences, Chinese Academy of Sciences, China

Abstract

Stein-rule and other improved estimators have scarcely been used in empirical work. One major reason is that it is not easy to obtain precision measures for these estimators. In this paper, we derive unbiased estimators for both the mean squared error (MSE) and the scaled MSE matrices of a class of Stein-type estimators. Our derivation provides the basis for measuring the estimators' precision and constructing confidence bands. Comparisons are made between these MSE estimators and the least squares covariance estimator. For illustration, the methodology is applied to data on energy consumption.

MSC: primary 62J05

Keywords: confidence ellipsoid, energy consumption, precision measure, Stein-rule

1. Introduction

The past few decades have seen the emergence of a large body of literature that concentrates on estimators that fall outside the tradition of unbiased estimation. Much of this is attributed to the seminal articles of Stein (1956) and James and Stein (1961), who demonstrated that the normal sample mean is inadmissible and uniformly dominated by the so- called Stein-rule (SR) estimator in terms of quadratic risk provided that the dimension is at least three. Stein's finding astounded the statistics profession when it was first published and subsequently opened up a new stream of research dedicated to obtain improved estimators over the traditional rules. Following the publication of James and Stein's results, many authors have proposed other classes of improved estimators for other loss functions, distributional assumptions and model specifications; in fact, some of the proposed estimators dominate the SR estimator, too (Baranchik, 1964, Stein, 1966, Berry, 1994). Hoffmann (2000), in a review article, suggested that the list of participants in this line of research reads like a "Who's who" in modern 20^{th} century statistics.

While Stein's finding is undoubtedly one of the most important results in modern statistical decision theory, it is also fair to say that the idea of Stein-rule has not really found its place in the applied statistics literature and the technique is vastly underutilized in applied work. A major drawback of the SR technique is that it is not straightforward to obtain a precision measure for the estimates. In consequence, Stein-type estimators cannot readily be used for the construction of confidence intervals or in hypothesis testing. One way to tackle this question is to carry out inference based on variance structure of a more conventional estimator, say, the ordinary least squares (OLS) estimator, with the Stein-type estimate replacing the least squares estimate. Clearly, given that the variance of the Stein-type estimator generally differs from that of the OLS estimator, inferences made in this way will inevitably be misleading. Another approach is to estimate the mean squared error (MSE) of the estimator directly and use the estimated MSE as the estimator's precision. Carter et al. (1990) derived an unbiased estimator of the MSE of the SR estimator in a normal linear regression model and used it to form confidence ellipsoids for the regression coefficients. They found that the resulting confidence set has a smaller expected squared volume than both the standard OLS confidence set and that based on the SR point estimator together with the OLS estimated covariance matrix.

In this paper we take an approach similar to the one developed in Carter et al. (1990). But unlike Carter et al. (1990), who focused solely on the SR estimator, we generalize the analysis to a wider family of improved estimators that includes the SR estimator, the double k-class (KK) estimator (Ullah and Ullah, 1978), the feasible and adjusted feasible minimum mean squared error (FMMSE and AFMMSE) estimators (Farebrother,1975; Ohtani, 1996a, 1996b) and the new SR estimator (Dasgupta and Sinha, 1999) as special cases. An unbiased estimator of the MSE matrix of this class of Stein-type estimators is derived in Section 2. Section 3 considers the empirical usefulness of the theoretical findings in the context of an energy consumption model discussed in Welsch (1989) and Adkins and Eells (1995). In Section 4, we consider the use of the estimated MSE matrices in the construction of the confidence ellipsoids for the coefficient vector. Section 5 concludes the paper.

2. Unbiased estimation of MSE matrix

To motivate discussion and establish notations, we consider the linear regression model,

$$y = X\boldsymbol{b} + u \quad ; \qquad\qquad u \sim N(0, \boldsymbol{s}^2 I) \qquad (2.1)$$

where y is a $n \times 1$ vector of observations on the dependent variable, X is a non-stochastic $n \times p$ matrix of full column rank, $\boldsymbol{b}$ is a $p \times 1$ vector of regression coefficients and u is a $n \times 1$ vector of disturbances. The least squares estimator of $\boldsymbol{b}$ is $b = (X'X)^{-1}X'y$, which is unbiased with covariance matrix $\boldsymbol{s}^2(X'X)^{-1}$. This covariance matrix is unbiasedly estimated by $\hat{V}(b) = s^2(X'X)^{-1}$, where $s^2 = \boldsymbol{J}/v$, $\boldsymbol{J} = (y - Xb)'(y - Xb)$ and $v = n - p$. Now, consider the class of estimators:

$$\hat{\boldsymbol{b}} = \left(1 - \frac{r(F)}{F}\right) b, \qquad (2.2)$$

where $F = b'X'Xb/\boldsymbol{J}$ and $r(F)$ is a function of F. Some specific forms of $r(F)$ lead to the following estimators:

- $r(F) = 0$ - the OLS estimator; (2.3)
- $r(F) = k,\ 0 < k < 2(p-2)/(n-p+2)$ - the SR estimator; (2.4)
- $r(F) = \dfrac{F}{1 + (n-p)F}$ - the FMMSE estimator; (2.5)
- $r(F) = \dfrac{pF}{p + (n-p)F}$ - the AFMMSE estimator; (2.6)
- $r(F) = \dfrac{k_1 F}{F + (1 - k_2)}$, k_1 and k_2 are arbitrary parameters - the KK estimator (2.7)

The SR estimator has been extensively analyzed in the literature of linear models; see, for example, Judge and Bock (1978). The FMMSE and AFMMSE estimators, due to Farebrother (1975) and Ohtani (1996b) respectively, are both adaptive versions of Theil's (1971) minimum mean squared error estimator. It is straightforward to show that the AFMMSE estimator is intrinsically identical to the new Stein-rule estimator given in Dasgupta and Sinha (1999). Both the FMMSE and AFMMSE estimators are consistent estimators of $\boldsymbol{b}$, but neither estimator has any minimum MSE property. In terms of quadratic risk, neither the FMMSE estimator nor the AFMMSE estimator strictly dominates the other, but the latter estimator is preferred to the former and the SR estimator over a large region of the parameter space (Ohtani, 1996b). Each of these estimators is a version of the KK estimator by varying the values of scalar parameters k_1 and k_2 in the latter's equation. Carter et al. (1993) showed that although there exists a range of k_1 and k_2 values such that the resulting KK estimator will dominate both the OLS and SR estimators, this range depends on unknown parameters, as do the values of k_1 and k_2 that minimize the KK estimator's risk. Carter et al. (1993) did suggest, however, that a 'nearly' optimal KK estimator by choosing $k_1 = (p-2)/(v+2)$ and $k_2 = 1 - k_1$. Ohtani (2000a) suggested an alternative "ad-hoc" choice by setting $k_1 = p/v$ and $k_2 = 1-(p-2)/v$.

