CANTER

9304

Department of Economics
UNIVERSITY OF CANTERBURY
CHRISTCHURCH, NEW ZEALAND

ISSN 1171-0705

PRE-TEST ESTIMATION OF THE REGRESSION SCALE PARAMETER WITH MULTIVARIATE STUDENT-t ERRORS AND INDEPENDENT SUB-SAMPLES*

Juston Z. Anderson and Judith A. Giles

Discussion Paper

No. 9304

Department of Economics, University of Canterbury
Christchurch, New Zealand

Discussion Paper No. 9304

March 1993

PRE-TEST ESTIMATION OF THE REGRESSION SCALE PARAMETER WITH MULTIVARIATE STUDENT-t ERRORS AND INDEPENDENT SUB-SAMPLES*

Juston Z. Anderson and Judith A. Giles

PRE-TEST ESTIMATION OF THE REGRESSION SCALE PARAMETER WITH MULTIVARIATE STUDENT-t ERRORS AND INDEPENDENT SUB-SAMPLES*

Juston Z. Anderson

and

Judith A. Giles

Department of Economics
University of Canterbury

March, 1993

SUMMARY

We consider the estimation of the error variance of a regression when additional information is available in the form of a second sample, which may be generated from a process with the same variance. This problem has received attention in the literature when the joint errors are members of the spherically symmetric family. We extend this assumption to one in which the errors in each sample are independent multivariate Student-t random vectors. We derive the exact risk under quadratic loss of the pre-test estimator which results after a test for homogeneity of the variances and we compare the risk of this estimator with that of its component estimators.

AMS (1980) Subject Classifications: Primary 62J05, 62F11; secondary 62P20.

Key words and phrases: Preliminary testing, conditional inference, non-normal disturbances, testing for homogeneity.

CORRESPONDENCE: Dr Judith Giles, Department of Economics, University of Canterbury, Private Bag 4800, Christchurch, New Zealand.

FAX:64-3-3642635; **VOICE:**64-3-3642520; **EMAIL:**j.giles@csc.canterbury.ac.nz

1. INTRODUCTION AND MODEL FRAMEWORK

Suppose we have two linear regressions of the form:

$$y_i = X_i\beta_i + \varepsilon_i, \qquad i = 1,2 \qquad (1)$$

where y_i is a $(T_i\times 1)$ vector of observations on the dependent variable; X_i is a $(T\times k_i)$ full column rank matrix of non-stochastic regressors with $k_i < T_i$; β_i is a $(k_i\times 1)$ vector of parameters; and ε_i is a $(T_i\times 1)$ vector of disturbance terms. We assume that ε_1 and ε_2 are independently generated from multivariate Student-t (Mt) distributions, with probablitity density functions (pdf):

$$f(\varepsilon_i|\nu_i,\sigma_i) = \left(\nu_i^{\nu_i/2}\,\Gamma((\nu_i+T_i)/2)\right)\left(\pi_i^{T_i/2}\,\Gamma(\nu_i/2)\sigma_i^{T_i}\right)^{-1}$$

$$\times\left(\nu_i+\varepsilon_i'\varepsilon_i/\sigma_i^2\right)^{-(T_i+\nu_i)/2} \quad ; \quad i=1,2.$$

ν_i is the degrees of freedom parameter and σ_i is the scale parameter of the distribution. If $\nu_i>2$, then $E(\varepsilon_i) = 0$ and $Var(\varepsilon_i) = \sigma^2_{\varepsilon_i} = \nu_i\sigma_i^2/(\nu_i-2)$. When $\nu_i = 1$, the pdf is Cauchy for which no finite integral moments exist; when $\nu_i = \infty$ it is normal.

