1990

Performance of the stein-rule estimators when the disturbances are misspecified as homoscedastic

A Chaturvedi
University o f Allahbad

Tran Van Hoa
University of Wollongong

Govind Shukla
University of Allahbad

THE UNIVERSITY OF WOLLONGONG
DEPARTMENT OF ECONOMICS

PERFORMANCE OF THE STEIN-RULE ESTIMATORS WHEN THE DISTURBANCES ARE MISSPECIFIED AS HOMOSCEDASTIC

A. Chaturvedi*
Department of Mathematics and Statistics
University of Allahbad
Allahbad, India

Tran Van Hoa
Department of Economics
University of Wollongong
Wollongong NSW 2500 Australia

and

Govind Shukla
Department of Mathematics and Statistics
University of Allahbad
Allahbad, India

Working Paper 90-5

Co-ordinated by Dr Charles Harvie & Di Kelly
PO Box 1144 [Northfields Avenue], Wollongong NSW 2500 Australia
Phone: [042] 270 725 or 270 555. Telex 29022. Cable: UNIOFWOL. Fax [042] 270 477

ISSN 1035-4581
ISBN 0-8648-109-4

ABSTRACT

The paper investigates the effects of misspecifying the disturbances in a linear regression model as homoscedastic on the efficiency properties of the Stein-rule estimators. Asymptotic distribution of the Stein-rule estimator based on the OLS estimator is derived when the disturbances covariance matrix is nonscalar. The effects of non-homoscedasticity of the disturbances on the dominance conditions of the Stein-rule estimator is also observed. The risks under quadratic loss function of the Stein-rule estimators based on the OLS and the FGLS estimators are compared under a Pitman drift criterion.

Keywords: Linear regression models with nonscalar distr. bances, heteroscedasticity, AR(1), quadratic loss, Stein-rule estimators, feasible Stein-rule generalised least squares, Edgeworth expansions, Pitman drift.

1. INTRODUCTION

The Stein-rule estimator of the regression coefficient vector in a linear regression model with homoscedastic disturbances dominates, under the criterion of quadratic loss or mean squared errors, the ordinary least squares (OLS) estimator when some conditions related to the observation matrix of the explanatory variables are satisfied. In many practical applications of the linear regression model however, the assumption of homoscedasticity is usually not met, and the consequences of ignoring nonscalar disturbances when they are in fact present are well known in estimation and inference.

In an early study, Rothenberg (1984) considered a linear regression model with the disturbance covariance matrix depending upon a few unknown parameters and derived the Edgeworth type asymptotic expansion for the distribution of a linear function of the feasible generalized least squares (FGLS) estimator of the regression coefficients. Grubb and Magee (1988) obtained the asymptotic distribution of the OLS and FGLS estimators when the covariance parameters follow a Pitman drift such that the disturbance covariance matrix approaches the identity matrix asymptotically. They compared the variances of the OLS and FGLS estimators when the disturbance covariance matrix is locally nonscalar and discussed a number of examples. Chaturvedi and Shukla (1989) proposed a family of Stein rule estimators based on the FGLS estimator and derived the approximate distribution of this family of estimators when the sample size is large. They have also provided the conditions for the dominance of the Stein rule estimator over the FGLS estimator under a quadratic loss function and under the concentration probability approach.

In this paper, we consider a general linear model with nonscalar disturbances and investigate the efficiency properties of a family of Stein rule estimators that are based erroneously on the assumption of homoscedastic disturbances. First , Edgeworth type asymptotic expansions of the distribution of these estimators are derived and the risk under a quadratic loss function of these estimators is compared with that of the OLS estimator.

Acknowledgement. This paper was written when the first author was visiting the Department of Economics, Wollongong University. The author expresses his gratitude to the Department for inviting him to visit the Department as a Visiting Fellow and for providing research support.

Secondly, when the disturbance covariance matrix is locally nonscalar, the risk of the Stein rule estimator based on the OLS estimator is compared with the risk of the Stein rule estimator based on the FGLS estimator. Some examples of the basic results are are finally given

2. THE MODEL AND STEIN-RULE ESTIMATOR

Let us consider a general linear regression model

$$y = X\beta + u,\ Eu = 0,\ Eu\,u' = \sigma^2\,\Omega^{-1}, \tag{2.1}$$

where y is an n x 1 vector of observations on dependent variable, X is an n x p matrix of observations on p independent variables, β is a p x 1 vector of regression coefficients and u is an n x 1 vector of disturbances assumed to follow the normal probability law. The elements of $\Omega \equiv \Omega\,(\theta)$ are functions of an unknown q x 1 parameter vector θ. The parameter space of θ is an open subset of the q - dimensional space.