It has been shown (see, for instance, Baranchik, 1970) that $\hat{\boldsymbol{b}}$ dominates b in the sense that

$$E\left[\left(\hat{\boldsymbol{b}}-\boldsymbol{b}\right)' X'X\left(\hat{\boldsymbol{b}}-\boldsymbol{b}\right)\right] < E\left[\left(b-\boldsymbol{b}\right)' X'X\left(b-\boldsymbol{b}\right)\right] \quad , \tag{2.8}$$

if i) $p \geq 3$, $0 \leq r(F) \leq 2(p-2)/(v+2)$ and $r(F)$ is not always equal to 0 or $2(p-2)/(v+2)$ and ii) $r'(F) \geq 0$. The estimator $\hat{\boldsymbol{b}}$ is biased. An unbiased estimator of the biased vector of $\hat{\boldsymbol{b}}$ is given by,

$$\hat{B}\left(\hat{\boldsymbol{b}}\right) = \hat{\boldsymbol{b}} - b = \frac{-r(F)}{F} b \quad , \tag{2.9}$$

while an unbiased estimator of the MSE matrix, $M(\hat{\boldsymbol{b}})$, of $\hat{\boldsymbol{b}}$ is,

$$\hat{M}\left(\hat{\boldsymbol{b}}\right) = s^2 (X'X)^{-1} + \left(\frac{vs^2}{b'X'Xb}\right)^2 r^2\left(\frac{b'X'Xb}{vs^2}\right) bb' - vs^2 \frac{vs^2}{b'X'Xb}\int_0^1 t^{v/2} r\left(\frac{b'X'Xb}{vs^2 t}\right) dt\,(X'X)^{-1}$$

$$-2\int_0^1 t^{v/2-2}\left[\frac{vs^2 t}{b'X'Xb} r\left(\frac{b'X'Xb}{vs^2 t}\right) - \left(\frac{vs^2 t}{b'X'Xb}\right)^2 r\left(\frac{b'X'Xb}{vs^2 t}\right)\right] dt\, bb'. \tag{2.10}$$

The proof of (2.10) is given in the Appendix. If $r(F)$ is a constant as in the case of SR estimator, then (2.10) reduces to the corresponding expression given in Carter et al. (1990). In general, the form of the estimator in (2.10) is not readily computable and requires numerical integration if it is to be calculated. In this case, one alternative is to derive an estimator of the "scaled" MSE of $\hat{\boldsymbol{b}}$, defined as $M\left(\hat{\boldsymbol{b}}\right)/\boldsymbol{s}^2$. An unbiased estimator of $M\left(\hat{\boldsymbol{b}}\right)/\boldsymbol{s}^2$ is,

$$\tilde{M}\left(\hat{\boldsymbol{b}}\right) = \left[1 - \frac{2vs^2}{b'X'Xb} r\left(\frac{b'X'Xb}{vs^2}\right)\right](X'X)^{-1} + \frac{vs^2}{(b'X'Xb)^2}\left\{(v+2) r^2\left(\frac{b'X'Xb}{vs^2}\right) + 4r\left(\frac{b'X'Xb}{vs^2}\right)\right.$$

$$\left. -4\frac{b'X'Xb}{vs^2} r'\left(\frac{b'X'Xb}{vs^2}\right)\left[r\left(\frac{b'X'Xb}{vs^2}\right)+1\right]\right\} bb'. \tag{2.11}$$

Details are given in the Appendix. Using (2.11), an alternative "indirect" estimator of $M(\hat{b})$ is given by,

$$\hat{M}^*\left(\hat{b}\right) = s^2\widetilde{M}\left(\hat{b}\right). \qquad (2.12)$$

Note that (2.11) involves no integrals and thus has a computational advantage over (2.10). Of course, the relevant question is whether the MSE values obtained from (2.12) are close enough to those obtained from (2.10). In the next section, the empirical usefulness of these results will be considered.

3. An illustrative example

A recent application of the Stein-rule technique was presented by Adkins and Eells (1995), in which the energy consumption models of Prosser (1985) and Welsh (1989) were estimated by the technique of Stein-rule. Adkins and Eells (1995) gave Stein-rule estimates of the energy models' parameters but did not consider the precision measure of the estimates. As an example to illustrate the utility of our results we choose the first model used by Adkins and Eells (1995), given as follows:

$$E_t = b_o + b_1 P_t + b_2 q_t + b_3 Y_t + b_4 E_{t-1} + e_t \qquad , \qquad (3.1)$$

where E = energy consumption, P = price index of energy (1990 = 100), q = price index of other goods and Y = gross domestic product at 1990 prices. Our data were obtained from the *Energy Balances of OECD Countries: Total Final Energy Consumption*, various years. In contrast to the results reported here, Adkins and Eells (1985) reported only SR estimates of the coefficients. Also, while Adkins and Eells' (1995) sample ended in 1989, we expanded their data to include observations from 1970 to 1997. But, unlike Adkins and Eells (1985), who considered the eight largest energy consuming nations of the OECD, our analysis was based on data of France, the Netherlands and the United Kingdom. Table 1 summarizes the results based on the OLS, SR, FMMSE, AFMMSE and the KK estimators using Carter et al. (1993)'s "nearly" optimal values of the characterizing parameters (i.e, $k_1 = (p-2)/(v+2)$ and $k_2 = 1-k_1$) and the "ad-hoc" values suggested by Ohtani (2000a) (i.e., $k_1 = p/v$ and $k_2 = 1-(p-2)/v$). In each case, we computed the point estimates of the regression coefficients and their root mean squared error (RMSE) estimates obtained from (2.10) and (2.12).

As remarked in Carter et al. (1990), the SR estimator dominates the OLS estimator only with respect to the estimator's risk of the entire coefficient vector but not necessarily with respect to the MSE of each coefficient. Hence we do not expect the estimated MSE of the SR estimates to be uniformly better than those of the OLS estimates. Indeed, in terms of the estimated MSE's, none of the six estimators uniformly dominate the others, though the "ad-hoc" version of the KK estimator seems to have the edge over the others in the majority of cases. Table 1 also shows that (2.12) gives MSE figures that are reasonably close to those obtained via the more "direct" expression (2.10). It therefore appears at least in the context of the example provided here that (2.12) is useful in assessing the estimation precision of the Stein-type estimators.

4. Construction of confidence ellipsoids

In this section, we assume that $X'X$ is $O(n)$. Note the following quadratic forms that define confidence ellipsoids for the vector of coefficients $\boldsymbol{b}$:

$$Q_s = \left(\hat{\boldsymbol{b}} - \boldsymbol{b}\right)' \left[\hat{M}\left(\hat{\boldsymbol{b}}\right)\right]^{-1} \left(\hat{\boldsymbol{b}} - \boldsymbol{b}\right) \qquad , \tag{4.1}$$

$$Q_a = \left(\hat{\boldsymbol{b}} - \boldsymbol{b}\right)' \left[s^2 (X'X)^{-1}\right]^{-1} \left(\hat{\boldsymbol{b}} - \boldsymbol{b}\right) \qquad , \tag{4.2}$$

and

$$Q_{ls} = (b - \boldsymbol{b})' \left[s^2 (X'X)^{-1}\right]^{-1} (b - \boldsymbol{b}) \qquad . \tag{4.3}$$

Note that (4.1) uses $\hat{\boldsymbol{b}}$ together with its estimated MSE matrix $\hat{M}\left(\hat{\boldsymbol{b}}\right)$, while (4.2) replaces $\hat{M}\left(\hat{\boldsymbol{b}}\right)$ by its asymptotic value $s^2 (X'X)^{-1}$, and (4.3), which defines a confidence ellipsoid of the OLS estimator, provides a benchmark for comparison.