We suppose that the researcher desires an estimate of $\sigma^2_{\varepsilon_1}$, the variance of the first sample, when it is suspected that ε_2 may have the same variance. When the error variances are unequal we use the so-called never-pool estimator (NPE), s_N^2, to estimate $\sigma^2_{\varepsilon_1}$:

$$s_N^2 = s_1^2 = \varepsilon_1'M_1\varepsilon_1/v_1 \qquad (2)$$

where $v_i = T_i-k_i$; $M_i = I_{T_i}-X_i(X_i'X_i)^{-1}X_i'$; $i=1,2$. s_N^2 uses only the information from the first sample. Alternatively, if $\sigma^2_{\varepsilon_1} = \sigma^2_{\varepsilon_2}$ it is then more efficient to use the "always-pool" estimator (APE), s_A^2:

$$s_A^2 = (v_1s_1^2 + v_2s_2^2)/(v_1+v_2) \qquad (3)$$

where s_2^2 is defined analogously to s_1^2. Given the uncertainty about the equality of the error variances a typical strategy is to pre-test for homogeneity and then use s_N^2 if we reject homogeneity or use s_A^2 if we cannot reject homogeneity. The estimator actually reported after such a procedure is the so-called pre-test estimator (PTE) s_P^2 :

$$s_P^2 = \begin{cases} s_A^2 & \text{if } J \le c \\ s_N^2 & \text{if } J > c \end{cases} \tag{4}$$

where $J = s_2^2/s_1^2$ is the test statistic used to test the hypothesis[1] $H_0 : \sigma^2_{\varepsilon_1} = \sigma^2_{\varepsilon_2}$ *vs.* $H_A : \sigma^2_{\varepsilon_1} < \sigma^2_{\varepsilon_2}$ and c is the critical value of the test corresponding to an $\alpha\%$ significance level.

If $\varepsilon' = (\varepsilon_1' \ \ \varepsilon_2')$ is distributed as a member of the elliptically symmetric family of distributions then it is well known that $f(J) = \phi^{-1} f\left(F_{(v_2, v_1)}\right)$ where $\phi = \sigma^2_{\varepsilon_1}/\sigma^2_{\varepsilon_2}$ is a measure of the hypothesis error and $F_{(v_2,v_1)}$ is a central F random variable with v_2 and v_1 degrees of freedom (see King (1979) and Chmielewski (1981)). However, if ε_1 and ε_2 are independent Mt random vectors then J has a non-standard distribution. Ohtani (1990) derives the density of J in this case:

$$f(J) = \left[B(v_1/2,\ v_2/2)\ B(\nu_1/2,\ \nu_2/2) \right]^{-1} \theta^{\nu_1/2}\ J^{\nu_1/2-1}$$

$$\times \int_0^{\infty} t^{(v_2+\nu_2)/2-1}\ (1-t)^{(v_1+\nu_1)/2-1}\ (t+\theta J(1-t))^{-(\nu_1+\nu_2)/2}\ dt$$

where $\theta = (v_2/v_1)[\nu_1\sigma_1^2/(\nu_2\sigma_2^2)]$ and $B(.,.)$ is the Beta function. Ohtani also tabulates a limited number of critical values for the test under this assumption, assuming a 5% significance level.

The particular pre-test problem considered here has been well investigated in the literature but not under the above specified error term assumptions. For example, Bancroft (1944), Toyoda and Wallace (1975), Ohtani and Toyoda (1978), and Bancroft and Han (1983) all consider this problem from various aspects under normal errors[2]. Bancroft (1944) concludes from his numerical evaluations that the PTE which uses a critical value of unity strictly dominates the NPE.

Toyoda and Wallace (1975) show that the risk of the APE always has two intersections with the risk of the NPE and that one intersection always lies in the range $\phi \in (0, 1)$. This implies that the NPE cannot strictly dominate the APE and *vice versa*. Toyoda and Wallace also show that a PTE with a critical value in the range $c \in (0, 2)$ strictly dominates the NPE, and that the PTE with c=1 "almost always" dominates the APE, except in the neighbourhood of the null hypothesis. They consequently suggest c=1 as a reasonable choice for the pre-test (p.399) and go on to show that this choice of critical value maximises relative average efficiency.