The OLS estimator of β is given by

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

with

$$E(b - \beta)\,(b - \beta)' = \sigma^2\,(X'X)^{-1}\,X'\,\Omega^{-1}\,X\,(X'X)^{-1}. \tag{2.3}$$

Let $\hat{\theta}$ be a consistent estimator of θ and $\hat{\Omega} = \Omega\,(\hat{\theta})$ is obtained by replacing θ by $\hat{\theta}$ in Ω. Then, a FGLS estimator of β is given by

$$\hat{\beta} = (X'\,\hat{\Omega}\,X)^{-1}\,X'\,\hat{\Omega}\,y. \tag{2.4}$$

Now, we consider the following family of Stein-rule estimators for which is based on the assumption that the disturbances are homoscedastic and obtained through shrinking the OLS estimator towards the null vector:

$$b(a) = \left[1 - \frac{a}{n}\,\frac{(y - Xb)'\,(y - Xb)}{b'X'Xb}\right] b. \tag{2.5}$$

3. ASYMPTOTIC DISTRIBUTION OF THE STEIN-RULE ESTIMATOR

Before obtaining the asymptotic distribution of b(a), we introduce the following notations:

$$D = \frac{1}{n} X'X, \quad D^* = \frac{1}{n} X' \Omega^{-1} X$$

$$\delta = \beta' D \beta, \quad \gamma = \frac{1}{n} \operatorname{tr} \Omega^{-1}$$

$$\zeta(a) = \frac{\sqrt{n}}{\sigma} D^{2} [b(a) - \beta]$$

$$\alpha = -\frac{a \gamma \sigma}{\sqrt{n}\, \delta} D^{\frac{1}{2}} \beta$$

$$\Sigma = D^{-\frac{1}{2}} D^* D^{-\frac{1}{2}} - \frac{2a \supseteq \gamma \sigma^2}{n \supseteq \delta} [D^{-\frac{1}{2}} D^* D^{-\frac{1}{2}} - \frac{1}{\delta} D^{\frac{1}{2}}$$
$$(\beta \beta' D^{-1} + D^{-1} \beta \beta') D^{-\frac{1}{2}}].$$

Theorem 1: Let $(\frac{1}{n} X'X)$ tends to a positive definite finite matrix as n tends to infinity. Then, the asymptotic distribution of $\zeta(a)$, to order $O_p(n^{-1})$, is normal with mean vector α and dispersion matrix Σ.

Proof: We can write b as

$$b = \beta + D^{-1} z, \tag{3.1}$$

where $z = n^{-\frac{1}{2}} X'u$ follows the normal distribution with mean vector 0 and dispersion matrix $\sigma^2 D^*$. Further

$$\frac{1}{n} (y - Xb)' (y - Xb)$$
$$= \frac{1}{n} u' \left[I_n - X (X'X)^{-1} X' \right] u$$
$$= \frac{1}{n} \operatorname{tr} \left[\{ I_n - X (X'X)^{-1} X' \} u u' \right]$$
$$= \gamma \sigma^2 + 0_p (n^{-1}), \tag{3.2}$$

and

$$\frac{1}{b'X'Xb} = \frac{1}{n(\beta + n^{\frac{1}{2}} D^{-1} z)' D (\beta + n^{-\frac{1}{2}} D^{-1} z)}$$
$$= \frac{1}{n\delta} \left[1 - \frac{2}{\delta \sqrt{n}} \beta' z \right] + 0_p (n^{-2}). \tag{3.3}$$

Utilizing (3.1), (3.2) and (3.3), we can write $\zeta(a)$, to order $0_p(n^{-1})$, as (3.4) z(a).

$$\zeta(a) = \frac{\sqrt{n}}{\sigma} D^{\frac{1}{2}}\left[b - \beta - \frac{a}{n}\frac{(y - Xb)'(y - Xb)}{b'X'Xb} b\right]$$

$$= \frac{1}{\sigma} D^{\frac{1}{2}}\left[D^{-1}z - \frac{a\gamma\sigma^2}{\sqrt{n}\,\delta}\left(1 - \frac{2}{\delta\sqrt{n}}\beta' z\right)(\beta + D^{-1}z)\right]$$

$$= \frac{1}{\sigma} D^{\frac{1}{2}} z - \frac{a\gamma\sigma}{\sqrt{n}\,\delta} D^{\frac{1}{2}}\beta - \frac{a\gamma\sigma}{n\,\delta}\left(D^{-\frac{1}{2}} z - \frac{2}{\delta}\beta\,\beta' z\right) \tag{3.4}$$

