Department of Economics
UNIVERSITY OF CANTERBURY
CHRISTCHURCH, NEW ZEALAND

PRE-TESTING FOR LINEAR RESTRICTIONS IN A REGRESSION MODEL WITH SPHERICALLY SYMMETRIC DISTURBANCES

JUDITH A. GILES

Discussion Paper

No.8909

PRE-TESTING FOR LINEAR RESTRICTIONS IN A REGRESSION MODEL WITH SPHERICALLY SYMMETRIC DISTURBANCES

Judith A. Giles*
University of Canterbury

Abstract

In this paper we derive the exact risk (under quadratic loss) of pre-test estimators of the prediction vector and of the error variance of a linear regression model with spherically symmetric disturbances. The pre-test in question is one of the validity of a set of exact linear restrictions on the model's coefficient vector. We demonstrate how the known results for the model with normal disturbances can be extended to this broader case. Numerical evaluations of the risk expressions in the particular case of multivariate Student-t errors suggest that sampling properties of these pre-test estimators under these conditions are qualitatively similar to those which apply under normal errors.

*This work forms part of the author's Ph.D. research. My thanks to David Giles, Tony Rayner, Robin Carter, George Judge, John Knight, Peter Morgan, Kazuhiro Ohtani, Peter Phillips, Viren Srivastava, Mike Veall and Victoria Zinde-Walsh for their constructive comments and suggestions. I am also extremely grateful to Robert Davies for suggesting the use of his algorithm AS155 and for providing his software.

Correspondence:

Judith A. Giles
Department of Economics
University of Canterbury,
Private Bag,
Christchurch 1, N.Z.

1. Introduction

In a linear regression model, suppose that the process generating a (T×1) vector of observations on a dependent variable y is

$$y = X\beta + e \ , \tag{1}$$

where X is a (T×k) full rank matrix of non-stochastic variables and β is a (k×1) vector of unknown parameters. We assume that the (T×1) vector of regression disturbances e is distributed according to the laws of the class of spherically symmetric distributions which can be expressed as a variance mixture of normals.[1] Further, assume that the probability density function (pdf) of e exists and that $E(e) = 0$ and $E(ee') = \sigma_e^2 I_T$. We write $e \sim SSD_N(0, I_T)$.[2]

Consider also m independent linear restrictions on β, summarised by the hypotheses

$$H_0: \ R\beta = r \quad \text{vs.} \quad H_1: \ R\beta \neq r \tag{2}$$

where R and r are (m×k) and (m×1) matrices of known constants and rank (R) = m (<k). The usual statistic for testing the linear restrictions (2) is

$$F = \frac{v(Rb-r)'[RS^{-1}R']^{-1}(Rb-r)}{m(y-Xb)'(y-Xb)} \ , \tag{3}$$

where $v = (T-k)$, $S = (X'X)$, $b = S^{-1}X'y$ is the unrestricted least squares estimator of β, and $\tilde{\sigma}_e^2 = (y-Xb)'(y-Xb)/v$ is the unrestricted unbiased estimator of σ_e^2. The restricted least squares estimator of β which imposes H_0 is $b^* = b-S^{-1}R'[RS^{-1}R']^{-1}(Rb-r)$ and the corresponding estimator of the error variance is $\sigma_e^{*2} = (y-Xb^*)'(y-Xb^*)/(v+m)$. σ_e^{*2} is unbiased only when H_0 is true.[3]

We are usually uncertain of the validity of the prior information, so the common procedure is to (pre-)test H_0 prior to estimating the parameters of the model. This results in pre-test estimators of β and σ_e^2, say $\hat{b}$ and $\hat{\sigma}_e^2$ respectively. The sampling properties of these estimators of the parameters

of the linear regression model, after a pre-test for linear restrictions on the coefficient vector, have been widely examined (see, for example, Brook (1972), Wallace (1977), Judge and Bock (1978), Ohtani (1988), Clarke, Giles and Wallace (1987a,b), and Gelfand and Dey (1988)). All of these studies assume that the regression disturbances are normally distributed, and it is this assumption which is generalised here.

There is a large body of literature suggesting that some economic data series may be generated by processes whose error distributions have fat tails, or even infinite variances. Examples include price-change analysis in the stock, financial and commodity markets (Fama (1963, 1965), Sharpe (1971), Praetz (1972), Blattberg and Gonedes (1974) and Sutradhar and Ali (1986)); cash flow analysis (Granger and Orr (1972)); and demand analysis (Coursey and Nyquist (1988)). The possibility of non-normal regression disturbances has led to searches for robust estimators, resulting in such estimators as the M-, L-, and R-estimators. See, for instance, Huber (1981), Koenker (1982), Hampel et al. (1983) and Judge et al. (1985).

There have also been many studies of the robustness of traditional estimators. In particular, these studies show that the least squares estimator is sensitive to the form of the underlying distribution, because it minimises squared deviations and so, gives a relatively heavy weight to the tails of the distribution.[4] Various alternative distributions to normality have been investigated. One that has received considerable attention in the literature is the spherically symmetric family of distributions (and its parent distribution, the elliptically symmetric family). Well known members are the normal and the multivariate Student-t distributions; the latter includes the Cauchy distribution.

The T-dimensional random vector e is said to have a (multivariate) spherically symmetric distribution (SSD) if e and He have the same

distribution for all (T×T) orthogonal matrices H. Hence, its distribution is independent of direction from the origin and is a function only of distance from the origin: that is, $r = (e'e)^{1/2}$. So, the joint pdf of e is of the form

$$f(e) = \phi(e'e) \tag{4}$$

with respect to the Lebesgue measure on R^T, where ϕ: $[0,\infty)\to[0,\infty)$ and $\int_0^\infty r^{T-1}\phi(r^2)dr = (1/2)\Gamma(T/2)\pi^{-T/2}$. All non-normal SSD's have components which are dependent but are uncorrelated: the normal distribution is the only spherically symmetric law for which the observations are independent. Discussions of this family of distributions are given by Kelker (1970), Devlin et al. (1976), King (1979), Chmielewski (1981), and Muirhead (1982), for instance.

This is a sensible extension of spherical normality to investigate as first, it is a class of density functions whose contours of equal density have the same spherical shape as the spherical normal; secondly we can generate members which have fat and thin tails relative to those of the normal; thirdly, all marginal and conditional densities of a spherical random vector are also spherically distributed and have the same shape; and finally, a subclass of the spherically symmetric family, say SSD_N, can be written in terms of a variance mixture of normal distributions:

$$f(e) = \int_0^\infty f_N(e)f(\tau)d\tau \tag{5}$$

where $f_N(e)$ is the pdf of e when $e \sim N(0,\tau^2 I_T)$ and $f(\tau)$ is the pdf of τ which is supported on $[0,\infty)$. In this case $\sigma_e^2 = E(\tau^2)$.

So, we may have non-normal regression disturbances even if each e_i $(i=1,\ldots,T)$ is normally distributed when the variance of e_i is itself a

random variable. In this paper we consider the family of SSD's which can be expressed in the form (5).

For instance, if $f(\tau)$ is an inverted gamma density with, say, scale parameter σ^2 and degrees of freedom parameter ν, then (5) is the pdf of a multivariate Student-t (Mt) distribution. That is,

$$f(e) = \frac{c_\nu}{\sigma^T}\left[1 + \frac{\nu^{-1}e'e}{\sigma^2}\right]^{-(T+\nu)/2}, \quad \sigma > 0,\ \nu > 0,\ -\infty < e_i < \infty. \qquad (6)$$

$c_\nu = \Gamma(T+\nu/2)[(\pi\nu)^{T/2}\Gamma(\nu/2)]^{-1}$ is the normalising constant and $\sigma_e^2 = \nu\sigma^2/(\nu-2)$ is the common variance of the e_i's, $i = 1,\ldots,T$, $\nu > 2$. The marginal distributions are univariate Student-t and for small values of ν they have thicker tails than under normality; as $\nu\to\infty$ the pdf approaches a normal form; and when $\nu = 1$, the pdf is Cauchy.[5]

Many studies have investigated linear regression models with spherically symmetric disturbances including Box (1952), Thomas (1970), Zellner (1976), King (1979), Ullah and Zinde-Walsh (1984), Judge, Miyazaki and Yancey (1985), Ullah and Phillips (1986), Sutradhar and Ali (1986), Singh (1988), Sutradhar (1988).[6] Box (1952) notes that the F-ratio (3), under H_0, is central $F_{(m,v)}$ for all SSD's. The non-null distribution, however, depends on the specific form of the SSD. This is shown by Thomas (1970). We provide an alternative derivation of the non-null distribution for SSD_N errors in Section 2. Unaware of Thomas's work, Ullah and Phillips (1986) and Sutradhar (1988), assuming Mt errors, also derive the non-null distribution of F.

Thomas also proves that the usual least squares estimator of β is the linear minimum variance unbiased estimator and the maximum likelihood estimator of β. See also Zellner (1976). King (1979) extends many of Thomas's results. In particular, he shows that if a test has an optimal power property for normal disturbances over all possible values of τ^2 then

it maintains this property when the errors are SSD_N. Consequently, the F-ratio given by (3) is a UMPI size-α test for SSD_N regression disturbances.

Judge et al. (1985) establish sampling properties of the James-Stein estimator of the location parameter vector (and its positive part counterpart) under a squared error loss measure and a Mt error density. They compare, via a Monte Carlo experiment, the finite sample behaviour (empirical risks) for their Stein-like, and some conventional robust, estimators. In general, the risk characteristics are found to be the same as for the normal errors case. However, there are no analytical results relating to the finite-sample properties of pre-test estimators when the model's disturbances are non-normally distributed. Accordingly, in this paper we derive the risk, under squared error loss, of the usual pre-test estimators of the prediction vector[7] and of the error variance of model (1) after a pre-test of H_0 when the regression disturbances are $SSD_N(0, I_T)$. These risk functions depend on the form of $f(\tau)$ and so to illustrate the results we numerically evaluate them for the important case of Mt errors. This enables us to investigate how departures from normality, as represented by the value of ν, affect the risk functions of the estimators. In the next section we derive the non-null distribution of F for the general case of SSD_N errors. Sections 3 and 4 present, discuss and evaluate the finite sample risk functions of the various estimators of the prediction vector and of the error variance respectively. Some concluding remarks are given in the final section, and the proofs of the theorems appear in an appendix.