Now, the confidence ellipsoids corresponding to (4.1), (4.2) and (4.3) are given by,

$$C_s = \{\boldsymbol{b} : Q_s \le c\} \ , \quad C_a = \{\boldsymbol{b} : Q_a \le c\} \quad \text{and} \quad C_{ls} = \{\boldsymbol{b} : Q_{ls} \le c\} \tag{4.4}$$

respectively, where c is a positive scalar.

As in Carter et al. (1990) and Chaturvedi et al. (1997), the confidence ellipsoids will be compared in terms of their expected squared volumes and coverage probabilities. In what follows, we assume, in addition to conditions i) and ii) in Section 2, that

iii) $\quad r'\left(\dfrac{\boldsymbol{b}'X'X\boldsymbol{b}}{n\boldsymbol{s}^2}\right) = r''\left(\dfrac{\boldsymbol{b}'X'X\boldsymbol{b}}{n\boldsymbol{s}^2}\right) = O\left(n^{-1}\right); \ r''\left(\dfrac{b'X'Xb}{vs^2}\right) = O_p\left(n^{-1}\right)$

and

iv) $\quad |\, r''(.)\, |$ is monotone and decreasing.

Note that conditions i) - iv) can be satisfied for all the estimators considered in Sections 2 and 3. For example, take the KK estimator, for which $r(F) = \dfrac{k_1 F}{F + (1 - k_2)}$, $r'(F) = \dfrac{k_1 (1 - k_2)}{[F + (1 - k_2)]^2}$ and $r''(F) = -\dfrac{2k_1 (1 - k_2)}{[F + (1 - k_2)]^3}$. Thus, conditions i) to iv) can be satisfied for $0 < k_1 \le 2(p-2)/(v+2)$ and $k_2 \le 1$.

From Carter et al. (1990), the approximate coverage probability of C_{ls}, up to $O(n^{-1})$, is given by,

$$P(Q_{ls} \le c) = G(\boldsymbol{c}_p^2) + \sum_{i=0}^{2} m_i G(\boldsymbol{c}_{p+2i}^2) \qquad , \tag{4.5}$$

where $G(\boldsymbol{c}_h^2) = P(\boldsymbol{c}_h^2 \le c)$, $\boldsymbol{c}_h^2$ is a central Chi-square variate with h degrees of freedom, $m_o = p(p-2)/(4n)$, $m_1 = -p^2/(2n)$ and $m_2 = p(p+2)/(4n)$.

It is shown in the Appendix that up to $O(n^{-1})$, the approximate coverage probability of C_a and C_s are,

$$P(Q_a \le c) = G(\boldsymbol{c}_p^2) + \sum_{i=0}^{2} m_i^* G(\boldsymbol{c}_{p+2i}^2) \tag{4.6}$$

and

$$P(Q_s \le c) = G(\boldsymbol{c}_p^2) + \sum_{i=0}^{2} m_i^{**} G(\boldsymbol{c}_{p+2i}^2) \tag{4.7}$$

respectively,

where $m_o^* = m_o - m,\ m_1^* = m_1 + m,\ m_2^* = m_2,\ m = \dfrac{\boldsymbol{s}^2 k^* n}{2\boldsymbol{b}'X'X\boldsymbol{b}}\left[k^* n - 2(p-2)\right] - 2r'\left(\dfrac{\boldsymbol{b}'X'X\boldsymbol{b}}{n\boldsymbol{s}^2}\right),$

$k^* = r\left(\dfrac{\boldsymbol{b}'X'X\boldsymbol{b}}{n\boldsymbol{s}^2}\right),\ m_o^{**} = m_o + m^*,\ m_1^{**} = m_1 - m^*,\ m_2^{**} = m_2$

and

$$m^* = \frac{n(p-2)\boldsymbol{s}^2}{\boldsymbol{b}'X'X\boldsymbol{b}}\left[r\left(\frac{\boldsymbol{b}'X'X\boldsymbol{b}}{n\boldsymbol{s}^2}\right) - \frac{(v+2)}{2}\int_0^1 t^{v/2} r\left(\frac{\boldsymbol{b}'X'X\boldsymbol{b}}{n\boldsymbol{s}^2 t}\right)dt\right] + 2\left[r'\left(\frac{\boldsymbol{b}'X'X\boldsymbol{b}}{n\boldsymbol{s}^2}\right) - \frac{n}{2}\int_0^1 t^{v/2-1} r'\left(\frac{\boldsymbol{b}'X'X\boldsymbol{b}}{n\boldsymbol{s}^2 t}\right)dt\right]. \tag{4.8}$$

Note that in the case of the SR estimator, i.e., $r(F) = k$, (4.6) and (4.7) reduce to the corresponding expressions given in Carter et al. (1990). Also, it is seen from (4.5) and (4.6) that,

$$P(Q_a \le c) - P(Q_{ls} \le c) = -m\left[G(\boldsymbol{c}_p^2) - G(\boldsymbol{c}_{p+2}^2)\right]. \tag{4.9}$$

Now, note that $r'(F) \ge 0$ for $F > 0$. Thus, $m < 0$ when $p \ge 3$ and $k^* < 2(p-2)/n$, which are implied by condition i). Thus, at least to order $O(n^{-1})$, $P(Q_a \le c) - P(Q_{ls} \le c) \ge 0$. In other words, the confidence ellipsoid based on $\hat{\boldsymbol{b}}$ and the estimated OLS covariance matrix has larger coverage probability than that based on b and the estimated OLS covariance matrix.

It is also seen from (4.6) and (4.7) that

$$P(Q_a \le c) - P(Q_s \le c) = -(m + m^*)\left[G(\boldsymbol{c}_p^2) - G(\boldsymbol{c}_{p+2}^2)\right]. \tag{4.10}$$

If $r(.)$ is a constant, then $P(Q_a \le c) - P(Q_s \le c) > 0$, as shown in Carter et al. (1990). On the other hand, if $r(.)$ is not a constant, then from condition ii),

$$\frac{(v+2)}{2}\int_0^1 t^{v/2} r\left(\frac{\boldsymbol{b}'X'X\boldsymbol{b}}{n\boldsymbol{s}^2 t}\right)dt > r\left(\frac{\boldsymbol{b}'X'X\boldsymbol{b}}{n\boldsymbol{s}^2}\right) \quad . \tag{4.11}$$

Thus,

$$m + m^* = \frac{\boldsymbol{s}^2 k^* n}{2\boldsymbol{b}'X'X\boldsymbol{b}}\left[k^* n - 2(p-2)\right] + \frac{n(p-2)\boldsymbol{s}^2}{\boldsymbol{b}'X'X\boldsymbol{b}}\left[r\left(\frac{\boldsymbol{b}'X'X\boldsymbol{b}}{n\boldsymbol{s}^2}\right) - \frac{(v+2)}{2}\int_0^1 t^{v/2} r\left(\frac{\boldsymbol{b}'X'X\boldsymbol{b}}{n\boldsymbol{s}^2 t}\right)dt\right]$$

$$-n\int_0^1 t^{v/2-1} r'\left(\frac{\boldsymbol{b}'X'X\boldsymbol{b}}{n\boldsymbol{s}^2 t}\right)dt \qquad , \tag{4.12}$$

which is negative, so long as $p \geq 3$ and $k^* < 2(p-2)/n$, which are implied by condition i). Thus, up to order $O(n^{-1})$, $P(Q_a \leq c) - P(Q_s \leq c) > 0$. So, for the entire class of estimators $\hat{\boldsymbol{b}}$ satisfying conditions i) to iv), the coverage probability for Q_a exceeds that of Q_s, up to order $O(n^{-1})$.