For any p x 1 vector h, the cumulant generating function of $\zeta(a)$, to order $0_p(n^{-1})$, is given by (3.5)

K(h)

$$= \log E\left[\exp\{i\, h'\, \zeta\,(a)\}\right]$$

$$= -i\frac{a\gamma\sigma^2}{\sqrt{n}\,\delta}(h' D^{\frac{1}{2}}\beta) + \log E\left[\exp\left(i\frac{1}{\sigma}h' D^{-\frac{1}{2}} z\right)\left\{1 - i\frac{a\gamma\sigma}{n\,\delta}\left(h' D^{-\frac{1}{2}} z - \frac{2}{\delta} h'\,\beta\beta' z\right)\right\}\right]$$

$$= -i\frac{a\gamma\sigma^2}{\sqrt{n}\,\delta} - \frac{1}{2} h D^{-\frac{1}{2}} D^* D^{-\frac{1}{2}} h + \log\left[1 + \frac{a\gamma\sigma}{n\,\sigma}\left\{h' D^{-\frac{1}{2}} D^* D^{-\frac{1}{2}} h - \frac{1}{\delta} h'(\beta\,\beta' D^* D^{-\frac{1}{2}} + D^{-\frac{1}{2}} D^*\beta\,\beta')\, h\right\}\right]$$

$$= i\, h'\alpha - \frac{1}{2} h'\Sigma\, h, \tag{3.5}$$

where we have utilized the following results

$$E\left[\exp\left(i\frac{1}{\sigma}h' D^{-\frac{1}{2}} z\right)\right] = \exp\left[-\frac{1}{2}h' D^{-\frac{1}{2}} D^* D^{-\frac{1}{2}} h\right]$$

$$E\left[z \exp\left(i\frac{1}{\sigma}h' D^{-\frac{1}{2}} z\right)\right]$$

$$= i\,\sigma D^* D^{-\frac{1}{2}} h \exp\left[-\frac{1}{2}h' D^{-\frac{1}{2}} D^* D^{-\frac{1}{2}} h\right]. \tag{3.5}$$

Therefore, to order 0 (n^{-1}), K(h) is same as the cumulant generating function of a normal variate with mean vector α and dispersion matrix Σ. Hence we follow the theorem.

The bias vector, to order $0(n^{-1})$, and the MSE matrix, to order $0(n^{-2})$, of b(a) are given by

$$E\left[b(a) - \beta\right] = -\frac{a\,\gamma\,\sigma^2}{n\,\delta}\,\beta \tag{3.6}$$

$$E\left[b(a) - \beta\right]\left[b(a) - \beta\right]'$$

$$= \frac{\sigma^2}{n}\left[(D^{-1} D^* D^{-1} - \frac{a\,\gamma\,\sigma^2}{n\,\delta}\right.$$

$$\left.\left\{2\, D^{-1} D^* D^{-1} - \frac{2}{\delta}(\beta\,\beta'\, D^* D^{-1} + D^{-1} D^*\, \beta\,\beta') - \frac{a\,\gamma}{\delta}\beta\,\beta'\right\}\right] \tag{3.7}$$

Consider the quadratic loss function

$$L(\tilde{\beta}, \beta) = (\tilde{\beta} - \beta)'\, Q\, (\tilde{\beta} - \beta), \tag{3.8}$$

where $\tilde{\beta}$ is any estimator of β and Q is a p x p positive definite, symmetric matrix. Then, to order $0(n^{-2})$, the risk of the estimator b(a) is given by

$$R[b(a)]$$

$$= \frac{\sigma^2}{n}\left[(tr\,(D^{-1} D^* D^{-1} Q) - \frac{a\,\gamma\,\sigma^2}{n\,\delta}\{2\, tr\,(D^{-1} D^* D^{-1} Q)\right.$$

$$\left. -\frac{4}{\delta}\,\beta' D^* D^{-1} Q\beta - \frac{a\,\gamma}{\delta}\beta' Q\,\beta\}\right] \tag{3.9}$$

Further, we have

$$R[b] = \frac{\sigma^2}{n}\, tr\,(D^{-1} D^* D^{-1} Q) \tag{3.10}$$

Therefore, to order $0(n^{-2})$, the difference between the risks of the estimators b and b(a) is given by

$$R[b] - R[b(a)]$$

$$= \frac{a\,\gamma\,\sigma^4}{n^2\,\delta}\left[2\, tr\,(D^{-1} D^* D^{-1} Q) - \frac{4}{\delta}\beta'\, D^* D^{-1}\, Q\,\beta - \frac{a\,\gamma}{\delta}\beta'\, Q\,\beta\right] \tag{3.11}$$