2. The non-null distribution of F

To determine the properties of the pre-test estimators, we need knowledge of the distribution of the test statistic under the alternative hypothesis.

Theorem 1. If $e \sim SSD_N(0,I_T)$ then,

$$f(F) = \sum_{r=0}^{\infty} \frac{\theta^r m^{\frac{m}{2}+r} v^{\frac{v}{2}} F^{\frac{m}{2}+r-1}}{r!(v+mF)^{\frac{m+v}{2}+r} B\left(\frac{m}{2}+r;\frac{v}{2}\right)} \int_0^{\infty} e^{-\theta/\tau^2}(\tau^2)^{-r} f(\tau)d\tau, \qquad (7)$$

where $\theta = (R\beta-r)'[RS^{-1}R']^{-1}[R\beta-r)/2$, and $B(.;.)$ is the usual Beta function.

Proof. See the appendix.

Corollary 1. If $e \sim Mt\left(0,\sigma^2\nu/(\nu-2)I_T\right)$ then,

$$f_{Mt}(F) = \sum_{r=0}^{\infty} \frac{(2\lambda/\nu)^r \Gamma\left(\frac{\nu}{2}+r\right) m^{\frac{m}{2}+r} v^{\frac{v}{2}} F^{\frac{m}{2}+r-1}}{r!(1+2\lambda/\nu)^{\nu/2+r} B\left(\frac{m}{2}+r,\frac{v}{2}\right)\Gamma\left(\frac{\nu}{2}\right)(v+mF)^{\frac{m+v}{2}+r}}, \qquad (8)$$

where $\lambda = \theta/\sigma^2$.

Proof. See the appendix.

We note the following points:

(i) $f_{Mt}(F)$ is equivalent to the expressions derived by Ullah and Phillips (1986) and Sutradhar (1988).

(ii) (7) and (8) both collapse to a non-central F pdf, $F'_{(m,v;\lambda)}$, with m and v degrees of freedom and non-centrality parameter λ, when $e \sim N(0,\sigma^2 I_T)$.

(iii) When H_0 is true $F \sim F_{(m,v)}$. (This information is also used in analysing the pre-test estimators' sampling properties.)

3. The risk functions of alternative estimators of E(y).

We define the pre-test estimator for $E(y) = X\beta$ as

$$X\hat{b} = \begin{cases} Xb & ; \text{ if } F > c \\ Xb^* & ; \text{ if } F \le c \end{cases}$$

$$= Xb I_{(c,\infty)}(F) + Xb^* I_{[0,c]}(F) \tag{9}$$

where $c = c(\alpha)$ satisfies $\int_0^c dF_{(m,v)} = (1-\alpha)$ and $I_{[a,b]}(F)$ is an indicator function with value unity if $F \in [a,b]$, zero otherwise. Then, if $X\bar{b}$ is any estimator of $E(y)$ its risk function, under quadratic loss, is

$$\rho\left[E(y),X\bar{b}\right] = E\left[X\bar{b}-E(y)\right]'\left[X\bar{b}-E(y)\right] , \tag{10}$$

which is the trace of the mean squared error matrix of $X\bar{b}$. We now derive the risk expressions for the various estimators of $E(y)$.

Theorem 2. If $e \sim SSD_N(0,I_T)$ then,

$$\rho\left[E(y),Xb\right] = kE(\tau^2) \tag{11}$$

$$\rho\left[E(y),Xb^*\right] = (k-m)E(\tau^2) + 2\theta \tag{12}$$

$$\rho\left[E(y),X\hat{b}\right] = kE(\tau^2) + 2\theta \int_0^\infty \left(2P^\tau_{20}-P^\tau_{40}\right)f(\tau)d\tau$$

$$- m \int_0^\infty \tau^2 P^\tau_{20} f(\tau)d\tau , \tag{13}$$

where $P^\tau_{ij} = \Pr.\left[F'_{(m+i,v+j;\lambda_\tau)} \le \left(cm(v+j)\right)/\left(v(m+i)\right)\right]$ and $\lambda_\tau = \theta/\tau^2$, $i,j = 0,1,2,\ldots$.

Proof. See the appendix.

Corollary 2. If $e \sim Mt\left(0,\nu\sigma^2/(\nu-2)I_T\right)$ then, for $\nu > 2$,

$$\rho_{Mt}\left[E(y),Xb\right] = \sigma^2 k\nu/(\nu-2) \tag{14}$$

$$\rho_{Mt}\left[E(y),Xb^*\right] = \sigma^2\left[(k-m)\nu+2\lambda(\nu-2)\right]/(\nu-2) \tag{15}$$

$$\rho_{Mt}\left[E(y),X\hat{b}\right] = \sigma^2\left[k\nu-m\nu P_{201}+2\lambda(\nu-2)(2P_{202}-P_{402})\right]/(\nu-2) , \tag{16}$$

where

$$P_{ijn} = \sum_{r=0}^{\infty} \frac{(2\lambda/\nu)^r \Gamma\left(\frac{\nu}{2}+r+n-2\right)}{r!(1+2\lambda/\nu)^{\frac{\nu}{2}+r+n-2} \Gamma\left(\frac{\nu}{2}+n-2\right)} . I_x\left[\tfrac{1}{2}(m+i)+r;\tfrac{1}{2}(v+j)\right] ,$$

$i,j,n = 0,1,2,\ldots,$, and $I_x(.;.)$ is Pearson's incomplete beta function with $x = cm/(v+cm)$.

Proof. See the appendix.

Comparing (11), (12) and (13) we note that each depends on $f(\tau)$. In addition,

(i) If $e \sim N(0,\tau^2 I_T)$ then (11), (12) and (13) reduce to the well known expressions

$$\rho_N\left[E(y),Xb\right] = k\sigma^2 \tag{17}$$

$$\rho_N\left[E(y),Xb^*\right] = \sigma^2\left[(k-m)+2\lambda\right] \tag{18}$$

$$\rho_N\left[E(y),X\hat{b}\right] = \sigma^2\left[k+(4\lambda-m)P_{20}-2\lambda P_{40}\right], \tag{19}$$

where $P_{ij} = Pr.\left[F'_{(m+i,v+j;\lambda)} \leq \left(cm(v+j)\right)/\left(v(m+i)\right)\right]$.

(See, for example, Wallace (1977) and Judge and Bock (1978)).

(17), (18) and (19) can also be obtained directly from (14), (15) and (16), respectively, as $e \sim N(0,\sigma^2 I_T)$ when $\nu = \infty$. In this case, $P_{ijn} = P_{ij}$.

(ii) When $\alpha \to 1(0)$, $c \to 0(\infty)$, $P^{\tau}_{ij} \to 0(1)$ for all i,j and $\rho\left[E(y),X\hat{b}\right] \to \rho\left[E(y),Xb\right]\left(\rho\left[E(y),Xb^*\right]\right)$.

(iii) When the null hypothesis is true ($\theta = 0$),

$$\rho_0\left[E(y),Xb^*\right] = (k-m)E(\tau^2) < \rho_0\left[E(y),X\hat{b}\right] = kE(\tau^2)$$

$$- m \int_0^{\infty} \tau^2 P^*_{20} f(\tau) d\tau < \rho_0\left[E(y), Xb\right] = kE(\tau^2) ,$$

where $P^*_{ij} = \Pr.\left[F_{(m+i,v+j)} \leq \left(cm(v+j)\right) / \left(v(m+i)\right)\right]$.

(iv) When $\theta \to \infty$, the risk of Xb^* is unbounded, while

$$\rho\left[E(y), X\hat{b}\right] \to \rho\left[E(y), Xb\right] \text{ as } P^{\tau}_{ij} \to 0 \text{ for all } i,j.$$

(v) $\rho\left[E(y), Xb\right] = \rho\left[E(y), Xb^*\right]$ when $\theta = mE(\tau^2)/2 = \theta^*$. If $e \sim N(0, \sigma^2 I_T)$ then $\theta^*_N = m\sigma^2/2$, as is well documented in the literature; while if $e \sim Mt\left(0, \nu\sigma^2/(\nu-2) I_T\right)$ then $E(\tau^2) = \nu\sigma^2/(\nu-2) > \sigma^2$ and $\theta^*_{Mt} = m\nu\sigma^2/\left(2(\nu-2)\right) > \theta^*_N$.

So, if we assume normality when in fact the distribution of the errors belongs to the wider class of SSD_N, there is a range of θ over which we would choose the incorrect estimator. For example, if $E(\tau^2) > \sigma^2$ (that is, the marginal distribution of e has fatter tails than under normality) then we should select Xb^* for $\theta < \theta^*$ to minimize risk but if we assume normality then we would incorrectly choose Xb for $\theta \in (\theta^*_N, \theta^*)$.

It is difficult to discuss further features of the risk functions without numerically evaluating them. Hence, to illustrate the results we assume Mt errors and evaluate the risk expressions (14),(15), and (16) for various choices of ν, α, m, k and T as functions of λ. We consider risk relative to σ^2 and parameterise with respect to λ rather than θ to eliminate the scale parameter σ^2.[8] So, the relative risk of an estimator $X\bar{b}$ of E(y) is $R\left[E(y), X\bar{b}\right] = \rho\left[E(y), X\bar{b}\right]/\sigma^2$. Some representative results, for various ν values, appear in Figures 1 to 4.

Comparing the figures, we see that a decrease in the value of ν from the normal errors case ($\nu = \infty$) causes an upward shift of the estimator risk functions, a decrease in the rate at which the risk of the pre-test

estimator approaches that of the unrestricted estimator, and an increase in the risk gain of the restricted estimator over the unrestricted estimator for all λ such that $R\left[E(y),Xb^*\right] < R\left[E(y),Xb\right]$. For the unrestricted and the restricted estimators these changes occur because of the increase in the estimators' variances as ν decreases (the marginal distribution has fatter tails). For the pre-test estimator, the increase in its variance and its absolute bias (for relatively large λ) both contribute to the observed differences. Our numerical evaluations suggest that, in general, the difference between an estimator's risk under normality and Mt errors is relatively insignificant for a ν value of at least 100.