Finally, from (4.5) and (4.7), we observe that,

$$P(Q_s \leq c) - P(Q_{ls} \leq c) = \boldsymbol{m}^*\left[G(\boldsymbol{c}_p^2) - G(\boldsymbol{c}_{p+2}^2)\right] \qquad , \tag{4.13}$$

which can be either positive or negative in general. But from (4.8) and (4.11), if $p \geq 3$, and $r'\left(\frac{\boldsymbol{b}'X'X\boldsymbol{b}}{n\boldsymbol{s}^2}\right) = r''\left(\frac{\boldsymbol{b}'X'X\boldsymbol{b}}{n\boldsymbol{s}^2}\right) = O(n^{-2})$, then up to order $O(n^{-1})$, $\boldsymbol{m}^* < 0$ so long as $r(.)$ is not a constant. So, provided that these conditions are satisfied, then $P(Q_s \leq c) - P(Q_{ls} \leq c) < 0$. In other words, C_{ls} is preferred to C_s in terms of coverage probability if $r(.)$ is not a constant. Note that this finding contrasts with that of Carter et al. (1990) for the case of the SR estimator, for which $r(.)$ is a constant and $P(Q_s \leq c) = P(Q_{ls} \leq c)$, up to order $O(n^{-1})$.

Next, let' s compare the squared volumes of C_a, C_{ls} and C_s, which are given by,

$$V_a^2 = \frac{|s^2(X'X)^{-1}|(c\boldsymbol{p})^p}{(\Gamma(p/2+1))^2} = gs^{2p}, \tag{4.14}$$

$$V_{ls}^2 = V_a^2 \tag{4.15}$$

and

$$V_s^2 = \frac{|\hat{M}(\hat{\boldsymbol{b}})|(c\boldsymbol{p})^p}{(\Gamma(p/2+1))^2} \tag{4.16}$$

respectively, where $g = \frac{(c\boldsymbol{p})^p}{|X'X|(\Gamma(p/2+1))^2}$. As g is of order $O(n^{-p})$, the standardized squared volumes of C_a and C_{ls}, up to order $O_p(n^{-1})$, can be written as,

$$n^pV_a^2 = n^pV_{ls}^2 = \boldsymbol{s}^{2p}n^pg\left(1 + pd + \frac{p(p-1)}{2}d^2\right) \qquad , \tag{4.17}$$

where $d = (s^2 - \boldsymbol{s}^2)/\boldsymbol{s}^2$. On the other hand, from (A.20) in the Appendix, we have, up to order $O_p(n^{-1})$,

$$n^pV_s^2 = n^pgs^{2p}\left|\frac{(X'X)\hat{M}(\hat{\boldsymbol{b}})}{s^2}\right|$$

$$= s^{2p} n^p g\left[1 + pd + \frac{p(p-1)}{2}d^2 + s^2 l^* k^* n - \frac{s^2 n(v+2)(p-2)}{b'X'Xb}\int_0^1 t^{v/2} r\left(\frac{b'X'Xb}{ns^2 t}\right)dt \right.$$
$$\left. - 2n\int_0^1 t^{v/2-1} r'\left(\frac{b'X'Xb}{ns^2 t}\right)dt\right] \quad . \qquad (4.18)$$

Therefore, using (4.17) and (4.18),

$$n^p\left(V_{ls}^2 - V_s^2\right) = -s^{2p} n^p g\left[s^2 l^* k^* n - \frac{s^2 n(v+2)(p-2)}{b'X'Xb}\int_0^1 t^{v/2} r\left(\frac{b'X'Xb}{ns^2 t}\right)dt - 2n\int_0^1 t^{v/2-1} r'\left(\frac{b'X'Xb}{ns^2 t}\right)dt\right],$$
$$(4.19)$$

which is positive whenever $p \geq 3$ and $k^* < 2(p-2)/n,$ which are implied by condition i). Hence up to order $O(n^{-1})$, $E(V_{ls}^2) = E(V_a^2) > E(V_s^2)$. In other words, in terms of expected squared volume, the confidence ellipsoid C_s is most preferred.

5. Concluding Remarks

In various guises, shrinkage estimation has occupied an important place in the statistical decision theory literature. The technique is, however, underutilized in applied work, owing to the fact most shrinkage estimators cannot be readily used for inference purposes. This paper derives unbiased estimators of the MSE and the scaled MSE matrices of a class of shrinkage estimators, including the SR, FMMSE, AFMMSE and KK estimators. One particular advantage afforded by the estimator of the scaled MSE matrix over that of the MSE matrix is the achievement of considerable ease in computation. A numerical example taken from the energy economics literature has been used to illustrate the utility of our findings. It is also found that the confidence ellipsoid based on $\hat{b}$ and the unbiased estimator of its MSE matrix, for making confidence statements about the whole vector $b,$ is most preferred in large samples in terms of expected squared volume; but it can be the least preferred ellipsoid in terms of coverage probability. On the other hand, the ellipsoid based on $\hat{b}$ and the OLS estimated covariance matrix has the largest coverage probability in large samples among the three confidence ellipsoids considered. It remains, however, to compare these ellipsoids in terms of their small sample properties. Alternatively, the confidence bands for these estimators may be constructed via bootstrapping. Recent work by Kazimi and Brownstone (1999) and Ohtani (2000b, Ch. 7) may offer some insights in this regard.

References

Adkins, L.C. and Eells, J.B. (1995), "Improved estimators of energy models", *Energy Economics* 17, 15-25.

Baranchik, A.J. (1964), "Multiple regression and estimation of the mean of a multivariate normal distribution", Technical Report No. 51, Department of Statistics, Stanford University.

Baranchik, A.J. (1970), "A family of minimax estimators of the mean of a multivariate normal distribution", *Annals of Mathematical Statistics* 41, 642-645.

Berry, J.C. (1994), “Improving the James-Stein estimator using the Stein variance estimator”, *Statistics and Probability Letters* 20, 241-245.

Carter, R.A.L., Srivastava, M.S., Srivastava, V.K. and Ullah, A. (1990), “Unbiased estimation of the MSE matrix of Stein-rule estimators, confidence ellipsoids and hypothesis testing”, *Econometric Theory* 6, 63-74.

Carter, R.A.L., Srivastava, V.K. and Chaturvedi, A. (1993), “Selecting a double k-class estimator for regression coefficients”, *Statistics and Probability Letters* 18, 363-371.

Chaturvedi, A., Hasegawa, H., Chaturvedi, A. and Shukla, G. (1997), “Confidence sets for the coefficients vector of a linear regression model with non-spherical disturbances”, *Econometric Theory* 13, 406-429.