Let λ_ℓ (.) and λ_s(.) denote, respectively, the maximum characteristic root and minimum characteristic root of the matrix inside the bracket. Then, we observe that

$$\mathrm{tr}\,(D^{-1} D^* D^{-1} Q) \geq \lambda_s (\Omega^{-1})$$

$$\frac{\beta' Q \beta}{\delta} \leq \lambda_\ell \, (D^{-1} Q)$$

$$\frac{\beta' D^* D^{-1} Q \beta}{\delta} \leq \frac{(\beta' D^* \beta)^{\frac{1}{2}} (\beta' Q D^{-1} D^* D^{-1} Q \beta)^{\frac{1}{2}}}{\beta' D \beta} \leq \lambda_\ell (\Omega^{-1}) \lambda_\ell (D^{-1} Q),$$

and

$$\gamma \leq \lambda_\ell \, (\Omega^{-1}).$$

Hence, we get the following lower bound for the difference between the risks of b and b(a):

$$R[b] - R[b\,(a)]$$

$$\geq \frac{a \gamma \sigma^4}{n^2 \delta} \lambda_\ell (\Omega^{-1}) \lambda_\ell (D^{-1} Q) [2 (g - 2) - a], \tag{3.12}$$

where

$$g = \frac{\lambda_s (\Omega^{-1})}{\lambda_\ell (\Omega^{-1})} \frac{\mathrm{tr}\,(D^{-1}Q)}{\lambda_\ell (D^{-1}Q)} . \tag{3.13}$$

Therefore, a sufficient condition for b(a) to have lower risk than b is given by

$$0 \leq a \leq 2 (g - 2) \; ; \; g > 2. \tag{3.14}$$

Notice that as the ratio $\lambda_s (\Omega^{-1})/\lambda_\ell (\Omega^{-1})$ decreases, the range of a in which b(a) dominates b narrows. If $\lambda_s (\Omega^{-1})$ is very small as compared to $\lambda_\ell (\Omega^{-1})$, g may become less than 2 and the dominance condition (3.14) may not be satisfied.

In particular, if we take matrix of the loss function Q = D, the expression (3.9) for the risk of b(a) reduces to

$$R\,[b\,(a)]$$

$$= \frac{\sigma^2}{n} \left[\mathrm{tr}\,(D^{-1} D^*) - \frac{a \gamma \sigma^2}{n \delta} \{2 \,\mathrm{tr}\,(D^{-1} D^*) - \frac{4}{d} \beta' D^* \beta - a \gamma\} \right] \tag{3.15}$$

and the dominance condition (3.14) becomes

$$0 \leq a \leq 2 \left[p \frac{\lambda_s (\Omega^{-1})}{\lambda_\ell (\Omega^{-1})} - 2 \right]; p > 2 \frac{\lambda_\ell (\Omega^{-1})}{\lambda_s (\Omega^{-1})} . \tag{3.16}$$

4. COMPARISON OF THE STEIN-RULE ESTIMATORS BASED ON OLS AND FGLS ESTIMATORS:

Now, we consider the following family of Stein-rule estimators based on the FGLS estimator:

$$\hat{\beta}(a) = \left[1 - \frac{a}{n} \frac{(y - X\hat{\beta})' \hat{\Omega} (y - X\hat{\beta})}{\hat{\beta}' X' \hat{\Omega} X \hat{\beta}} \right] \hat{\beta}. \tag{4.1}$$

Following Chaturvedi and Shukla (1989), the approximate expressions for the bias vector and MSE matrix of $\hat{\beta}(a)$ are given by

$$E[\hat{\beta}(a) - \beta) = - \frac{a\sigma^2}{\phi n} \beta \tag{4.2}$$

$$E\left[\hat{\beta}(a) - \beta\right]\left[\hat{\beta}(a) - \beta\right]'$$

$$= \frac{\sigma^2}{n} \left[A^{-1} + \frac{1}{n} A^{-1} \left(\sum_{j.k=1}^{q} Q_{jk} \lambda_{jk} \right) A^{-1} - \frac{a\sigma^2}{n\phi} \left\{ 2A^{-1} - \frac{4+a}{\phi} \beta\beta' \right\} \right], \tag{4.3}$$

where

$$A = \frac{1}{n} X' \Omega X, \phi = \beta' A \beta$$

$$\Omega_j = \frac{\partial}{\partial\theta_j} \Omega, \; A_j = \frac{1}{n} X' \Omega_j X$$

$$\Omega_{jk} = \frac{1}{n} X' \Omega_j \Omega^{-1} \Omega_k X - A_j A^{-1} A_k,$$

and λ_{jk} is the (j,k) -th element of the matrix.