Comparing the risk functions of the pre-test estimator and its component estimators for a given ν we find that the conclusions observed when the errors are normally distributed hold for all values of ν. When the null hypothesis is true the pre-test estimator is risk inferior to the restricted estimator but superior to the unrestricted estimator. However, the pre-test estimator is dominated by the unrestricted estimator over a wide range of λ, and by both of its component estimators over part of the parameter space. Hence, no one of the estimators strictly dominates the other two. This latter feature suggests choosing an estimator according to some optimality criterion. Such a study is beyond the scope of this paper (see, for example, Brook (1972), Ohtani and Toyoda (1978), Ohtani (1988), FToyoda and Wallace (1975)). So, aside from appropriate scaling, we find that the risk characteristics of the estimators are similar for all choices of ν.

4. The risk functions of alternative estimators of σ_e^2.

Let the pre-test estimator of σ_e^2 be given by

$$\hat{\sigma}_e^2 = \begin{cases} \tilde{\sigma}_e^2 & ; \text{ if } F > c \\ \sigma_e^{*2} & ; \text{ if } F \leq c \end{cases}$$

$$= \tilde{\sigma}_e^2 I_{(c,\infty)}(F) + \sigma_e^{*2} I_{[0,c]}(F)$$

where c, F and $I_{[a,b]}(F)$ are as previously defined. The risk, under quadratic loss, of any estimator $\bar{\sigma}_e^2$ of σ_e^2 is

$$\rho\left(\sigma_e^2, \bar{\sigma}_e^2\right) = E\left(\bar{\sigma}_e^2 - \sigma_e^2\right)^2 ,$$

and is the mean squared error of $\bar{\sigma}_e^2$. We now consider the risks of the various estimators of σ_e^2.

Theorem 3. If $e \sim SSD_N(0, I_T)$ then,

$$\rho\left(\sigma_e^2, \tilde{\sigma}_e^2\right) = \left[(v+2)E(\tau^4) - v\left(E(\tau^2)\right)^2\right]/v \tag{20}$$

$$\rho\left(\sigma_e^2, \sigma_e^{*2}\right) = \left[(v+m)(v+m+2)E(\tau^4) - (v+m)^2\left(E(\tau^2)\right)^2 + 2\theta\left(\theta + 2E(\tau^2)\right)\right]/(v+m)^2 \tag{21}$$

$$\begin{aligned}\rho\left(\sigma_e^2, \hat{\sigma}_e^2\right) = \Bigg\{&(v+m)^2\left[(v+2)E(\tau^4) - v\left(E(\tau^2)\right)^2\right] \\ &+ \int_0^\infty \Big[-m(v+2)(2v+m)\tau^4 P_{04}^\tau + 2mv(v+m)E(\tau^2)\tau^2 P_{02}^\tau \\ &+ v\left[m(m+2)\tau^4 P_{40}^\tau + 4(m+2)\theta\tau^2 P_{60}^\tau + 4\theta^2 P_{80}^\tau\right] + 2v^2\tau^2(m\tau^2 P_{22}^\tau + 2\theta P_{42}^\tau) \\ &- 2v(v+m)E(\tau^2)(m\tau^2 P_{20}^\tau + 2\theta P_{40}^\tau)\Big] f(\tau)d\tau\Bigg\} \Big/ \left[v(v+m)^2\right].\end{aligned} \tag{22}$$

Proof. See the appendix.

Corollary 3. If $e \sim Mt(0, \sigma_e^2 I_T)$, $\sigma_e^2 = \nu\sigma^2/(\nu-2)$, $\nu > 4$, then

$$\rho_{Mt}\left(\sigma_e^2, \tilde{\sigma}_e^2\right) = 2\sigma_e^4(v+\nu-2)/\left(v(\nu-4)\right) \tag{23}$$

$$\rho_{Mt}\left(\sigma_e^2, \sigma_e^{*2}\right) = 2\sigma_e^4\left[\nu^2(v+m)(v+m+2) + 2\lambda(\nu-2)(\nu-4)\left[(\nu-2)\lambda + 2\nu\right]\right] / \left[\nu^2(\nu-4)(v+m)^2\right] \tag{24}$$

$$\rho_{Mt}\left(\sigma_e^2,\hat{\sigma}_e^2\right) = \sigma_e^4\Big\{2\nu^2(v+m)^2(v+\nu-2) + m\nu^2\Big[-(v+2)(m+2v)(\nu-2)P_{040}$$

$$+ v(\nu-2)(m+2)P_{400} + 2v^2(\nu-2)P_{220}+2v(v+m)(\nu-4)(P_{021}-P_{201})\Big]$$

$$+ 4\lambda\nu v(\nu-2)(\nu-4)\Big[(m+2)P_{601}-(v+m)P_{402} + vP_{421}\Big]$$

$$+ 4\lambda^2 v(\nu-2)^2(\nu-4)P_{802}\Big\}\Big/\Big[\nu^2(\nu-4)v(v+m)^2\Big] \quad . \tag{25}$$

Proof. See the appendix.

Each of the risk functions in Theorem 3 depends on $f(\tau)$. We also note, from comparing these risk functions, that:

(i) If $e \sim N(0,\sigma^2 I_T)$ then (20), (21) and (22) collapse to the expressions derived by Clarke et al. (1987b):

$$\rho_N(\sigma^2,\tilde{\sigma}^2) = 2\sigma^4/v \quad . \tag{26}$$

$$\rho_N(\sigma^2,\sigma^{*2}) = 2\sigma^4(2\lambda^2+4\lambda+v+m)/(v+m)^2 \tag{27}$$

$$\rho_N(\sigma^2,\hat{\sigma}^2) = \sigma^4\Big\{4v\lambda^2P_{80} + 4v\lambda\Big[(m+2)P_{60}+vP_{42}-(v+m)P_{40}\Big]$$

$$+ 2(v+m)^2-2mv(v+m)(P_{20}-P_{02})-m(v+2)(m+2v)P_{04}+2mv^2P_{22}$$

$$+ mv(m+2)P_{40}\Big\}\Big/\Big[v(v+m)^2\Big] \quad . \tag{28}$$

(26), (27) and (28) also follow, respectively, from (23), (24) and (25) as $e \sim N(0,\sigma^2 I_T)$ when $\nu = \infty$. In this case, $P_{ijn} = P_{ij}$.

(ii) $\rho\left(\sigma_e^2,\hat{\sigma}_e^2\right)$ converges to $\rho\left(\sigma_e^2,\tilde{\sigma}_e^2\right)$ when $\alpha \to 1$, and to $\rho\left(\sigma_e^2,\sigma_e^{*2}\right)$ when $\alpha \to 0$.

(iii) $\rho\left(\sigma_e^2,\hat{\sigma}_e^2\right) \to \rho\left(\sigma_e^2,\tilde{\sigma}_e^2\right)$ as $\theta \to \infty$, while $\rho\left(\sigma_e^2,\sigma_e^{*2}\right)$ is unbounded.

(iv) $\rho\left(\sigma_e^2,\sigma_e^{*2}\right) < \rho\left(\sigma_e^2,\tilde{\sigma}_e^2\right)$ when the restrictions are valid.

To illustrate the results we have numerically evaluated, relative to σ_e^4, the risk expressions $\rho_{Mt}\left(\sigma_e^2,\tilde{\sigma}_e^2\right)$, $\rho_{Mt}\left(\sigma_e^2,\sigma_e^{*2}\right)$ and $\rho_{Mt}\left(\sigma_e^2,\hat{\sigma}_e^2\right)$ as functions of λ, for the values of ν, α, m and v as before. So the relative risk of an estimator $\bar{\sigma}_e^2$ of σ_e^2 is $R\left(\sigma_e^2,\bar{\sigma}_e^2\right) = \rho\left(\sigma_e^2,\bar{\sigma}_e^2\right)/\sigma_e^4$. Figures 5 to 8 depict typical cases. We note that Figure 6, which considers the risk functions when $\nu = 5$, is drawn on a different scale from that of Figures 5, 7 and 8, to enable the features for all four cases to be distinguishable.

Consider first Figure 5, which illustrates the risk functions of the estimators when the errors are normally distributed. We see that there exists a family of pre-test estimators, with $c \in (0,1]$, which strictly dominate the unrestricted estimator for all λ, and dominate the restricted estimator over part of this parameter space.[9] This feature is not observed in the evaluations undertaken by Clarke et al. (1987b) but is noted in subsequent work by Ohtani (1988).[10]

Turning to the consequences of decreasing the value of ν from infinity, we find that the risk functions change in a similar way to that observed when estimating the prediction vector. That is, the estimator risk functions shift upwards, there is a decrease in the rate at which the risk of the pre-test estimator approaches that of the unrestricted estimator, and there is an increase in the risk gain of the restricted estimator over the unrestricted estimator for all λ such that $R\left(\sigma_e^2,\sigma_e^{*2}\right) < R\left(\sigma_e^2,\tilde{\sigma}_e^2\right)$. These effects occur because of the increase in the variance of all of the estimators when ν decreases and changes in the bias functions of the restricted and pre-test estimators. The bias of the restricted estimator decreases for all λ while the bias function of the pre-test estimator shifts down for relatively small λ (which may increase or decrease absolute bias) but becomes unbiased at a slower rate.

In Section 3, when comparing the risk functions of an estimator of the prediction vector for different values of ν, we found that, in general, there was little difference between the normal risk function and those for $\nu = 100$. However, when estimating the error variance we find that the differences may be up to 20% for this ν value and only become relatively insignificant for much larger values of ν, say 10,000.

Nevertheless, when comparing the risk functions of the pre-test estimator and its component estimators for a given ν, the conclusions for when the errors are normally distributed continue to hold for all values of ν. Namely, there exists a family of pre-test estimators, with $c \in (0,1]$ which strictly dominate the unrestricted estimator for all λ. Further, some members of this family,[11] for some ν, also strictly dominate the restricted estimator. The numerical evaluations suggest that the pre-test estimator with critical value 1 strictly dominates all other members of this family.[12]

Moreover, for some ν, the restricted estimator is also strictly dominated by those pre-test estimators with $1 < c < \infty$. Comparing equations (24) and (25) this will depend on m and v as well as ν. For the cases analyzed we find, in general, that the restricted estimator is strictly dominated by all pre-test estimators, except for those with c around 0 and $c = \infty$, if ν is at most 15.