Ohtani and Toyoda (1978) consider optimal critical values for this pre-test problem according to a minimax regret criterion. They show that the risk function for the PTE declines monotonically for $c \in (0, 1]$, and obtains a local minimum at c=1. Using a minimax regret criterion they solve numerically for an "optimal" critical value, c*. They find that c* depends on v_1 and v_2 and ranges from 1.7 to 2.8 (for the cases evaluated) which correspond to sizes from 6% to 22%. Bancroft and Han (1983) also consider the choice of an optimal critical value according to another criterion - so called relative efficiency. Their results suggest significance levels ranging from 24% to 48%, depending on v_1 and v_2.

As the processes generating many time-series are non-normal, Giles (1990, 1992) extends the above cited work to one where the joint error term in the model is distributed according to the scale mixture of normals family of distributions, which are members of the elliptically symmetric family of distributions. It is then possible to consider error term assumptions which result in more or less kurtosis than under a normality assumption. One special member of this family is the Mt distribution which results in uncorrelated but dependent errors. She shows that the risk function for the PTE has a minimum when c=1 and also that for small values of the Mt shape parameter, ν, (*i.e.* "fat-tailed" distributions) this PTE strictly dominates *both* the NPE and the APE.

Here we extend the error term assumptions further by assuming that the errors in each sample are Mt but are *independent*. This allows for the error distributions for the samples to have potentially different shape, as well as scale, parameters. The outline of this paper is as follows: Section 2 derives the risk functions of the NPE, APE and PTE under the independent Mt assumption and undertakes some comparisons of the risk properties of the three estimators. In Section 3 we consider some numerical evaluations of these risk functions. Here we use Ohtani's (1990) critical values where possible but we also consider that a researcher may incorrectly assume normality of the errors and so use critical values from the central-F distribution. For the latter, of course, the *true* significance level will differ from the assigned *nominal* significance level. We conclude with some final remarks in Section 4 followed by an appendix which contains brief proofs of the theorems.

2. THE RISK FUNCTIONS

Let $\bar{s}^2$ be any estimator of $\sigma^2_{\varepsilon_1}$ and let its risk under quadratic loss be defined by $R(\bar{s}^2) = E(\bar{s}^2 - \sigma^2_{\varepsilon_1})^2$. Then, if ε_1 and ε_2 are generated from independent Mt distributions with $\nu_1>4$ and $\nu_2>4$ we have:

Theorem 1:

$$R(s_N^2) = 2\nu_1^2\sigma_1^4(v_1+\nu_1-2)/\left(v_1(\nu_1-2)^2(\nu_1-4)\right) \tag{5}$$

$$R(s_A^2) = \nu_1^2\sigma_1^4\Big\{(\nu_2-4)\Big[v_1(2v_1+\nu_1)+(v_2^2+v_1)(\nu_1-4)\Big]+v_2(v_2+2)(\nu_2-2)(\nu_1-4)/\phi^2$$

$$-2v_2^2(\nu_1-4)(\nu_2-4)/\phi\Big\}/\left((v_1+v_2)^2(\nu_1-2)^2(\nu_1-4)(\nu_2-4)\right) \tag{6}$$

$$R(s_P^2) = \nu_1^2\sigma_1^4\Big\{(v_1+v_2)^2(v_1+2)(\nu_1-2)+v_1v_2(v_2+2)P_{0440}(\nu_1-2)^2(\nu_1-4)$$

$$-v_2(2v_1+v_2)(v_1+2)P_{4004}(\nu_1-2)^2(\nu_1-4)+2v_1^2v_2P_{2222}(\nu_1-2)^2(\nu_1-4)$$

$$-2v_1(v_1+v_2)\Big((v_1+v_2)(\nu_1-4)+v_2P_{0220}(\nu_1-2)(\nu_1-4)$$

$$-v_2P_{2002}(\nu_1-2)(\nu_1-4)\Big)+v_1(v_1+v_2)^2(\nu_1-4)\Big\}/$$

$$\left(v_1(v_1+v_2)^2(\nu_1-2)(\nu_1-4)\right) \tag{7}$$

where

$$P_{abij} = \frac{\Gamma\left(\frac{\nu_1+\nu_2-a-b}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}\left(\frac{\nu_2-2}{\phi}\right)^{\nu_2/2} v_1^{(\nu_2-b)/2}\,(cv_2)^{(\nu_1-a)/2}$$