Dasgupta, A. and Sinha, B.K. (1999), “A new general interpretation of the Stein-estimate and how it adapts: applications”, *Journal of Statistical Planning and Inference* 75, 247-268.

Farebrother, R.W. (1975), “The minimum mean square error linear estimator and ridge regression”, *Technometrics* 17, 127-128.

Hoffmann, K. (2000), “Stein estimation – a review”, *Statistical Papers* 41, 127-158.

James, W. and Stein, C. (1961), “Estimation with quadratic loss”, in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press: Berkeley, 361-379.

Judge, G.G. and Bock, M.E. (1978), *The Statistical Implications of Pre-test and Stein-rule estimators in Econometrics*, Amsterdam: North Holland.

Kazimi, C. and Brownstone, D. (1999), “Bootstrap confidence bands for shrinkage estimators”, *Journal of Econometrics* 90, 99-127.

Ohtani, K. (1996a), “Exact small sample properties of an operational variant of the minimum mean squared error estimator”, *Communications in Statistics: Theory and Methods* 25, 1223-1231.

Ohtani, K. (1996b), “On an adjustment of degrees of freedom in the minimum mean squared error estimator”, *Communications in Statistics: Theory and Methods* 25, 3049-3058.

Ohtani, K. (2000a), “Pre-test double k-class estimators in linear regression”, *Journal of Statistical Planning and Inference* 87, 287-299.

Ohtani, K. (2000b), *Shrinkage Estimation of a Linear Regression Model in Econometrics*, Nova Science: New York.

Prosser, R. (1985), “Demand elasticities in OECD: dynamical aspects”, *Energy Economics* 7, 9-12.

Stein, C. (1956), “Inadmissibility of the usual estimator for the mean of a multivariate normal distribution”, in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press: Berkeley, 197-206.

Stein, C. (1966), "An approach to the recovery of interblock information in balance incomplete block designs", in David, F.N. (ed.), *Research Papers in Statistics: Festschrift for J. Neyman*, John Wiley: New York, 351-366.

Theil, H. (1971), *Principles of Econometrics*, John Wiley: New York.

Ullah, A. and Ullah, S. (1978), "Double k-class estimators of coefficients in linear regression", *Econometrica* 12, 143-150.

Welsch, J. (1989), "The reliability of aggregate energy demand functions", *Energy Economics* 11, 285-292.

Appendix

<u>Proof of (2.10)</u>

Note that,

$$\begin{aligned} M\left(\hat{\boldsymbol{b}}\right) &= E\left[\left(\hat{\boldsymbol{b}}-\boldsymbol{b}\right)\left(\hat{\boldsymbol{b}}-\boldsymbol{b}\right)'\right] \\ &= \boldsymbol{s}^2 (X'X)^{-1} - E\left[\left\{(b-\boldsymbol{b})b' + b(b-\boldsymbol{b})'\right\}\frac{r(F)}{F}\right] + E\left[bb'\frac{r^2(F)}{F^2}\right] \quad , \qquad \text{(A.1)} \end{aligned}$$

where unbiased estimators of the first and third terms in (A.1) are $\dfrac{\boldsymbol{J}}{v}(X'X)^{-1}$ and $bb'\dfrac{r^2(F)}{F^2}$ respectively. Now, to obtain an unbiased estimator of the second term in (A.1), write $z=(X'X)^{1/2}b$, $\boldsymbol{q}=(X'X)^{1/2}\boldsymbol{b}$ and $r^*(F)=\dfrac{r(F)}{F}$. Note that $z \sim N(\boldsymbol{q}, \boldsymbol{s}^2 I_{p\times p})$, $\boldsymbol{J}/\boldsymbol{s}^2 \sim \boldsymbol{c}_v^2$ and Z and $\boldsymbol{J}$ are independently distributed. Thus, we have,

$$\begin{aligned} E\left[(b-\boldsymbol{b})b'\frac{r(F)}{F}\right] &= (X'X)^{-1/2} E\left[(z-\boldsymbol{q})z'r^*\left(\frac{z'z}{\boldsymbol{J}}\right)\right](X'X)^{-1/2} \\ &= (X'X)^{-1/2}\boldsymbol{s}^2 E\left[\frac{\partial}{\partial z}\left\{z'r^*\left(\frac{z'z}{\boldsymbol{J}}\right)\right\}\right](X'X)^{-1/2} \\ &= \boldsymbol{s}^2(X'X)^{-1/2}\left[E\left(r^*\left(\frac{z'z}{\boldsymbol{J}}\right)\right)I_{p\times p} + 2E\left(zz'\frac{1}{\boldsymbol{J}}r^{*\prime}\left(\frac{z'z}{\boldsymbol{J}}\right)\right)\right](X'X)^{-1/2} \\ &= \boldsymbol{s}^2(X'X)^{-1}E_b\left[E_J\left(r^*\left(\frac{b'X'Xb}{\boldsymbol{J}}\right)\right)\right] + 2\boldsymbol{s}^2 E_b\left[bb'E_J\left(\frac{1}{\boldsymbol{J}}r^{*\prime}\left(\frac{b'X'Xb}{\boldsymbol{J}}\right)\right)\right], \\ & \qquad \text{(A.2)} \end{aligned}$$

where $r^{*\prime}(F)=\dfrac{\partial}{\partial F}r^*(F)$. Hence the second term of $M(\hat{\boldsymbol{b}})$ becomes,

$$-2(X'X)^{-1}E_b\left[E_J\left(\boldsymbol{s}^2 r^*\left(\frac{b'X'Xb}{\boldsymbol{J}}\right)\right)\right] - 4E_b\left[bb'E_J\left(\frac{\boldsymbol{s}^2}{\boldsymbol{J}}r^{*\prime}\left(\frac{b'X'Xb}{\boldsymbol{J}}\right)\right)\right]. \qquad \text{(A.3)}$$

Now, let $H(\boldsymbol{J})$ be a continuous differentiable function of $\boldsymbol{J}$, then it can be shown that,

$$\frac{1}{\boldsymbol{s}^2} E_J\left[\left(\boldsymbol{J} - v\boldsymbol{s}^2\right)H(\boldsymbol{J})\right] = 2E_J\left[\boldsymbol{J}\frac{\partial}{\partial \boldsymbol{J}}H(\boldsymbol{J})\right] \quad , \qquad \text{(A.4)}$$

or

$$E_J\left[\boldsymbol{J}H(\boldsymbol{J})\right] = \boldsymbol{s}^2 E_J\left[vH(\boldsymbol{J}) + 2\boldsymbol{J}\frac{\partial}{\partial \boldsymbol{J}}H(\boldsymbol{J})\right] \quad . \qquad \text{(A.5)}$$