These numerical evaluations show that, apart from appropriate scaling, the risk properties of the estimators are qualitatively similar for all values of ν. Regardless of the value of ν, our recommendation is to pre-test rather than to impose the restrictions without testing their validity, particularly if ν is believed to be relatively small. Further, when using the least squares component estimators, a critical value of one seems to be the appropriate choice for the pre-test.

5. Conclusions

In this paper we have considered the sampling properties of various estimators of the linear regression model, after a preliminary test of restrictions on the coefficients, when the regression disturbances are spherically symmetric. The risk expressions were numerically evaluated for the case of Mt errors, and these results suggested that, qualitatively, the estimator properties are similar for all values of ν.

Our analysis assumes that the researcher estimates the parameters of the model separately whereas one usually wishes to simultaneously estimate $X\beta$ and σ_e^2. This suggests considering a joint risk function for $X\beta$ and σ_e^2. This problem remains to be investigated, not only for our choice of error distribution, but generally in the pre-test literature.

Appendix

Proof of Theorem 1.

We note equation (5) and so,

$$f(F) = \int_0^\infty f_N(F) f(\tau) d\tau \ , \tag{A.1}$$

where $f_N(F)$ is the joint density function of F when $e \sim N(0,\tau^2 I_T)$, which is $F'_{(m,v;\lambda_\tau)}$. Using this and (A.1) gives the result directly. #

Proof of Corollary 1.

To obtain Corollary 1 from Theorem 1 let $f(\tau)$ be an inverted gamma density function with scale parameter σ^2 and degrees of freedom parameter ν. Then,

$$f(\tau) = \left[2/\Gamma(\nu/2)\right]\left(\nu\sigma^2/2\right)^{\nu/2}\tau^{-(\nu+1)}e^{-\nu\sigma^2/2\tau^2} \tag{A.2}$$

and so,

$$f_{Mt}(F) = \sum_{r=0}^{\infty} \frac{\theta^r m^{\frac{m}{2}+r} v^{\frac{v}{2}} F^{\frac{m}{2}+r-1}}{r!(v+mF)^{\frac{m+v}{2}+r} B\left(\frac{m}{2}+r;\frac{v}{2}\right)\Gamma\left(\frac{\nu}{2}\right)} \int_0^\infty e^{-[2\theta+\nu\sigma^2]/2\tau^2}(\tau^2)^{-(r+\nu/2+1/2)}d\tau, \tag{A.3}$$

where B(.;.) is the usual Beta function.

Let $\tau^2 = 1/Z$ so that the integral in (A.3) becomes

$$\frac{1}{2}\int_0^\infty e^{-[2\theta+\nu\sigma^2]Z/2} Z^{r+\nu/2-1} dZ$$

$$= \frac{1}{2}\left[\frac{2}{2\theta+\nu\sigma^2}\right]^{r+\nu/2} \int_0^\infty e^{-t} t^{r+\nu/2-1} dt$$

with the change of variable $t = [2\theta+\nu\sigma^2]Z/2$.

Now, $\int_0^\infty e^{-t} t^{f-1} dt = \Gamma(f)$ so (A.3) becomes

$$f_{Mt}(F) = \sum_{r=0}^{\infty} \frac{\theta^r m^{\frac{m}{2}+r} v^{\frac{v}{2}} F^{\frac{m}{2}+r-1}}{r!(v+mF)^{\frac{m+v}{2}+r} B\left(\frac{m}{2}+r;\frac{v}{2}\right)} \cdot \left(\frac{\nu\sigma^2}{2}\right)^{\nu/2} \cdot \frac{\Gamma\left(\frac{\nu}{2}+r\right)}{\Gamma\left(\frac{\nu}{2}\right)} \cdot \left[\frac{2}{2\theta+\nu\sigma^2}\right]^{r+\nu/2} .$$

Collecting terms and allowing for the change from θ to λ completes the proof. #

Proof of Theorem 2.

First,

$$\rho\left[E(y),Xb\right] = tr\left[v(Xb)\right]$$

$$= kE(\tau^2) ,$$

as $E(ee') = E(\tau^2)I_T$, and b is unbiased.

Secondly,

$$\rho\left[E(y),Xb^*\right] = E\left(Xb^*-E(y)\right)'\left(Xb^*-E(y)\right)$$

$$= E\Big(e'(XS^{-1}X'-XS^{-1}R'[RS^{-1}R']^{-1}RS^{-1}X')e$$

$$+ (R\beta-r)'[RS^{-1}R']^{-1}(r\beta-r)\Big)$$

$$= (k-m)E(\tau^2) + 2\theta ,$$

as $Xb^* - E(y) = XS^{-1}X'e - XS^{-1}R'[RS^{-1}R']^{-1}\left[(R\beta-r)+RS^{-1}X'e\right]$, and $\theta = (R\beta-r)'[RS^{-1}R']^{-1}(R\beta-r)/2$.

Finally, to derive the risk of $X\hat{b}$ we have that

$$X\hat{b} = Xb - XS^{-1}R'[RS^{-1}R']^{-1}(Rb-r)I_{[0,c]}(F) ,$$

so

$$\rho\left[E(y),X\hat{b}\right] = \rho\left[E(y),Xb\right] + E\Big\{\Big[2(R\beta-r)'[RS^{-1}R']^{-1}(Rb-r)$$

$$- (Rb-r)'[RS^{-1}R']^{-1}(Rb-r)\Big]I_{[0,c]}(F)\Big\}$$

$$= \rho\left[E(y),Xb\right] + E\left\{G\right\} . \qquad (A.4)$$

Now, let $E_N\left\{G\right\} = E\left\{G\right\}$ when $e \sim N(0,\tau^2 I_T)$ so that

$$E\left\{G\right\} = \int_0^\infty E_N\left\{G\right\}(\tau)d\tau \ .$$

Then, using the results of, for instance, Judge and Bock (1978) we have

$$E_N\left\{G\right\} = \tau^2\left[2\lambda_\tau\left(2P^\tau_{20}-P^\tau_{40}\right) - mP^\tau_{20}\right] \ , \tag{A.5}$$

where $\lambda_\tau = (R\beta-r)'[RS^{-1}R']^{-1}(r\beta-r)/2\tau^2$. Substituting (A.5) into (A.4), and noting that $\theta = \lambda_\tau/\tau^2$, yields the expression $\rho\left[E(y),X\hat{b}\right]$. #

To establish some of the remaining results we require the following lemma:

Lemma A1. If $\tau \sim IG(\sigma^2,\nu)$ then $e \sim Mt(0,\nu\sigma^2/(\nu-2)I_T)$ and

$$\int_0^\infty (\tau^2)^h P^\tau_{ij} f(\tau)d\tau = \left(\frac{\nu\sigma^2}{2}\right)^h \sum_{r=0}^\infty \frac{(2\lambda/\nu)^r \Gamma\left(\frac{\nu}{2}+r-h\right)}{r!(1+2\lambda/\nu)^{\nu/2+r-h}\Gamma\left(\frac{\nu}{2}\right)}$$

$$\times\ I_x\left[\tfrac{1}{2}(m+i)+r;\tfrac{1}{2}(v+j)\right] \ , \tag{A.6}$$

$h,i,j = 0,1,2,\ldots$.

Proof.

If $\tau \sim IG(\sigma^2,\nu)$ then

$$\int_0^\infty (\tau^2)^h P^\tau_{ij} f(\tau)d\tau = \int_0^\infty (\tau^2)^h \sum_{r=0}^\infty e^{-\lambda_\tau}\left(\lambda_\tau^r/r!\right)(r!)^{-1}$$

$$\cdot\ I_x\left[\tfrac{1}{2}(m+i)+r;\tfrac{1}{2}(v+j)\right]\left(\Gamma\left(\frac{\nu}{2}\right)\right)^{-1}2(\nu\sigma^2/2)^{\nu/2}$$

$$\cdot\ \tau^{-(\nu+1)}e^{-\nu\sigma^2/2\tau^2}d\tau$$

$$= \sum_{r=0}^{\infty} \theta^r I_x\left[\tfrac{1}{2}(m+i)+r;\tfrac{1}{2}(v+j)\right] 2(\nu\sigma^2/2)^{\nu/2}\left(r!\Gamma\left(\frac{\nu}{2}\right)\right)^{-1}$$

$$\cdot \int_0^{\infty} e^{(2\theta+\nu\sigma^2)/2\tau^2}(\tau^2)^{-(r-h+\nu/2+1/2)}d\tau \ . \qquad \text{(A.7)}$$

Now, using the same change of variables as in the proof of Theorem 1, the integral in (A.7) is

$$(1/2)\left[2/(2\theta+\nu\sigma^2)\right]^{\frac{\nu}{2}+r-h}\Gamma\left(\frac{\nu}{2}+r-h\right) , \qquad \text{(A.8)}$$

and so, substituting (A.8) into (A.7), collecting terms and allowing for the change from θ to λ completes the proof of Lemma A1. #

Proof of Corollary 2.

The desired expressions are directly obtained from Theorem 2 as $e \sim Mt(0,\nu\sigma^2/(\nu-2)I_T)$ when τ has an inverted gamma distribution with scale parameter σ^2 and degrees of freedom parameter ν. To establish $\rho_{Mt}\left[E(y),Xb\right]$ and $\rho_{Mt}\left[E(y),Xb^*\right]$ we merely need to note that

$$E(\tau^2) = \nu\sigma^2/(\nu-2) \qquad \text{(A.9)}$$

To obtain $\rho_{Mt}\left[E(y),X\hat{b}\right]$ we have from Lemma A1 that

$$\int_0^{\infty} P^{\tau}_{ij} f(\tau)d\tau = P_{ij2} \qquad \text{(A.10)}$$

and

$$\int_0^{\infty} \tau^2 P^{\tau}_{ij} f(\tau)d\tau = \nu\sigma^2 P_{ij1}/(\nu-2) , \qquad \text{(A.11)}$$

where

$$P_{ijn} = \sum_{r=0}^{\infty} \frac{(2\lambda/\nu)^r\Gamma\left(\frac{\nu}{2}+r+n-2\right)}{r!(1+2\lambda/\nu)^{\nu/2+r+n-2}\Gamma\left(\frac{\nu}{2}+n-2\right)} I_x\left[\tfrac{1}{2}(m+i)+r;\tfrac{1}{2}(v+j)\right] ,$$

$i,j,n = 0,1,2,\ldots,$ and $I_x(.;.)$ is Pearson's incomplete beta function with $x = cm/(v+cm)$. Substituting (A.6), (A.10) and (A.11) into equation (13) of Theorem 1 completes the proof. #

Proof of Theorem 3.