$$\times\, 2^{-(a+b)/2}\,(\nu_1-2)^{(\nu_1-a-b)/2}\int_0^1 t_1^{(\nu_2-b)/2-1}\,(1-t_1)^{(\nu_1-a)/2-1}$$

$$\times\left(v_1t_1(\nu_2-2)/\phi+(\nu_1-2)cv_2(1-t_1)\right)^{(a+b-\nu_1-\nu_2)/2}$$

$$\times\, I_{t_1}\left(\tfrac{1}{2}(v_2+i);\tfrac{1}{2}(v_1+j)\right)dt_1$$

Proof: See the appendix.

Remarks:

1. When c=0 (α=1) we always reject H_0 and the risk of the PTE collapses to that of the NPE. Conversely, if $c\to\infty$ ($\alpha\to 0$) we never reject H_0 and $R(s_P^2) = R(s_A^2)$.

2. The risk functions of s_N^2 and s_A^2 have two intersections with respect to ϕ. Let these be ϕ_1 and ϕ_2. Their values are $\phi_1 = (\omega + \tau^{1/2})/\zeta$ and $\phi_2 = (\omega - \tau^{1/2})/\zeta$, where $\omega = -2v_1v_2^2(\nu_1-4)(\nu_2-4)$, $\tau = 8v_1v_2^2(\nu_2-4)(\nu_1-4) \times\Big(2v_1^2(\nu_2-2)(v_1+2) - v_1\Big[v_2^2(\nu_1-\nu_2-2)+v_2(2-\nu_2)(\nu_1+2)-4(\nu_2-2)(\nu_1-2)\Big] + v_2(\nu_2-2)(\nu_1-2)(\nu_2+2)\Big)$, and $\zeta = 2v_2(\nu_2-4)\Big(4v_1^2 + v_1(4(\nu_1-2) - v_2(\nu_1-6)) + 2v_2(\nu_1-2)\Big)$. As $R(s_A^2)$ is a quadratic in $1/\phi$, with an asymptote at ϕ=0, one of these intersections will lie in the range (0, $-\infty$) and the other in the range (0, $+\infty$). There are three possibilities. First, one intersection lies in the range (0, 1). This implies that neither of the NPE or the APE strictly dominates the other, and accords with the results under spherically symmetric disturbances. The second alternative is that the positive intersection is greater than unity. Then, the NPE strictly dominates the APE over the range $\phi \in (0, 1]$. The third possibility is that there are no real intersections. This occurs when τ is negative and in this case the NPE again strictly dominates the APE. We consider these cases further in the next section.

3. When $\nu_1\to\infty$ and $\nu_2\to\infty$, the risk functions collapse to their normal counterparts (see, for example, Toyoda and Wallace (1975)).[3]

4. As $\phi\to 0$, $R(s_P^2) \to R(s_N^2)$ while $R(s_A^2) \to \infty$. That is, pre-testing leads us to follow the correct strategy when the prior information is very false.

5. Typically $R(s_A^2) < R(s_N^2)$ when ϕ=1, although there are exceptions as

noted in the above point 2. It is also possible for $R(s_P^2) < R(s_A^2)$ when $\phi=1$. We illustrate such cases in the next section.

6. **Theorem 2:**

   Extrema of $R(s_P^2)$ result when $c=0$, $c \to \infty$, and $c=1$.

   Proof: See the appendix.

   So, for any particular value of ϕ, the minimum risk estimator among those considered in this paper may be either the APE, the NPE, or the PTE with $c=1$.