Equation (A.5) implies that if

$$vH(\boldsymbol{J}) + 2\boldsymbol{J}\frac{\partial}{\partial \boldsymbol{J}}H(\boldsymbol{J}) = r^*\left(\frac{b'X'Xb}{\boldsymbol{J}}\right) \quad , \qquad \text{(A.6)}$$

then for fixed values of b, an unbiased estimator of $\boldsymbol{s}^2 E_J\left[r^*\left(\frac{b'X'Xb}{\boldsymbol{J}}\right)\right]$ is $\boldsymbol{J}H(\boldsymbol{J})$. Note that (A.6) is a first order differential equation and its solution is given by,

$$H(\boldsymbol{J}) = \frac{1}{2\boldsymbol{J}^{v/2}}\int_0^{J} \boldsymbol{f}^{v/2-1} r^*\left(\frac{b'X'Xb}{\boldsymbol{f}}\right)d\boldsymbol{f}$$

$$= \frac{1}{2}\int_0^1 r^*\left(\frac{b'X'Xb}{\boldsymbol{J}t}\right)t^{v/2-1}dt$$

$$= \frac{1}{2}\int_0^1 t^{v/2}\,\frac{r(F/t)}{F}dt \quad ,$$

from which it is straightforward to show that an unbiased estimator of $2E_b\left[E_J\left(\boldsymbol{s}^2 r^*\left(\frac{b'X'Xb}{\boldsymbol{J}}\right)\right)\right](X'X)^{-1}$ is,

$$\left[\boldsymbol{J}\int_0^1 t^{v/2}\,\frac{r(F/t)}{F}dt\right](X'X)^{-1} \quad . \qquad \text{(A.7)}$$

Similarly, an unbiased estimator of $4E_b\left[bb'E_J\left(\frac{\boldsymbol{s}^2}{\boldsymbol{J}}r^{*\prime}\left(\frac{b'X'Xb}{\boldsymbol{J}}\right)\right)\right]$ is,

$$2\left[\int_0^1 t^{v/2-2} r^{*\prime}\left(\frac{b'X'Xb}{\boldsymbol{J}t}\right)dt\right]bb' = 2\left[\int_0^1 t^{v/2-1}\,\frac{r'(F/t)}{F}dt - \int_0^1 t^{v/2}\,\frac{r(F/t)}{F^2}dt\right]bb' \quad . \qquad \text{(A.8)}$$

Hence (2.10) follows.

Proof of (2.11)

Note that from (A.1), we have,

$$M\left(\hat{\boldsymbol{\beta}}\right)/\sigma^2 = \left(X'X\right)^{-1} - \frac{1}{\sigma^2}E\left\{\left[(b-\boldsymbol{\beta})b' + b(b-\boldsymbol{\beta})'\right]\frac{r(F)}{F}\right\} + \frac{1}{\sigma^2}E\left[bb'\frac{r^2(F)}{F^2}\right]. \qquad \text{(A.9)}$$

From (A.2), an unbiased estimator of the second term in (A.9) is given by,

$$-2\left(X'X\right)^{-1}r^*\left(\frac{b'X'Xb}{\boldsymbol{\vartheta}}\right) - 4bb'\frac{1}{\boldsymbol{\vartheta}}r^{*\prime}\left(\frac{b'X'Xb}{\boldsymbol{\vartheta}}\right) \qquad . \qquad \text{(A.10)}$$

Also, by virtue of (A.5), we obtain,

$$\begin{aligned}
\frac{1}{\sigma^2}E\left[bb'\frac{r^2(F)}{F^2}\right] &= \frac{1}{\sigma^2}E_b\left[bb'E_\vartheta\left(\boldsymbol{\vartheta}\frac{r^{*2}(F)}{\boldsymbol{\vartheta}}\right)\right] \\
&= \frac{1}{\sigma^2}E_b\left[bb'\sigma^2 E_\vartheta\left(v\frac{r^{*2}(F)}{\boldsymbol{\vartheta}} + 2\boldsymbol{\vartheta}\frac{\partial}{\partial\boldsymbol{\vartheta}}\left(\frac{r^{*2}(F)}{\boldsymbol{\vartheta}}\right)\right)\right] \\
&= E\left[bb'\left((v-2)\frac{r^{*2}(F)}{\boldsymbol{\vartheta}} + 4r^*(F)\frac{\partial r^*(F)}{\partial\boldsymbol{\vartheta}}\right)\right] \qquad . \qquad \text{(A.11)}
\end{aligned}$$

Hence an unbiased estimator of the third term of (A.9) is,

$$bb'\left((v-2)\frac{r^{*2}(F)}{\boldsymbol{\vartheta}} - 4r^*(F)\frac{r^{*\prime}(F)F}{\boldsymbol{\vartheta}}\right) \qquad . \qquad \text{(A.12)}$$

Equation (2.11) is readily obtained by combining (A.9), (A.10) and (A.12).

<u>Proof of (4.6) and (4.7)</u>

Following Carter et al. (1990), we assume that $\boldsymbol{\beta}$ is non-zero. Now, as $X'X$ is $O(n)$, so $X'u$ is $O_p(n^{1/2})$, $s^2 = \sigma^2 + (s^2 - \sigma^2) = O(1) + O_p(n^{-1/2})$ and

$$\begin{aligned}
b'X'Xb &= u'X(X'X)^{-1}X'u + 2u'X\boldsymbol{\beta} + \boldsymbol{\beta}'X'X\boldsymbol{\beta} \\
&= O_p(1) + O_p(n^{1/2}) + O(n). \qquad \text{(A.13)}
\end{aligned}$$

Thus,

$$\begin{aligned}
\frac{b'X'Xb}{vs^2} &= \frac{\boldsymbol{\beta}'X'X\boldsymbol{\beta}}{n\sigma^2} + \frac{1}{n\sigma^2}\left(2u'X\boldsymbol{\beta} - \boldsymbol{\beta}'X'X\boldsymbol{\beta}\frac{(s^2-\sigma^2)}{\sigma^2}\right) + O_p(n^{-1}) \\
&= \frac{\boldsymbol{\beta}'X'X\boldsymbol{\beta}}{n\sigma^2} + O_p(n^{-1/2}) \qquad \text{(A.14)}
\end{aligned}$$

and

$$\frac{vs^2}{b'X'Xb} = \frac{n\sigma^2}{\beta'X'X\beta} + \frac{1}{\beta'X'X\beta}\left[n\left(s^2-\sigma^2\right) - \frac{2u'X\beta}{\beta'X'X\beta}n\sigma^2\right] + O_p(n^{-1}) \qquad . \qquad \text{(A.15)}$$

Recognizing (A.14) and conditions iii) and iv) in Section 4, we have,

$$r\left(\frac{b'X'Xb}{vs^2}\right) = r\left(\frac{\beta'X'X\beta}{n\sigma^2}\right) + \frac{1}{n\sigma^2}\left(2u'X\beta - \beta'X'X\beta\frac{(s^2-\sigma^2)}{\sigma^2}\right)r'\left(\frac{\beta'X'X\beta}{n\sigma^2}\right) + O_p(n^{-2})\,. \qquad \text{(A.16)}$$