$$\Lambda = \lim_{n\to\infty} \left[n E(\hat{\theta} - \theta)(\hat{\theta} - \theta') \right]. \tag{4.4}$$

Now, we compare the risks of the estimators b(a) and $\hat{\beta}(a)$ under a Pitman drift process for θ. Let θ_0 be a value of θ such that $\Omega(\theta_0) = I_n$. Further, we assume that $\theta \equiv \theta_n$ follows the process

$$\theta_n = \theta_0 + n^{-\frac{1}{2}}\, \eta, \tag{4.5}$$

where η is an n x 1 vector of order 0(1). Thus

$$\lim_{n\to\infty} \Omega(\theta_n) = I_n.$$

For all j, k = 1, 2, ..., q, we define

$$\Omega_j(\theta_0) = \frac{\partial}{\partial\theta_j}\Omega \Big|_{\theta = \theta_0}, \quad \Omega_{jk}(\theta_0) = \frac{\partial^2}{\partial\theta_j\, \partial\theta_k}\, \Omega \Big|_{\theta = \theta_0}. \tag{4.6}$$

Along with the regularity assumptions given by Rothenberg (1984, p817) we also assume that for all $t \geq 3$, $n^{-t/2} X' \Omega^*_t X$ and $n^{-t/2} X' \Omega_{(t)} X$ are of order 0(1) as $n\to\infty$, where Ω^*_t is any product of the matrices $\Omega_j(\theta_0)$ and $\Omega_{jk}(\theta_0)$ (j, k = 1,...,q), such that total number of matrix subscripts is t, and $\Omega_{(t)}$ is any t-th order derivative of $\Omega(\theta)$ evaluated at $\theta = \theta_0$.

Now we write

$$A_j(\theta_0) = \frac{1}{n} X' \Omega_j(\theta_0) X, \quad A_{jk}(\theta_0) = \frac{1}{n} X' \Omega_{jk}(\theta_0) X$$

$$A^*_{jk}(\theta_0) = \frac{1}{n} X' \Omega_j(\theta_0)\, \Omega_k(\theta_0)\, X.$$

<u>Theorem 2:</u> If θ_n follows the process (4.5), the expressions for the MSE matrices of the estimators b(a) and $\hat{\beta}(a)$, to order $0(n^{-2})$, are given by

$$E\big[b(a) - \beta\big]\, \big[b(a) - \beta\big]'$$

$$= \frac{\sigma^2}{n} \Big[(D^{-1} - \frac{1}{\sqrt{n}} D^{-1} \sum_j A_j(\theta_0)\, \eta_j D^{-1} - \frac{1}{2n} D^{-1} \sum_{j.k} \{A_{jk}(\theta_0) - 2 A^*_{jk}(\theta_0)\}$$

$$\eta_j \eta_k D^{-1} - \frac{A\sigma^2}{n\delta} \Big(2 D^{-1} - \frac{4+a}{\phi} \beta \beta' \Big) \Big], \tag{4.7}$$

$$E\big[\hat{\beta}(a) - \beta\big]\, \big[\hat{\beta}(a) - \beta\big]'$$

$$= \frac{\sigma^2}{n}\Big[D^{-1} - \frac{1}{\sqrt{n}} D^{-1} \sum_j A_j(\theta_0)\, \eta_j D^{-1} - \frac{1}{2n} D^{-1} \sum_{j.k} \{A_{jk}(\theta_0)$$

$$- 2 A_j(\theta_0) D^{-1} A_k(\theta_0)\} \eta_j \eta_k D^{-1} + \frac{1}{n} D^{-1}$$

$$\sum_{j.k} \{A^*_{jk}(\theta_0) - A_j(\theta_0) D^{-1} A_k(\theta_0))\} \lambda_{jk} D^{-1}$$

$$- \frac{a\sigma^2}{n\delta}\left(2D^{-1} - \frac{4+a}{\phi}\beta\beta'\right), \tag{4.8}$$

where η_j is the j-th element of η.

Proof: When θ_n follows the process (4.5), we observe that

$$\Omega^{-1} = I_n - \frac{1}{\sqrt{n}} \sum_j \Omega_j(\theta_0)\, \eta_j + \frac{1}{n} \sum_{j.k} \Omega_j(\theta_0)\, \Omega_k(\theta_0)\, \eta_j \eta_k$$

$$- \frac{1}{2n} \sum_j \Omega_{jk}(q_0)\, \eta_j \eta_k + 0(n^{-3/2}), \tag{4.9}$$

$$D^* = D - \frac{1}{\sqrt{n}} \sum_j A_j(\theta_0)\, \eta_j + \frac{1}{n} \sum_{j.k} A^*_{jk}(\theta_0)\, \eta_j \eta_k$$

$$- \frac{1}{2n} \sum_{j.k} A_{jk}(\theta_0)\, \eta_j \eta_k + 0(n^{-3/2}), \tag{4.10}$$

$$\gamma = 1 + 0(n^{-\frac{1}{2}}). \tag{4.11}$$

Substituting the values of D^* and γ from (4.10) and (4.11) into the expression (3.7), we get the expression (4.7) for the MSE matrix of b(a).