3. NUMERICAL EVALUATIONS OF THE RISK FUNCTIONS

To illustrate the results we have numerically evaluated the risk functions. Various values of the arguments were considered: v_1, v_2 = 10, 30, 40; ν_1, ν_2 = 5, 10, 50, 100, ∞; those critical values corresponding to a *true* size of 5% (Ohtani (1990)), those from the central-F distribution corresponding to *nominal* sizes of 1%, 5%, 30% and 75% and a critical value of unity. Full details of these results are available on request. The evaluations were undertaken using a FORTRAN program written by the authors, which utilises several subroutines from Press *et al.* (1986). We executed the program on a VAX 7610 and a VAXstation 4000. Figures 1 to 4 provide representative results. The horizontal axis, in each figure, measures the extent of the hypothesis error $\phi \in (0, 1]$. The vertical axis measures risk and we have assumed, without loss of generality, that $\sigma_1^2 = 1$. The following points can be noted:

1. The figures illustrate the possible cases referred to in the previous Section as point 2. Specifically, the risk functions of the NPE and APE can intersect at a value of $\phi \in (0, 1)$ (see for example Figure 1). Our results suggest that this will occur for all values of v_1 and v_2

when $\nu_2 \geq \nu_1$. Figure 2 provides an example where the NPE strictly dominates the APE for all $\phi \in (0, 1]$. The evaluations suggest that this case is likely when ν_2 is sufficiently smaller than ν_1. So, if the error term of the model for the second sample has marginal distributions which have "fatter" tails than that for the first sample then it is *never* optimal to pool the data, even if the variances are equal. This result contrasts to that found when the joint disturbance is spherically symmetric. Then it is always preferable to pool the samples when the variances are equal rather than to simply ignore the prior information.

2. An increase in v_1, v_2 shifts the risk functions downwards, as does also an increase in ν_1, ν_2.
3. There is no strictly dominating estimator when the disturbances are normal or ν_1 and ν_2 are "large" (see for example Figure 3). Then, the APE has the smallest risk around the neighbourhood of H_0, while it is generally preferable to employ the PTE with c=1 otherwise.
4. For small ν_1, ν_2 (*e.g.* 5, 10) the PTE can strictly dominate both the NPE and the APE (see for example Figure 4). Typically, the PTE which uses c=1 strictly dominates all other PTE's. This result accords with those of Toyoda and Wallace (1975), Ohtani and Toyoda (1978), and Giles (1992).
5. Using the central-F critical values, as opposed to the values provided by Ohtani (1990), typically has a significant effect on the risk function; that is, there is a significant difference between the *nominal* and *true* sizes for the F-test. The distortion in size increases as ν_1, ν_2 decrease.
6. The PTE which uses c=1 *always* strictly dominates the NPE. It is never optimal to ignore the prior information.

4. CONCLUDING REMARKS

In this paper we have examined the risks of estimators of the regression error variance after a preliminary test for homogeneity, when the disturbances in each sample are Mt but independent. In summary our investigation suggests that for large ν_1 and ν_2 the results under the independent Mt assumption are qualitatively similar to those under normal errors (*e.g.* Toyoda and Wallace (1975)) - no strictly dominating estimator exists; the APE has the smallest risk around the neighbourhood of the null hypothesis; and the PTE with c=1 strictly dominates the NPE. Secondly, our results suggest that for $\nu_2 \geq \nu_1$ we have similar qualitative conclusions to Giles (1992) - neither the NPE nor the APE can strictly dominate the other; and the PTE with c=1 can strictly dominate both the APE and the NPE. Finally, if $\nu_1 > \nu_2$ the NPE strictly dominates the APE, but both are then strictly dominated by the PTE which uses c=1. So, the optimal strategy when $\nu_1 > \nu_2$ is to pre-test with c=1.

There remain a number of issues for future work. For example, it would be interesting to consider this problem with a different variance mixing distribution[4]; to consider other pre-test problems under a similar disturbance assumption as used here; to investigate the choice of an optimal critical value according to some explicit optimality criterion; to assume that the disturbances are non-normal though identically independently distributed; and to consider the case where ν_1 and ν_2 are estimated rather than assumed to be known.