Making use of (A.15) and (A.16), and after some manipulations, we obtain,

$$\hat{\beta} - \beta = e_{-1/2} + e_{-1} + e_{-3/2} + O_p(n^{-2}), \qquad \text{(A.17)}$$

where $e_{-1/2} = (X'X)^{-1}X'u,\ e_{-1} = -\lambda^*\sigma^2\beta,$

$$e_{-3/2} = \sigma^2\lambda^*\left[\left(\frac{2\beta'X'u}{\beta'X'X\beta} - \frac{s^2-\sigma^2}{\sigma^2}\right)\beta - (X'X)^{-1}X'u\right] - r'\left(\frac{\beta'X'X\beta}{n\sigma^2}\right)\left(\frac{2\beta'X'u}{\beta'X'X\beta} - \frac{s^2-\sigma^2}{\sigma^2}\right)\beta\,,$$

and

$$\lambda^* = \frac{k^* n}{\beta'X'X\beta} \qquad .$$

In addition, it can be shown that,

$$\int_0^1 t^{v/2} r\left(\frac{b'X'Xb}{vs^2 t}\right)dt = \int_0^1 t^{v/2} r\left(\frac{\beta'X'X\beta}{n\sigma^2 t}\right)dt + \frac{1}{n\sigma^2}\left(2u'X\beta - \beta'X'X\beta\frac{(s^2-\sigma^2)}{\sigma^2}\right)$$

$$\times\int_0^1 t^{v/2-1} r'\left(\frac{\beta'X'X\beta}{n\sigma^2 t}\right)dt + O_p(n^{-3}) \qquad \text{(A.18)}$$

and

$$\int_0^1 t^{v/2-1} r'\left(\frac{b'X'Xb}{vs^2 t}\right)dt = \int_0^1 t^{v/2-1} r'\left(\frac{\beta'X'X\beta}{n\sigma^2 t}\right)dt + O_p(n^{-5/2}) \qquad . \qquad \text{(A.19)}$$

Using (A.18) and (A.19), we can write,

$$\hat{M}\left(\hat{\beta}\right) = s^2(X'X)^{-1} + \sigma^2\Delta_{-2} + O_p(n^{-5/2}), \qquad \text{(A.20)}$$

where,

$$\Delta_{-2} = \frac{n\sigma^2}{\beta'X'X\beta}\left\{\left[\frac{nk^{*2} + 2(v+2)\int_0^1 t^{v/2} r\left(\frac{\beta'X'X\beta}{n\sigma^2 t}\right)dt}{\beta'X'X\beta} - \frac{2}{\sigma^2}\int_0^1 t^{v/2-1} r'\left(\frac{\beta'X'X\beta}{n\sigma^2 t}\right)dt\right]\beta\beta'\right.$$

$$-(v+2)\int_0^1 t^{v/2} r\left(\frac{b'X'Xb}{n\sigma^2 t}\right)dt\,(X'X)^{-1}\Bigg\} \quad . \tag{A.21}$$

Upon substituting (A.17) and (A.20) in (4.1), we obtain, up to $O_p(n^{-1})$,

$$Q_s = w'Dw + (\boldsymbol{d}'_{-1/2} + \boldsymbol{h}'_{-1})w + \boldsymbol{d}_{-1}, \tag{A.22}$$

where

$$\begin{aligned}
&w = \sigma^{-1}(X'X)^{1/2} e_{-1/2} \sim N(0, I_{p\times p}),\\
&\boldsymbol{d}'_{-1/2} = -2\boldsymbol{l}^* \sigma b'(X'X)^{1/2} \quad ,\\
&\boldsymbol{h}'_{-1} = \frac{2}{\sigma} r'\left(\frac{b'X'Xb}{n\sigma^2}\right)\left(\frac{s^2-\sigma^2}{\sigma^2}\right)b'(X'X)^{1/2},\\
&\boldsymbol{d}_{-1} = \sigma^2 \boldsymbol{l}^* k^* n,\\
&D = (1-d+d^2)I + (X'X)^{1/2}(\Delta^*_{-2} - \Delta_{-2})(X'X)^{1/2},
\end{aligned}$$

and

$$\Delta^*_{-2} = \frac{4\left[\sigma^2 \boldsymbol{l}^* - r'\left(\frac{b'X'Xb}{n\sigma^2}\right)\right]}{b'X'Xb}\boldsymbol{bb}' - 2\sigma^2 \boldsymbol{l}^*(X'X)^{-1}.$$

As in Carter et al. (1990), it can be shown that the characteristic function of Q_s is,

$$\begin{aligned}
\boldsymbol{f}(t) &= E[\exp(itQ_s)]\\
&= \exp(it\boldsymbol{d}_{-1})E_{s^2}E_{w|s^2}\{\exp[it(w'Dw + (\boldsymbol{d}'_{-1/2} + \boldsymbol{h}'_{-1})w)]\}\\
&= \exp(it\boldsymbol{d}_{-1})E_{s^2}\left\{|I-2itD|^{-1/2}\exp\left[-\frac{t^2}{2}(\boldsymbol{d}_{-1/2}+\boldsymbol{h}_{-1})'(I-2itD)^{-1}(\boldsymbol{d}_{-1/2}+\boldsymbol{h}_{-1})\right]\right\}.
\end{aligned} \tag{A.23}$$

Now, if the characteristic roots of the matrix D^* are of order $O_p(n^{-1/2})$, then,

$$|I-D^*|^{-1/2} = 1 + \frac{1}{2}tr(D^*) + \frac{1}{8}\left(tr(D^*)\right)^2 + \frac{1}{4}tr(D^{*2}) + O_p(n^{-3/2}). \tag{A.24}$$

Making use of (A.24) and after some manipulations, we obtain, up to $O(n^{-1})$,

$$\begin{aligned}
\boldsymbol{f}(t) = {}& \boldsymbol{f}(\boldsymbol{c}_p^2)(1 + it\sigma^2 \boldsymbol{l}^* k^* n) + \boldsymbol{f}(\boldsymbol{c}_{p+2}^2)\left\{it\left[\frac{2p}{n} - 2\sigma^2 \boldsymbol{l}^*(p-2) - 4r'\left(\frac{b'X'Xb}{n\sigma^2}\right) - tr(\Delta_{-2}X'X)\right]\right.\\
&\left. - 2t^2\sigma^2 \boldsymbol{l}^* k^* n\right\} - \boldsymbol{f}(\boldsymbol{c}_{p+4}^2)\frac{t^2 p(p+2)}{n},
\end{aligned} \tag{A.25}$$

where $\boldsymbol{f}(\boldsymbol{c}_p^2) = (1-2it)^{-p/2}$ is the characteristic function of a central Chi-square variate with p degrees of freedom.

Using these results, the density function of Q_s can be written as,

$$g(Q_s) = g\left(\boldsymbol{c}_p^2\right) + \left\{ \frac{p(p-2)}{4n} + \left[\boldsymbol{s}^2 \boldsymbol{l}^* (p-2) - \frac{1}{2} \boldsymbol{s}^2 \boldsymbol{l}^* k^* n + 2r' \left(\frac{\boldsymbol{b}' X' X \boldsymbol{b}}{n \boldsymbol{s}^2} \right) + \frac{1}{2} tr\left(\Delta_{-2} X' X\right) \right] \right\} g(\boldsymbol{c}_p^2)$$
$$+ \left\{ -\frac{p^2}{2n} - \left[\boldsymbol{s}^2 \boldsymbol{l}^* (p-2) - \frac{1}{2} \boldsymbol{s}^2 \boldsymbol{l}^* k^* n + 2r' \left(\frac{\boldsymbol{b}' X' X \boldsymbol{b}}{n \boldsymbol{s}^2} \right) + \frac{1}{2} tr\left(\Delta_{-2} X' X\right) \right] \right\} g(\boldsymbol{c}_{p+2}^2) + \frac{p(p+2)}{4n} g(\boldsymbol{c}_{p+4}^2), \qquad \text{(A.26)}$$