For obtaining the expression (4.8) for the MSE matrix of $\hat{\beta}(a)$, we observe that

$$A^{-1} = D^{-1} - \frac{1}{\sqrt{n}} D^{-1} \sum_j A_j(\theta_0)\eta_j D^{-1} - \frac{1}{2n} D^{-1} \sum_{j.k} A_{jk}(\theta_0)\, \eta_j \eta_k D^{-1}$$

$$+ \frac{1}{n} D^{-1} \sum_{j.k} A_j(\theta_0) D^{-1} A_k(\theta_0)\, \eta_j \eta_k D^{-1} + 0(n^{-3/2}),$$

$$Q_{jk} = \frac{1}{n} X' \Omega_j(\theta_0)\, \Omega_k(\theta_0) X - A_j(\theta_0) D^{-1} A_k(\theta_0) + 0(n^{-\frac{1}{2}}).$$

Substitution of the above approxmate expression for A^{-1} and Q_{jk} in (4.3) leads to the expressions (4.8) for the approximate MSE matrix of $\hat{\beta}(a)$.

Using (4.7) and (4.8), we get the following expression for the difference between the approxmate MSE matrices of b(a) and $\hat{\beta}(a)$:

$$E[b(a) - \beta] [b(a) - \beta]' - E[\hat{\beta}(a) - \beta] [\hat{\beta}(a) - \beta]'$$

$$= \frac{\sigma^2}{n^2} D^{-1} \sum_{j.k} P_{jk}(\theta_0) (\eta_j \eta_k - \lambda_{jk}) D^{-1} \tag{4.12}$$

where

$$P_{jk}(\theta_0) = A^*_{jk}(\theta_0) - A_j (\theta_0) D^{-1} A_k(\theta_0)$$

$$= \frac{1}{n} X' \Omega_j(\theta_0) \left[I_n - \frac{1}{n} X D^{-1} X' \right] \Omega_k(\theta_0) X. \tag{4.13}$$

Notice that, to order $0(n^{-2})$, the difference between the MSE matrices of b(a) and $\hat{\beta}(a)$ is the same as the difference between the MSE matrices of b and $\hat{\beta}$. Under the quadratic loss function (3.8), the difference between the risks of b(a) and $\hat{\beta}(a)$, to order $0(n^{-2})$, is given by

$$R[b(a)] - R[\hat{\beta}(a)]$$

$$= \frac{\sigma^2}{n^2} \left[\eta' G(\theta_0) \eta - tr \{G(\theta_0) \Lambda \} \right] \tag{4.14}$$

where $G(\theta_0)$ is a q x q matrix with (j,k)-th element tr $\{D^{-1} P_{jk}(\theta_0) D^{-1} Q\}$. Hence b(a) dominates $\hat{\beta}(a)$ whenever

$$\eta' G(\theta_0) \eta \leq tr \{G(\theta_0) \Lambda \}, \tag{4.15}$$

and $\hat{\beta}(a)$ dominates b(a) whenever the reverse inequality holds.

In particular, if q = 1, we observe that

$$E[b(a) - \beta] [b(a) - \beta]' - E[\hat{\beta}(a) - \beta] [\hat{\beta}(a) - \beta]'$$

$$= \frac{\sigma^2}{n^2} D^{-1} P_{11}(\theta_0) D^{-1} (\eta^2 - \Lambda). \tag{4.16}$$

Since the matrix $P_{11}(\theta_0)$ is positive definite, the difference between the MSE matrices of b(a) and $\hat{\beta}(a)$ is positive (negative) definite whenever $\eta^2 > (<) \Lambda$, or in other words, whenever

$$\lim_{n\to\infty} \frac{n(\theta - \theta_0)^2}{\Lambda} > (<) 1. \tag{4.17}$$

5. EXAMPLES

In this section two examples are considered to illustrate how the results of the paper can be applied in special cases.