First, $\tilde{\sigma}_e^2 = e'Me/v$, where $M = I - XS^{-1}X'$ is an idempotent matrix of rank v. So,

$$\begin{aligned}\rho\left(\sigma_e^2,\tilde{\sigma}_e^2\right) &= E\left(\tilde{\sigma}_e^2 - \sigma_e^2\right)^2 \\ &= E\left(e'Me - vE(\tau^2)\right)^2/v^2 \\ &= E\left(G'\right)/v^2 \\ &= \frac{1}{v^2}\int_0^\infty E_N\left(G'\right)f(\tau)d\tau \qquad \text{(A.12)}\end{aligned}$$

where

$$G' = (e'Me)^2 - 2vE(\tau^2)(e'Me) + v^2\left(E(\tau^2)\right)^2$$

and $E_N\left(G'\right) = E\left(G'\right)$ when $e \sim N(0,\tau^2 I_T)$. Under these conditions $e'Me/\tau^2 \sim \chi_v^2$, and so

$$E_N\left(G'\right) = (v+2)\tau^4 - v\left[2E(\tau^2)\tau^2 - \left(E(\tau^2)\right)^2\right]. \qquad \text{(A.13)}$$

Substituting (A.13) into (A.12) completes the derivation of $\rho\left(\sigma_e^2,\tilde{\sigma}_e^2\right)$.

Secondly, we can write

$$\sigma_e^{*2} = (e_1'Me_1 + e_1'Ce_1)/(v+m) \qquad \text{(A.14)}$$

where $e_1 = d + e$, $d = X(\beta-\beta_0)$, $\beta_0 = R\bar{r}$ is any solution of $R\beta_0 = r$, and $C = XS^{-1}R'[RS^{-1}R']^{-1}RS^{-1}X'$ is an idempotent matrix of rank m. In writing (A.14) we have used the fact that $e_1'Me_1 = e'Me$, as $d'M = 0$.

So,

$$\rho\left(\sigma_e^2,\sigma_e^{*2}\right) = E\left\{(e_1'Me_1+e_1'Ce_1)^2 - 2(v+m)E(\tau^2)(e_1'Me_1+e_1'Ce_1)\right.$$

$$+ (v+m)^2\left(E(\tau^2)\right)^2\Big\}/(v+m)^2$$

$$= E\left\{G''\right\}/(v+m)^2$$

$$= \left[\int_0^\infty E_N\left\{G''\right\}f(\tau)d\tau\right]/(v+m)^2 , \qquad (A.15)$$

where $E_N\left\{G''\right\} = E\left\{G''\right\}$ when $e \sim N(0,\tau^2 I_T)$. Under this assumption $(e_1'Me_1+e_1'Ce_1)/\tau^2 \sim \chi^2_{v+m;\lambda_\tau}$ and so,

$$E_N\left\{G''\right\} = \tau^4\left[(v+m)(v+m+2) + 4(v+m+2)\lambda_\tau + 4\lambda_\tau^2\right] - 2(v+m)(v+m+2\lambda_\tau)E(\tau^2)$$

$$+ (v+m)^2\left(E(\tau^2)\right)^2 . \qquad (A.16)$$

Substituting (A.16) into (A.15), noting that $\lambda_\tau = \theta/\tau^2$, and integrating over τ completes the derivation of $\rho\left(\sigma_e^2,\sigma_e^{*2}\right)$.

Finally, to establish $\rho\left(\sigma_e^2,\hat{\sigma}_e^2\right)$ we write

$$\hat{\sigma}_e^2 = \left\{(e_1'Me_1)(v+m) + \left(ve_1'Ce_1-me_1'Me_1\right)I_{[0,c]}\left(ve_1'Ce_1/me_1'Me_1\right)\right\}/\left[v(v+m)\right],$$

so that,

$$\rho\left(\sigma_e^2,\hat{\sigma}_e^2\right) = \left[\int_0^\infty \tau^4 E_N\left\{G'''\right\}f(\tau)d\tau\right]/\left[v^2(v+m)^2\right] \qquad (A.17)$$

where

$$G''' = \Big\{(e_1'Me_1/\tau^2)^2(v+m)^2 + v^2(v+m)^2\left(E(\tau^2)\right)^2/\tau^4$$

$$- 2v(v+m)^2\left(E(\tau^2)/\tau^2\right)(e_1'Me_1/\tau^2) + \Big[-m(2v+m)(e_1'Me_1/\tau^2)$$

$$+ v^2(e_1'Ce_1/\tau^2) + 2v^2(e_1'Me_1/\tau^2)(e_1'Ce_1/\tau^2)$$

$$+ 2mv(v+m)\left(E(\tau^2)/\tau^2\right)(e_1'Me_1/\tau^2) - 2v^2(v+m)\left(E(\tau^2)/\tau^2\right).$$

$$(e_1'Ce_1/\tau^2)\Big]I_{[0,c]}\left((ve_1'Ce_1/\tau^2)/(e_1'Me_1/\tau^2)\right)\Big\} ,$$

and $E_N\left\{G'''\right\} = E\left\{G'''\right\}$ when $e \sim N(0,\tau^2 I_T)$.

Using the results of Clarke, Giles and Wallace (1987a,b) we have

$$E_N\left\{G'''\right\} = v(v+m)^2(v+2) + v^2(v+m)^2\left(E(\tau^2)\right)^2/\tau^4 - 2v^2(v+m)^2E(\tau^2)/\tau^2$$

$$- mv(2v+m)(v+2)P^{\tau}_{04} + v^2\left[m(m+2)P^{\tau}_{40} + 4(m+2)\lambda_{\tau}P^{\tau}_{60} + 4\lambda^2_{\tau}P^{\tau}_{80}\right]$$

$$+ 2v^3\left(mP^{\tau}_{22} + 2\lambda_{\tau}P^{\tau}_{42}\right) + 2mv^2(v+m)\left(E(\tau^2)/\tau^2\right)P^{\tau}_{02}$$

$$- 2v^2(v+m)\left(E(\tau^2)/\tau^2\right)\left(mP^{\tau}_{20} + 2\lambda_{\tau}P^{\tau}_{40}\right), \qquad \text{(A.18)}$$

and so, substituting (A.18) into (A.17), noting that $\lambda_{\tau} = 0/\tau^2$, completes the derivation of $\rho\left(\sigma^2_e,\hat{\sigma}^2_e\right)$ and the proof of Theorem 3. #

Proof of Corollary 3.

This corollary follows from Theorem 3 in the same way that Corollary 2 was established from Theorem 2. The only additional information we require is that

$$E(\tau^4) = \nu^2\sigma^4/\left[(\nu-2)(\nu-4)\right]$$

and, from Lemma A1,

$$\int_0^{\infty} \tau^4 P^{\tau}_{ij} f(\tau)d\tau = \nu^2\sigma^4 P_{ij0}/\left[(\nu-2)(\nu-4)\right]. \qquad \#$$

Footnotes

1. This subclass of the spherically symmetric family is sometimes called the compound normal family. See, for instance, Muirhead (1982).

2. We require the existence of the first two moments if risk, under squared error loss, is to be a meaningful basis for the comparison of the estimators.

3. Note that these estimators of σ_e^2 correspond to the usual estimators of the error variance for normally distributed regression disturbances. The same is not true for the maximum likelihood (ML) or minimum mean squared error (MMSE) estimators of σ_e^2: the ML and MMSE estimators of σ_e^2 when the errors are SSD depend on the specific form of the SSD. See, for instance, Zellner (1976) and King (1979) for further discussion and some examples.

4. Clearly, if the error distribution has an infinite variance (for example, if it is the Cauchy distribution) then the least squares estimator will have zero efficiency.

5. We exclude this latter member by the assumption of finite variance.

6. King considers the wider class of elliptically symmetric disturbances while Zellner, Ullah and Zinde-Walsh, Ullah and Phillips, Sutradhar and Ali, Singh and Sutradhar investigate Mt errors.

7. We consider the estimation of the prediction vector rather than the location vector so that our results are independent of the design matrix. This is equivalent to assuming orthonormal regressors in the β space.

8. We numerically evaluated the risk functions for v = 10, 16, 20, 30; m = 1, 2, 3, 4, 5; α = 0.01, 0.05, 0.25, 0.30, 0.50, 0.75, 0.90 and that value associated with a critical value of unity; ν = 5, 10, 50, 100, 500, 1000, 5000, 10,000, 100,000, ∞, and $\lambda \in [0,3(0.1);3,20(0.5)]$.

Full results are available on request. All evaluations were carried out using double-precision FORTRAN on an AT computer. Davies (1980) algorithm was used to evaluate the P_{ij}'s and the subroutines GAMMLN and BETAI from Press et al. (1986) were utilized to obtain the P_{ijn}'s. Using these programs we found that the risk expressions were efficiently evaluated with no observed convergence problems.

9. That the risk of the pre-test estimator can dominate both of its components over any or all of the parameter space may seem counter intuitive. We may believe that as the pre-test estimator is a weighted sum of its component estimators then its risk function should be enveloped by those of its components. This, however, confuses the distinction between a weighted sum of the moments of the component estimators and the moments of their weighted sum. The dominance of the pre-test estimator, for suitably chosen c, over the unrestricted estimator for all λ, and over the restricted estimator for some λ, also occurs when estimating the error variance after a pre-test for homogeneity. See, for example, Bancroft (1944), Ohtani and Toyoda (1978), and Toyoda and Wallace (1975).

10. Ohtani considers the question of the optimal significance level for the pre-test problem examined by Clarke et al. (1987b) when the component estimators are based on the minimum mean squared error principle. He compares the sampling properties of the pre-test estimator with those of the Stein (1964) estimator (extended to the linear regression case) and shows first, that the Stein estimator can be written as a pre-test estimator with critical value of $v/(v+2)$, and secondly, that the numerical evaluations suggest that the "Stein pre-test" estimator is optimal in some sense. See also Gelfand and Dey (1988).