APPENDIX

Proof of Theorem 1.

$$R(s_N^2) = E\left(s_1^2 - \sigma_{\varepsilon_1}^2\right)^2.$$

Now,

$$f(\varepsilon_1) = \int_0^\infty f_N(\varepsilon_1) f(\tau_1) d\tau_1 \tag{A.1}$$

where $f_N(\varepsilon_1)$ is the pdf of ε_1 when $\varepsilon_1 \sim N(0, \tau_1^2 I_{T_1})$ with τ_1^2 a positive scalar. (A.1) is the density of a Mt random variable when τ_1 is an inverted gamma variate. Then,

$$f(\tau_1) = \left[\frac{2}{\Gamma(\nu_1/2)}\right]\left(\frac{\nu_1\sigma_1^2}{2}\right)^{\nu_1/2} \tau_1^{-(\nu_1+1)} e^{-\nu_1\sigma_1^2/2\tau_1^2},$$

and we write $\tau_1 \sim IG(\nu_1, \sigma_1^2)$. So,

$$R(s_N^2) = E(s_1^4) - 2E(\tau_1^2)E(s_1^2) + \left(E(\tau_1^2)\right)^2.$$

Now,

$$E(s_1^2) = \int_0^\infty E_N(s_1^2) f(\tau_1) d\tau_1$$

where $E_N(A)$ is the expected value of A under the normality assumption. As $s_1^2 = \varepsilon_1' M_1 \varepsilon_1$, $M_1 = I_{T_1} - X_1(X_1'X_1)^{-1}X_1'$, we have $\varepsilon_1' M_1 \varepsilon_1 / \tau_1^2 \sim \chi^2_{v_1}$ under the assumption that $\varepsilon_1 \sim N(0, \tau_1^2 I_{T_1})$. So, $E_N(s_1^2) = \tau_1^2$ and $E(s_1^2) = E(\tau_1^2)$.

Likewise, $E(s_1^4) = (v_1+2)E(\tau_1^4)/v_1$. Then,

$$R(s_N^2) = \left[(v_1+2)E(\tau_1^4) - v_1\left(E(\tau_1^2)\right)^2\right]/v_1$$

and equation (5) follows directly as $E(\tau_1^2) = \nu_1\sigma_1^2/(\nu_1-2)$ and $E(\tau_1^4) = \nu_1^2\sigma_1^4/\left((\nu_1-2)(\nu_1-4)\right)$ when $\tau_1 \sim IG(\nu_1, \sigma_1^2)$.

Similarly,

$$R(s_A^2) = E\left(s_A^2-\sigma_{\varepsilon_1}^2\right)^2$$

$$= E(s_A^4) - 2E(\tau_1^2)E(s_A^2) + \left(E(\tau_1^2)\right)^2.$$

Now,

$$s_A^2 = (\varepsilon_1' M_1\varepsilon_1+\varepsilon_2' M_2\varepsilon_2)/(v_1+v_2)$$

where M_2 is defined analogously to M_1. Then,

$$E(s_A^2) = \int_0^\infty\int_0^\infty E_N(s_A^2)f(\tau_1\tau_2)d\tau_1 d\tau_2$$

$$= \int_0^\infty\int_0^\infty E_N(s_A^2)f(\tau_1)f(\tau_2)d\tau_1 d\tau_2$$

as ε_1 and ε_2 are independent. As $\varepsilon_1' M_1\varepsilon_1/\tau_1^2 \sim \chi_{v_1}^2$ and $\varepsilon_2' M_2\varepsilon_2/\tau_2^2 \sim \chi_{v_2}^2$ when $\varepsilon_1 \sim N(0,\ \tau_1^2 I_{T_1})$ and $\varepsilon_2 \sim N(0,\ \tau_2^2 I_{T_2})$ respectively, we have that