5.1 Heteroscedastic Errors

Let us consider the following linear model with n observations classifed into q groups and n_j observatons in the j-th group (j = 1, 2,...,q):

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_q \end{bmatrix} \beta + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_q \end{bmatrix} \tag{5.1}$$

where y_i and u_j are n_j x 1 and X_j is n_j x p. Again, we assume that $\lim_{n\to\infty} (n_j/n) > 0$ and

$$\begin{aligned} E\, u_j u'_k &= \theta_j^{-1} I_{nj} \quad \forall\ j = k = 1,\ldots,q \\ &= 0 \quad \forall\, j \neq k. \end{aligned} \tag{5.2}$$

Then we have

$$W = \begin{bmatrix} \theta_1 I_{n1} & 0 & \cdots & 0 \\ 0 & \theta_2 I_{n2} & \cdots & 0 \\ \vdots & \vdots & & \\ 0 & 0 & \cdots & \theta_q I_{nq} \end{bmatrix}$$

$$\gamma = \frac{1}{n} \sum_j \frac{n_j}{\theta_j}, \qquad D = \frac{1}{n} \sum_j X'_j X_j$$

$$D^* = \frac{1}{n} \sum_j \frac{1}{\theta_j} X'_j X_j,$$

$$\lambda_{jk} = \frac{2\, n\, \theta_j^2}{n_j} \quad \text{if } j = k$$

$$= 0 \quad \text{if } j \neq k$$

The expression (3.7) for the MSE matrix of b(a) becomes

$$E[b(a) - \beta]\, [b(a) - \beta]'$$

$$= (\Sigma X_j' X_j)^{-1} \Big[\Sigma \theta_j^{-1} X_j' X_j - \frac{a\gamma}{n\,\delta} \{2\, \Sigma \theta_j^{-1} X_j' X_j$$

$$- \frac{2}{\delta} (\Sigma X_j'X_j)\ \beta\, \beta' (\Sigma \theta_j^{-1} X_j' X_j) - \frac{2}{\delta} (\Sigma \theta_j^{-1} X_j'X_j)\, \beta\, \beta' (\Sigma \theta_j^{-1} X_j' X_j)$$

$$\beta\, \beta' (\Sigma X_j' X_j) - \frac{a\gamma}{n\delta} (\Sigma X_j' X_j)\ \beta\, \beta'\ (\Sigma X_j' X_j)\}\ \Big]\ (\Sigma X_j' X_j)^{-1} \qquad (5.3)$$

Further, the dominance condition (3.14) reduces to

$$0 \le a \le \left[\frac{\theta_s}{\theta_\ell} \quad \frac{tr\,(D^{-1}Q)}{\lambda_\ell\,(D^{-1}Q)} - 2\right];\ \frac{\theta_s}{\theta_\ell}\ \frac{tr\,(D^{-1}Q)}{\lambda_\ell\,(D^{-1}Q)} > 2, \qquad (5.4)$$

where θ_s = minimum of $(\theta_1,...,\theta_q)$; θ_ℓ = maximum of $(\theta_1,...,\theta_q)$.

Let us write $\eta_j = \sqrt{n}\,(\theta_j - \theta_s)$. Then $\eta_j \ge 0$ for all j = 1,2,..., q and Σ becomes a scalar matrix if $\eta_1 = \eta_2 = ... = \eta_q = 0$. Thus, we take $\theta_0 = \theta_s.\ell$ where ℓ is a q x 1 vector will all elements unity. Now, for all j, k = 1, 2,...,q we have

$$A_j\,(\theta_0) = \frac{1}{n}\, X_j'\, X_j$$

$$A_{jk}(\theta_0) = 0$$

$$A^*_{jk}(\theta_0) = \frac{1}{n}\, X_j'\, X_j \quad \text{if } j = k$$

$$= 0 \qquad \text{if } j \neq k,$$

$$\lambda_{jk} = \frac{2n\theta_s^2}{n_j} + 0(n^{-\frac{1}{2}}) \qquad \text{if } j = k$$

$$= 0 \qquad \text{if } j \neq k,$$

$$P_{jk}(\theta_0) = \frac{1}{n}\left[X_j' X_j - X_j' X_j (\Sigma X_\ell' X_\ell)^{-1} X_j' X_j\right] \text{ if } j = k$$

$$= -\frac{1}{n} (X_j' X_j) (\Sigma X_\ell' X_\ell)-1 X_k' X_k \quad \text{if } j \neq k.$$

Therefore, if θ follows the process (4.5), the difference between the risks of b(a) and $\hat{\beta}(a)$, to order $0(n^{-2})$, is given by