11. The exceptions are those pre-test estimators with a critical value in the neighbourhood of $c = 0$.

12. It is straightforward to show that this feature holds under the null hypothesis for any ν, but the proof for $\lambda \neq 0$ is not obvious.

References

Bancroft, T.A., 1944, On biases in estimation due to the use of preliminary tests of significance, Annals of Mathematical Statistics 15, 190-204.

Blattberg, R.C. and N.J. Gonedes, 1974, A comparison of the stable and Student distributions as statistical models for stock prices, Journal of Business 47, 244-280.

Box, G.E.P., 1952, Multi-factor designs of first order, Biometrika 39, 49-57.

Brook, R.J., 1972, On the use of minimax regret functions to set significance points in prior tests of estimation, Ph.D. thesis, North Carolina State University.

Chmielewski, M.A., 1981, Elliptically symmetric distributions: A review and bibliography, International Statistical Review 49, 67-74.

Clarke, J.A., D.E.A. Giles and T.D. Wallace, 1987a, Estimating the error variance in regression after a preliminary test of restrictions on the coefficients, Journal of Econometrics 34, 293-304.

Clarke, J.A., D.E.A. Giles and T.D. Wallace, 1987b, Preliminary-test estimation of the error variance in linear regression, Econometric Theory 3, 299-304.

Coursey, D. and H. Nyquist, 1988, Applications of robust estimation techniques in demand analysis, Applied Economics 20, 595-610.

Davies, R.B., 1980, The distribution of a linear combination of χ^2 random variables (Algorithm AS 155), Applied Statistics 29, 323-333.

Devlin, S.J., R. Gnanadesikan and J.R. Kettenring, 1976, Some multivariate applications of elliptical distributions, in S. Ikeda (ed.), Essays in probability and statistics (Shinko Tsusho, Tokyo).

Fama, E.F., 1963, Mandelbrot and the stable Paretian hypothesis, Journal of Business 36, 420-429.

Fama, E.F., 1965, The behaviour of stock market prices, Journal of Business 38, 34-105.

Gelfand, A.E. and D.K. Dey, 1988, Improved estimation of the disturbance variance in a linear regression model, Journal of Econometrics 39, 387-395.

Granger, C.W.J. and D. Orr, 1972, "Infinite variance" and research strategy in time-series analysis, Journal of the American Statistical Association 67, 275-285.

Hampel, F.R., E.M. Ronchetti, P.J. Rousseeuw and W.A. Stahel, 1983, Robust statistics (John Wiley, New York).

Huber, P.J., 1981, Robust statistics (John Wiley, New York).

Judge, G.G. and M.E. Bock, 1978, The statistical implications of pre-test and Stein-rule estimators in econometrics (North-Holland, Amsterdam).

Judge, G.G., W.E. Griffiths, R.C. Hill, H. Lutkepohl and T-C. Lee, 1985, The theory and practice of econometrics, 2nd edition (John Wiley, New York).

Judge, G.G., S. Miyazaki and T. Yancey, 1985, Minimax estimators for the location vectors of spherically symmetric densities, Econometric Theory 1, 409-417.

Kelker, D., 1970, Distribution theory of spherical distributions and a location-scale parameter generalization, Sankhya *A* 32, 419-430.

King, M.L., 1979, Some aspects of statistical inference in the linear regression model, Ph.D. thesis, University of Canterbury.

Koenker, R.W., 1982, Robust methods in econometrics, Communications in Statistics - Econometric Reviews 1, 213-255.

Muirhead, R.J., 1982, Aspects of multivariate statistical theory (Addison-Wesley, Reading, Mass.).

Ohtani, K., 1988, Optimal levels of significance of a pre-test in estimating the disturbance variance after the pre-test for a linear hypothesis on coefficients in a linear regression, Economics Letters 28, 151-156.

Ohtani, K. and T. Toyoda, 1978, Minimax regret critical values for a preliminary test in pooling variances, Journal of the Japan Statistical Society 8, 15-20.

Praetz, P.D., 1972, The distribution of share price changes, Journal of Business 45, 49-55.

Press, W.H., B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, 1986, Numerical recipes: The art of scientific computing (Cambridge University Press, New York).

Sharpe, W.F., 1971, Mean-absolute deviation characteristic lines for securities and portfolios, Management Science 18, B1-B13.

Singh, R.S., 1988, Estimation of error variance in linear regression models with errors having multivariate Student-t distribution with unknown degrees of freedom, Economics Letters 27, 47-53.

Stein, C., 1964, Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean, Annals of the Institute of Statistical Mathematics 16, 155-160.

Sutradhar, B.C., 1988, Testing linear hypothesis with t error variable, Sankhya *B* 50, 175-180.

Sutradhar, B.C. and M.M. Ali, 1986, Estimation of the parameters of a regression model with a multivariate t error variable, Communications in Statistics - Theory and Methods 15, 429-450.

Thomas, D.H., 1970, Some contributions to radial probability distributions, statistics, and the operational calculi, Ph.D. thesis, Wayne State University.

Toyoda, T. and T.D. Wallace, 1975, Estimation after a preliminary test of homogeneity and optimal levels of significance for the pre-test, Journal of Econometrics 3, 395-404.

Ullah, A. and V. Zinde-Walsh, 1984, On the robustness of LM, LR and W tests in regression models, Econometrica 52, 1055-1066.

Ullah, A. and P.C.B. Phillips, 1986, Distribution of the F-ratio, Econometric Theory 2, 449-452.

Wallace, T.D., 1977, Pre-test estimation in regression: A survey, American Journal of Agricultural Economics 59, 431-443.

Zellner, A., 1976, Bayesian and non-Bayesian analysis of the regression model with multivariate Student-t error terms, Journal of the American Statistical Association 71, 400-405.

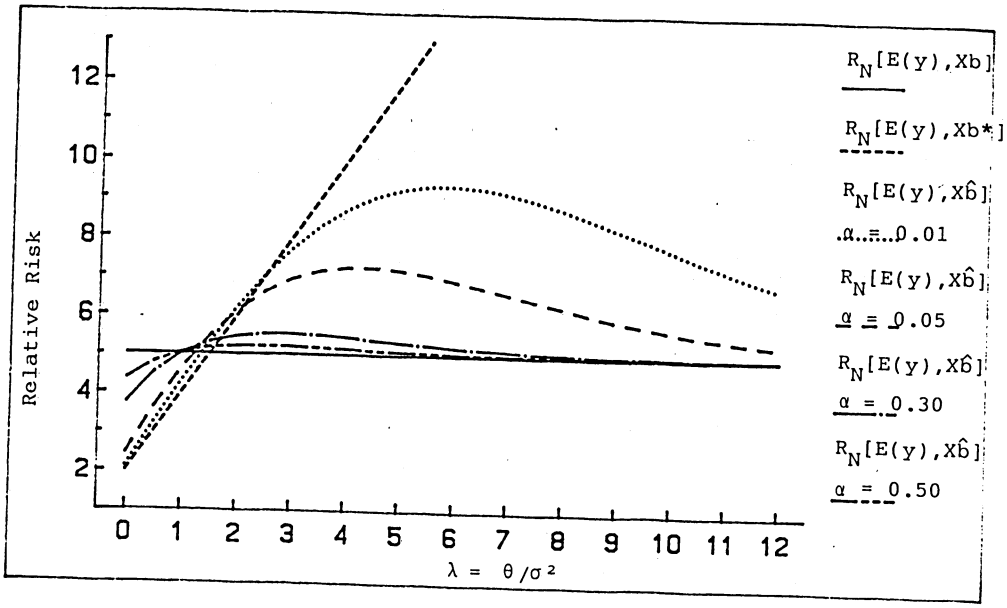

FIGURE 1. Relative risk functions for the unrestricted, restricted and pre-test estimators of E(y) when $e \sim N(0,\sigma^2 I_T)$: $T = 30$, $k = 5$, $m = 3$.

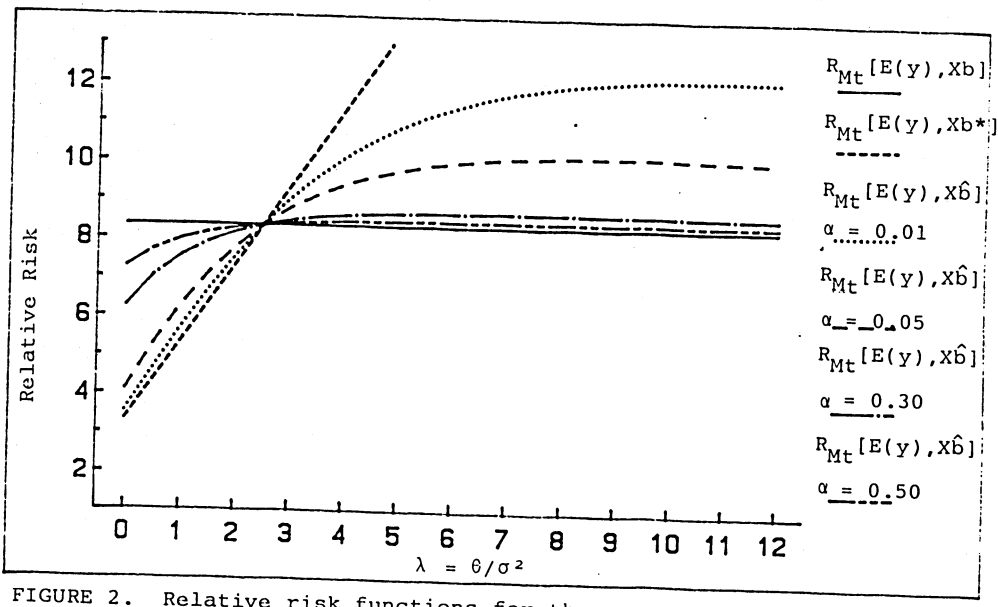

FIGURE 2. Relative risk functions for the unrestricted, restricted and pre-test estimators of E(y) when $e \sim Mt(0, \nu\sigma^2/(\nu-2)I_T)$: $T = 30$, $k = 5$, $m = 3$ and $\nu = 5$.