where $g(\boldsymbol{c}_h^2)$ is the density function of $\boldsymbol{c}_h^2$. Noting that,

$$tr\left(\Delta_{-2} X' X\right) = \boldsymbol{s}^2 \boldsymbol{l}^* k^* n - \frac{\boldsymbol{s}^2 n(v+2)(p-2)}{\boldsymbol{b}' X' X \boldsymbol{b}} \int_0^1 t^{v/2} r\left(\frac{\boldsymbol{b}' X' X \boldsymbol{b}}{n \boldsymbol{s}^2 t} \right) dt - 2n \int_0^1 t^{v/2-1} r'\left(\frac{\boldsymbol{b}' X' X \boldsymbol{b}}{n \boldsymbol{s}^2 t} \right) dt, \qquad \text{(A.27)}$$

and upon substituting (A.27) in (A.26), we obtain (4.7). Finally, equation (4.6) is obtained by setting $\Delta_{-2} = 0$ in (A.26).

Table 1. Estimates of Welsh's model

		OLS Estimates			SR Estimates		
Variable	Country	b	Bias	S.E.	$\hat{b}$	Bias	RMSE
ln(P)	UK	-0.8999708	-------	0.3459288	-0.6280049	0.2719659	0.4737593
ln(Y)		0.1897910		0.5188299	0.1324373	-0.0573537	0.3526133
ln(Q)		0.0363146		0.1970082	0.0253406	-0.0109741	0.1307734
ln(E(t-1))		-0.2471781		0.1753142	-0.1724823	0.0746958	0.1622926
Constant		0.3498934		0.1533703	0.2441577	-0.1057357	0.1904759
ln(P)	France	-0.0049931	-------	0.5894568	-0.0030423	0.0019509	0.3048398
ln(Y)		0.6124382		0.2994780	0.3731538	-0.2392845	0.3969687
ln(Q)		0.0509805		0.1838585	0.0310620	-0.0199185	0.0998282
ln(E(t-1))		-0.2201152		0.1686845	-0.1341144	0.0860008	0.1576927
Constant		-0.1794523		0.1380348	-0.1093389	0.0701135	0.1287081
ln(P)	Netherlands	-1.8248900	-------	0.7179774	-1.226295	0.5985944	1.0171120
ln(Y)		0.1389163		0.2109085	0.0933490	-0.0455669	0.1482198
ln(Q)		0.0724731		0.3817025	0.0487007	-0.0237724	0.2395982
ln(E(t-1))		-0.0795282		0.1685849	-0.0534416	0.0260866	0.1119330
Constant		-0.2998867		0.1816053	-0.2015188	0.0983679	0.1878146

		FMMSE Estimates				AFMMSE Estimates		
Variable	Country	$\hat{b}$	Bias	RMSE	RMSE*	$\hat{b}$	Bias	RMSE
ln(P)	UK	-0.8126538	0.0873170	0.3692356	0.3484700	-0.5854482	0.3145226	0.4928233
ln(Y)		0.1713771	-0.0184139	0.4707947	0.4669710	0.1234627	-0.0663283	0.3131455
ln(Q)		0.0327913	-0.0035233	0.1782517	0.1769791	0.0236234	-0.0126913	0.1147438
ln(E(t-1))		-0.2231963	0.0239817	0.1673933	0.1632661	-0.1607941	0.0863840	0.1596863
Constant		0.3159460	-0.0339474	0.1582642	0.1507801	0.2276123	-0.1222811	0.1962352
ln(P)	France	-0.0043841	0.0006090	0.5171256	0.5125413	-0.0029465	0.0020466	0.2714406
ln(Y)		0.5377367	-0.0747015	0.3129659	0.2932958	0.3614075	-0.2510307	0.3954401
ln(Q)		0.0447622	-0.0062183	0.1619170	0.1602615	0.0300842	-0.0208963	0.0901062
ln(E(t-1))		-0.1932667	0.0268483	0.1601107	0.1544858	-0.1298927	0.0902224	0.1541949
Constant		-0.1575638	0.0218885	0.1309481	0.1263698	-0.1058971	0.0735553	0.1258287
ln(P)	Netherlands	-1.634286	0.1906038	0.7739432	0.7706960	-1.1527020	0.6721877	1.0470951
ln(Y)		0.1244069	-0.0145094	0.1918056	0.1904445	0.0877473	-0.0511690	0.1361696
ln(Q)		0.0649036	-0.0075696	0.3424594	0.3399343	0.0457781	-0.0266951	0.2104453
ln(E(t-1))		-0.0712217	0.0083064	0.1522232	0.1511208	-0.0502344	0.0292938	0.1007126
Constant		-0.2685645	0.0313222	0.1774245	0.1764040	-0.1894251	0.1104616	0.1876349

RMSE = root mean squared error based on (2.10)
RMSE* = root mean squared error based on (2.12)

Table 1. Estimates of Welsh's model (cont'd)								
		CARTER's Estimates				AD-HOC Estimates		
Variable	Country	$\hat{b}$	Bias	RMSE	RMSE*	$\hat{b}$	Bias	RMSE
ln(P)	UK	-0.6911188	0.2088520	0.4266298	0.3655789	-0.5343351	0.3656357	0.5345029
ln(Y)		0.1457471	-0.0440439	0.3943797	0.3838483	0.1126836	-0.0771074	0.2694498
ln(Q)		0.0278873	-0.0084274	0.1479017	0.1446080	0.0215609	-0.0147537	0.0960803
ln(E(t-1))		-0.1898166	0.0573615	0.1608102	0.1473576	-0.1467558	0.1004223	0.1627225
Constant		0.2686953	-0.0811981	0.1747810	0.1520309	0.2077404	-0.1421530	0.2107531
ln(P)	France	-0.0035904	0.0014028	0.4016360	0.3901676	-0.0025452	0.0024481	0.1495028
ln(Y)		0.4403787	-0.1720595	0.3513918	0.2935929	0.3121784	-0.3002598	0.4322627
ln(Q)		0.0366579	-0.0143225	0.1275162	0.1230247	0.0259863	-0.0249942	0.0585501
ln(E(t-1))		-0.1582756	0.0618395	0.1542121	0.1361068	-0.1121994	0.1079158	0.1588108
Constant		-0.1290367	0.0504157	0.1259841	0.1112414	-0.0914723	0.0879800	0.1295081
ln(P)	Netherlands	-1.374146	0.4507431	0.9038249	0.8757168	-1.0365505	0.7883399	1.1431374
ln(Y)		0.1046043	-0.0343120	0.1634763	0.1595013	0.0789054	-0.0600109	0.1221847
ln(Q)		0.0545725	-0.0179007	0.2794115	0.2729862	0.0411652	-0.0313079	0.1665764
ln(E(t-1))		-0.0598849	0.0196433	0.1268641	0.1238635	-0.0451725	0.0343556	0.0857147
Constant		-0.2258155	0.0740713	0.1793351	0.1742233	-0.1703377	0.1295490	0.1966418

RMSE = root mean squared error based on (2.10)
RMSE* = root mean squared error based on (2.12)