$R[b(a)] - R[\hat{\beta}(a)]$

$$= \frac{1}{n^3} \sum_j \text{tr} [D^{-1} \{X_j' X_j - \frac{1}{n} X_j' X_j D^{-1} X_j' X_j\} D^{-1} Q]$$

$$(\eta_j^2 - \frac{2 n \theta_s}{n_j}) - \frac{1}{n^4} \sum_{\substack{j.k \\ j \neq k}} \text{tr} [D^{-1} X_j' X_j D^{-1} X_k' X_k D^{-1} Q] \eta_j \eta_k. \quad (5.5)$$

The above expression is negative whenever $\eta_j^2 < (2n/n_j)$ for all j = 1,...,q. In other words, to the order of our approximation, b(a) has lower risk than $\hat{\beta}(a)$ whenever

$$\theta_j \leq \theta_s + \sqrt{2n/n_j} \;\forall\; j = 1,...,q. \quad (5.6)$$

If n_l = maximum of $(n_1,...,n_q)$, a sufficient conditon for the dominance of b(a) over $\hat{\beta}(a)$ is given by

$$\theta_l \leq \theta_s + \sqrt{2/n_\ell}. \quad (5.7)$$

5.2 AR(1) Errors:

Suppose u_t follows the AR(1) process

$$u_t = \theta u_{t-1} + \epsilon_t; \; t = 1,...,n, \quad (5.8)$$

where $|\theta| < 1$ and for all t = 1,2,...,n

$E(\epsilon_t) = 0$

$E(\epsilon_t \epsilon_{t+s}) = \psi$ if s = 0

$= 0$ if $s \neq 0$.

Then $E u u' = \sigma^2 \Omega^{-1}$ with $\sigma^2 = \psi/(1 - \theta^2)$

and

$$\Omega = \frac{1}{1-\theta^2} \begin{bmatrix} 1 & -\theta & 0 & \cdots & 0 & 0 \\ -\theta & 1+\theta^2 & -\theta & \cdots & 0 & 0 \\ 0 & -\theta & 1+\theta^2 & \cdots & 0 & 0 \\ \cdot & & & & & \\ \cdot & & & & & \\ \cdot & & & & & \\ 0 & 0 & 0 & \cdots & 1+\theta^2 & -\theta \\ 0 & 0 & 0 & \cdots & -\theta & 1 \end{bmatrix}$$

A consistent estimator of θ is given by

$$\hat{\theta} = \frac{\sum_{t=1}^{n-1} \hat{u}_t \hat{u}_{t+1}}{\sum_{t=1}^{n-1} \hat{u}_t^2},$$

where $\hat{u}_t$ is the t-th element of $\hat{u}$ y-Xb. Further

$$\Lambda = \lim_{n\to\infty} \left[n \, E(\hat{\theta} - \theta)^2\right]$$

$$= 1 - \theta^2.$$

Since $\Omega = I_n$ whenever $\theta = 0$, we take $\theta_0 = 0$ and $\eta = \sqrt{n}\,\theta$. Again

$$\Lambda = 1 - \theta^2$$

$$= 1 + 0(n^{-\frac{1}{2}})$$

$$\Omega_1(\theta_0) = \frac{\partial}{\partial\theta} W \Big| \, \theta = 0$$

$$= -B$$

$$P_{11}(\theta_0) = \frac{1}{n} X'B\,(I_n - \frac{1}{n} X\,D^{-1}\,X')\,BX,$$

where B is an n x n matrix with (j,k)-th element, 1 if $j = k \pm 1$ and 0 otherwise.

Hence, to order $0(n^{-2})$, the difference between the MSE matrices of b(a) and $\hat{\beta}(a)$ is given by

$$E[b(a) - \beta]\,[b(a) - b]' - E[\hat{\beta}(a) - \beta]\,[\hat{\beta}(a) - \beta']$$

$$= \frac{\sigma^2}{n} D^{-1}\,X'\,B(I_n - \frac{1}{n} X\,D^{-1}\,X')\,B\,X\,D^{-1}\,(\eta^2 - 1) \qquad (5.9)$$

Since the matrix $D^{-1} X' B (I_n - \frac{1}{n} X D^{-1} X') B X D^{-1}$ is positive semidefinite, b(a) dominates $\hat{\beta}(a)$ according to the MSE matrix criterion whenever $|\theta| < n^{-\frac{1}{2}}$.

REFERENCES

[1] Chaturvedi, A and Shukla, G. (1989): 'Stein-Rule Estimation in Linear Model with Nonscalar Error Covariance Matrix', Unpublished manuscript.

[2] Grubb, D and Magee, L (1988): 'A Variance Comparison of OLS and Feasible GLS Estimators', Econometric Theory, Vol. 4, p.329-335.

[3] Rothenberg, T.J. (1984): 'Approximate Normality of Generalized Least Squares Estimates', Econometrica, Vol 52, No. 4, p.811-825.