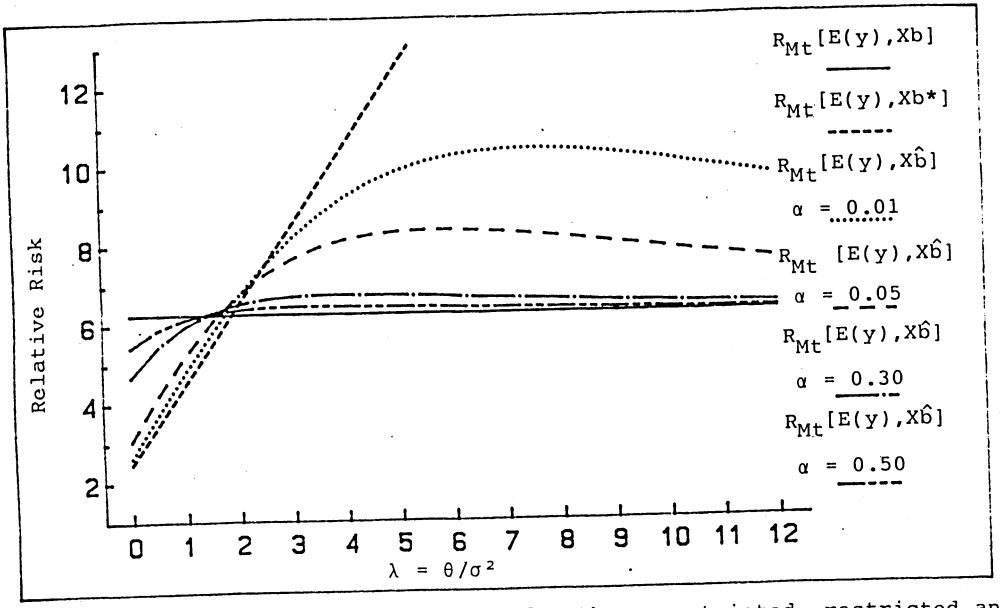

FIGURE 3. Relative risk functions for the unrestricted, restricted and pre-test estimators of E(y) when $e \sim Mt(0, \nu\sigma^2/(\nu-2)I_T)$: $T = 30$, $k = 5$, $m = 3$ and $\nu = 10$.

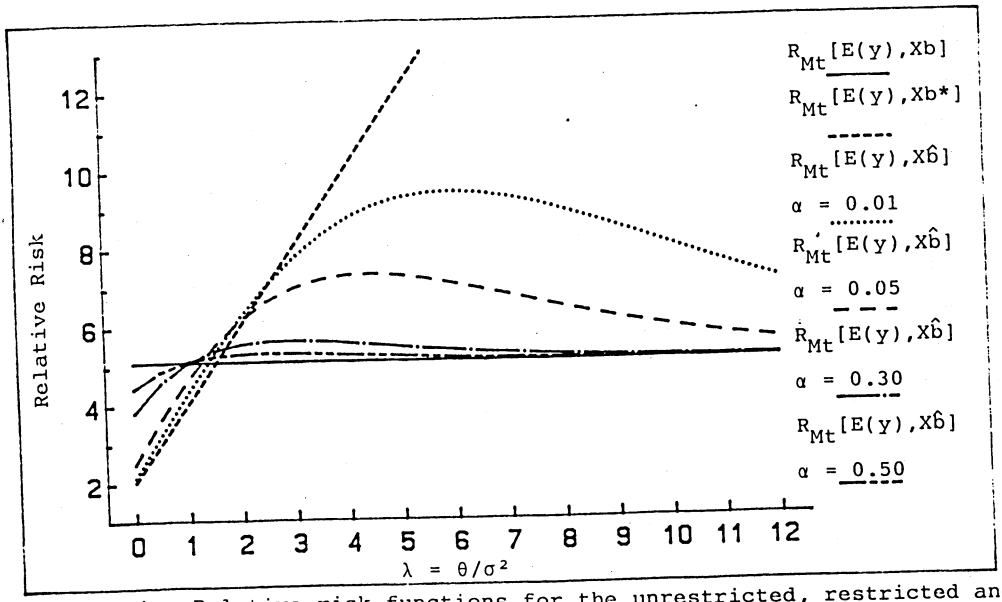

FIGURE 4. Relative risk functions for the unrestricted, restricted and pre-test estimators of E(y) when $e \sim Mt(0, \nu\sigma^2/(\nu-2)I_T)$: $T = 30$, $k = 5$, $m = 3$, and $\nu = 100$.

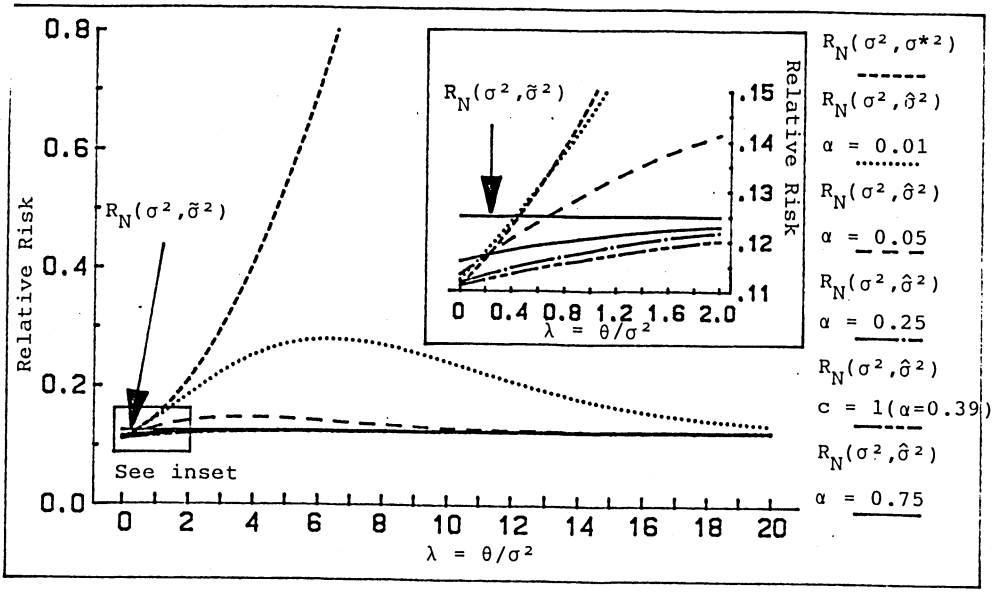

FIGURE 5. Relative risk functions for the unrestricted, restricted and pre-test estimators of the error variance when $e \sim N(0,\sigma^2 I_T)$: $T = 20$, $k = 4$, $m = 2$.

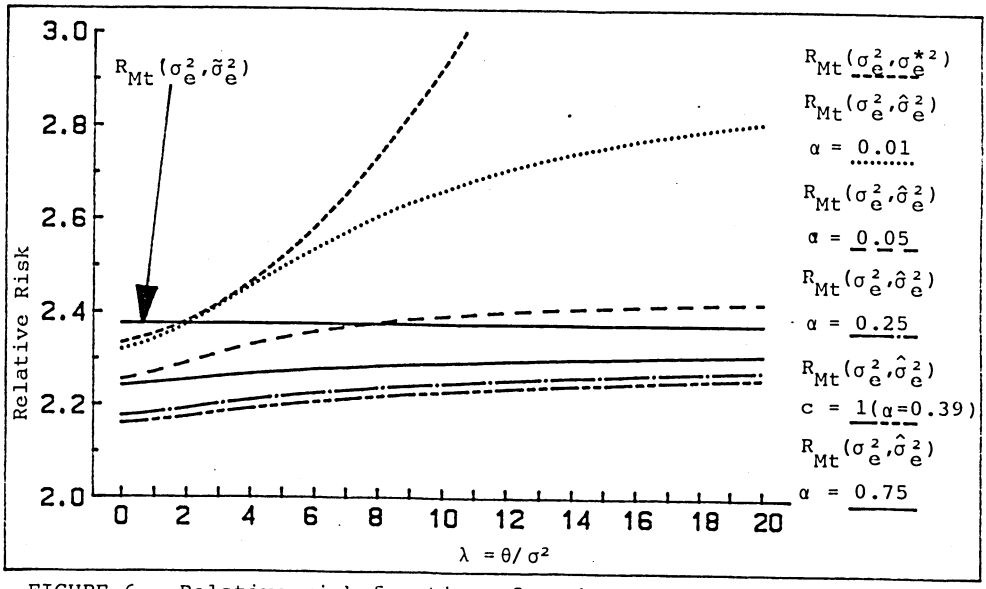

FIGURE 6. Relative risk functions for the unrestricted, restricted and pre-test estimators of the error variance when $e \sim Mt(0,\nu\sigma^2(\nu-2)I_T)$: $T = 20$, $k = 4$, $m = 2$, and $\nu = 5$.

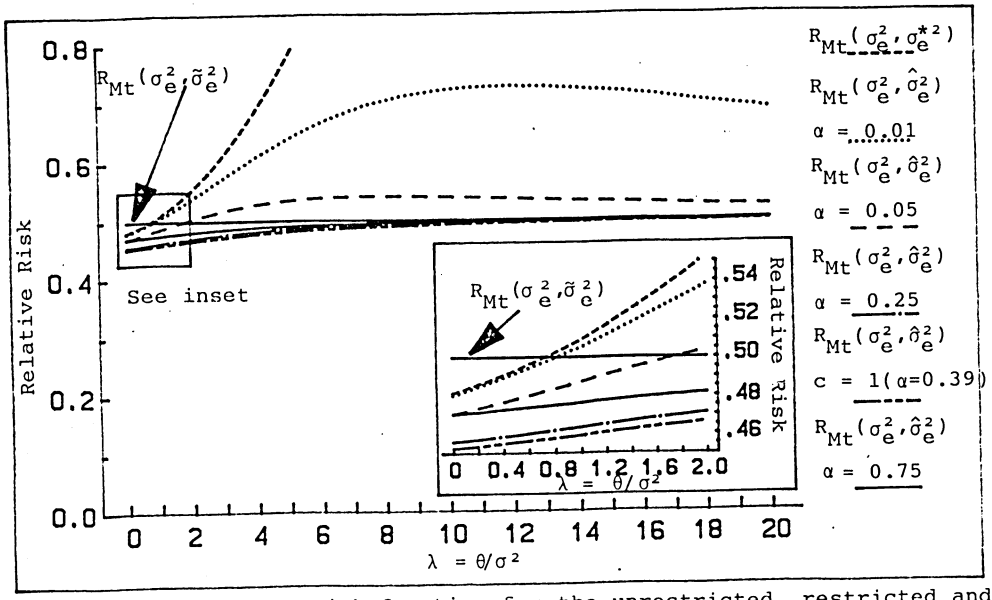

FIGURE 7. Relative risk function for the unrestricted, restricted and pre-test estimators of the error variance when $e \sim Mt(0/\sigma^2/(\nu-2)I_T)$: $T = 20$, $k = 4$, $m = 2$, and $\nu = 10$.

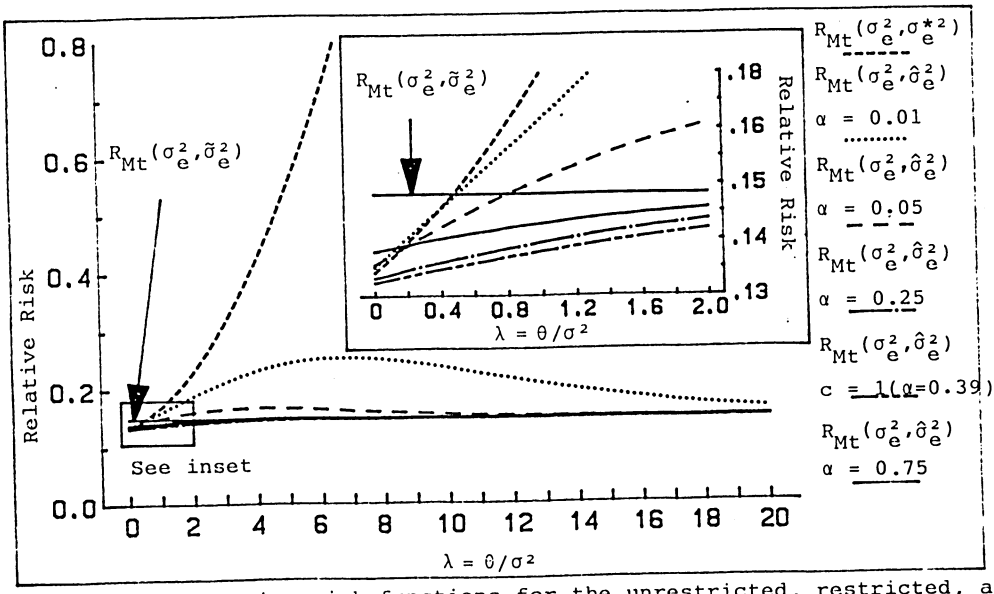

FIGURE 8. Relative risk functions for the unrestricted, restricted, and pre-test estimators of the error variance when $e \sim Mt(0,\nu\sigma^2/(\nu-2)I_T)$: $T = 20$, $k = 4$, $m = 2$, and $\nu = 100$.

LIST OF DISCUSSION PAPERS*

No. 8401 Optimal Search, by Peter B. Morgan and Richard Manning.

No. 8402 Regional Production Relationships During the Industrialization of New Zealand, 1935-1948, by David E. A. Giles and Peter Hampton.

No. 8403 Pricing Strategies for a Non-Replenishable Item Under Variable Demand and Inflation, by John A. George.

No. 8404 Alienation Rights in Traditional Maori Society, by Brent Layton.

No. 8405 An Engel Curve Analysis of Household Expenditure in New Zealand, by David E. A. Giles and Peter Hampton.

No. 8406 Paying for Public Inputs, by Richard Manning, James R. Markusen, and John McMillan.

No. 8501 Perfectly Discriminatory Policies in International Trade, by Richard Manning and Koon-Lam Shea.

No. 8502 Perfectly Discriminatory Policy Towards International Capital Movements in a Dynamic World, by Richard Manning and Koon-Lam Shea.

No. 8503 A Regional Consumer Demand Model for New Zealand, by David E. A. Giles and Peter Hampton.

No. 8504 Optimal Human and Physical Capital Accumulation in a Fixed-Coefficients Economy, by R. Manning.

No. 8601 Estimating the Error Variance in Regression After a Preliminary Test of Restrictions on the Coefficients, by David E. A. Giles, Judith A. Mikolajczyk and T. Dudley Wallace.

No. 8602 Search While Consuming, by Richard Manning.

No. 8603 Implementing Computable General Equilibrium Models: Data Preparation, Calibration, and Replication, by K. R. Henry, R. Manning, E. McCann and A. E. Woodfield.

No. 8604 Credit Rationing: A Further Remark, by John G. Riley.

No. 8605 Preliminary-Test Estimation in Mis-Specified Regressions, by David E. A. Giles.

No. 8606 The Positive-Part Stein-Rule Estimator and Tests of Linear Hypotheses, by Aman Ullah and David E. A. Giles.

No. 8607 Production Functions that are Consistent with an Arbitrary Production-Possibility Frontier, by Richard Manning.

No. 8608 Preliminary-Test Estimation of the Error Variance in Linear Regression, by Judith A. Clarke, David E. A. Giles and T. Dudley Wallace.

No. 8609 Dual Dynamic Programming for Linear Production/Inventory Systems, by E. Grant Read and John A. George.

No. 8610 Ownership Concentration and the Efficiency of Monopoly, by R. Manning.

No. 8701 Stochastic Simulation of the Reserve Bank's Model of the New Zealand Economy, by J. N. Lye.

No. 8702 Urban Expenditure Patterns in New Zealand, by Peter Hampton and David E. A. Giles.

No. 8703 Preliminary-Test Estimation of Mis-Specified Regression Models, by David E. A. Giles.

No. 8704 Instrumental Variables Regression Without an Intercept, by David E. A. Giles and Robin W. Harrison.

No. 8705 Household Expenditure in Sri Lanka: An Engel Curve Analysis, by Mallika Dissanayake and David E. A Giles.

No. 8706 Preliminary-Test Estimation of the Standard Error of Estimate in Linear Regression, by Judith A. Clarke.

No. 8707 Invariance Results for FIML Estimation of an Integrated Model of Expenditure and Portfolio Behaviour, by P. Dorian Owen.

No. 8708 Social Cost and Benefit as a Basis for Industry Regulation with Special Reference to the Tobacco Industry, by Alan E. Woodfield.

No. 8709 The Estimation of Allocation Models With Autocorrelated Disturbances, by David E. A. Giles.

No. 8710 Aggregate Demand Curves in General-Equilibrium Macroeconomic Models: Comparisons with Partial-Equilibrium Microeconomic Demand Curves, by P. Dorian Owen.

No. 8711 Alternative Aggregate Demand Functions in Macro-economics: A Comment, by P. Dorian Owen.

No. 8712 Evaluation of the Two-Stage Least Squares Distribution Function by Imhof's Procedure by P. Cribbett, J. N. Lye and A. Ullah.

No. 8713 The Size of the Underground Economy: Problems and Evidence, by Michael Carter.

(Continued on back cover)

No. 8714 A Computable General Equilibrium Model of a Fisherine Method to Close the Foreign Sector, by Ewen McCann and Keith Mclaren.

No. 8715 Preliminary-Test Estimation of the Scale Parameter in a Mis-Specified Regression Model, by David E. A. Giles and Judith A. Clarke.

No. 8716 A Simple Graphical Proof of Arrow's Impossibility Theorem, by John Fountain.

No. 8717 Rational Choice and Implementation of Social Decision Functions, by Manimay Sen.

No. 8718 Divisia Monetary Aggregates for New Zealand, by Ewen McCann and David E. A. Giles.

No. 8719 Telecommunications in New Zealand: The Case for Reform, by John Fountain.

No. 8801 Workers' Compensation Rates and the Demand for Apprentices and Non-Apprentices in Victoria, by Pasquale M. Sgro and David E. A. Giles.

No. 8802 The Adventures of Sherlock Holmes, the 48% Solution, by Michael Carter.

No. 8803 The Exact Distribution of a Simple Pre-Test Estimator, by David E. A. Giles.

No. 8804 Pre-testing for Linear Restrictions in a Regression Model With Student-t Errors, by Judith A. Clarke.

No. 8805 Divisia Monetary Aggregates and the Real User Cost of Money, by Ewen McCann and David Giles.

No. 8806 The Management of New Zealand's Lobster Fishery, by Alan Woodfield and Pim Borren.

No. 8807 Poverty Measurement: A Generalization of Sen's Result, by Prasanta K. Pattanaik and Manimay Sen.

No. 8808 A Note on Sen's Normalization Axiom for a Poverty Measure, by Prasanta K. Pattanaik and Manimay Sen.

No. 8809 Budget Deficits and Asset Sales, by Ewen McCann.

No. 8810 Unorganized Money Markets and 'Unproductive' Assets in the New Structuralist Critique of Financial Liberalization, by P. Dorian Owen and Otton Solis-Fallas.

No. 8901 Testing for Financial Buffer Stocks in Sectoral Portfolio Models, by P. Dorian Owen.

No. 8902 Provisional Data and Unbiased Prediction of Economic Time Series by Karen Browning and David Giles.

No. 8903 Coefficient Sign Changes When Restricting Regression Models Under Instrumental Variables Estimation, by David E. A. Giles.

No. 8904 Economies of Scale in the New Zealand Electricity Distribution Industry, by David E. A. Giles and Nicolas S. Wyatt.

No. 8905 Some Recent Developments in Econometrics: Lessons for Applied Economists, by David E. A. Giles.

No. 8906 Asymptotic Properties of the Ordinary Least Squares Estimator in Simultaneous Equations Models, by V. K. Srivastava and D. E. A. Giles.

No. 8907 Unbiased Estimation of the Mean Squared Error of the Feasible Generalised Ridge Regression, by V. K. Srivasatva and D. E. A. Giles.

No. 8908 An Unbiased Estimator of the Covariance Matrix of the Mixed Regression Estimator, by D. E. A. Giles and V. K. Srivastava.

No. 8909 Pre-testing for Linear Restrictions in a Regression Model with Spherically Symmetric Disturbances.

* Copies of these Discussion Papers may be obtained for $4 (including postage, price changes occasionally) each by writing to the Secretary, Department of Economics, University of Canterbury, Christchurch, New Zealand.