$$E_N(s_A^2) = (v_1\tau_1^2+v_2\tau_2^2)/(v_1+v_2)$$

and

$$E(s_A^2) = \frac{v_1}{v_1+v_2}\int_0^\infty\int_0^\infty \tau_1^2 f(\tau_1)d\tau_1 f(\tau_2)d\tau_2$$

$$+ \frac{v_2}{v_1+v_2}\int_0^\infty\int_0^\infty \tau_2^2 f(\tau_1)d\tau_1 f(\tau_2)d\tau_2$$

$$= \left[v_1 E(\tau_1^2)+v_2 E(\tau_2^2)\right]/(v_1+v_2)$$

$$= \left[v_1 E(\tau_1^2)+v_2 E(\tau_1^2)/\phi\right]/(v_1+v_2)$$

as $\phi = \sigma_{\varepsilon_1}^2/\sigma_{\varepsilon_2}^2$ and $E(\tau_2^2) = \sigma_{\varepsilon_2}^2 = E(\tau_1^2)/\phi$.

Likewise,

$$E(s_A^4) = \Big\{v_1(v_1+2)E(\tau_1^4)+v_2(v_2+2)E(\tau_2^2)$$

$$+ 2v_1v_2E(\tau_1^2)E(\tau_2^2)\Big\}/(v_1+v_2)^2$$

$$= \Big\{v_1(v_1+2)E(\tau_1^4)+v_2(v_2+2)E(\tau_2^4)$$

$$+ (v_2^2-v_1^2)\Big(E(\tau_1^2)\Big)^2+2v_2^2\Big(E(\tau_1^2)\Big)^2/\phi\Big\}/(v_1+v_2)^2.$$

Using $\tau_1 \sim IG(\nu_1, \sigma_1^2)$, $\tau_2 \sim IG(\nu_2, \sigma_2^2)$ then

$$E(s_A^4) = \left\{\frac{v_1(v_1+2)\nu_1^2\sigma_1^4}{(\nu_1-2)(\nu_1-4)} + \frac{v_2(v_2+2)\nu_1^2(\nu_2-2)\sigma_1^4}{(\nu_1-2)^2(\nu_2-4)\phi^2}\right.$$

$$\left. + (v_2^2-v_1^2)\frac{\nu_1^2\sigma_1^4}{(\nu_1-2)^2} - 2v_2^2\frac{(\nu_1^2\sigma_1^4)}{(\nu_1-2)^2}\right\}/(v_1+v_2)^2$$

so that $R(s_A^2)$ follows.

We now turn to the risk of the pre-test estimator:

$$R(s_P^2) = E(s_P^4) - 2E(\tau_1^2)E(s_P^2) + \Big(E(\tau_1^2)\Big)^2. \tag{A.2}$$

We have

$$s_P^2 = s_N^2 + (s_A^2-s_N^2)I_{[0,c]}(J)$$

$$= (\varepsilon_1'M_1\varepsilon_1)(v_1+v_2) + (v_1\varepsilon_2'M_2\varepsilon_2-v_2\varepsilon_1'M_1\varepsilon_1)$$

$$\times I_{[0,c]}(v_1\varepsilon_2'M_2\varepsilon_2/v_2\varepsilon_1'M_1\varepsilon_1)$$

where $I_{[a,b]}(J)$ is an indicator function which takes the value one when J lies within the subscripted range, 0 otherwise. So,

$$R(s_P^2) = \Big\{\tau_1^2(v_1+v_2)(\varepsilon_1'M_1\varepsilon_1/\tau_1^2)$$

$$+ \Big[\tau_2^2v_1(\varepsilon_2'M_2\varepsilon_2/\tau_2^2)-v_2\tau_1^2(\varepsilon_1'M_1\varepsilon_1/\tau_1^2)\Big]$$

$$\times I_{[0,c^*]}\Big((v_1\varepsilon_2'M_2\varepsilon_2/\tau_2^2)/(v_2\varepsilon_1'M_1\varepsilon_1/\tau_1^2)\Big)\Big\}/\Big(v_1(v_1+v_2)\Big)$$

where $c^* = c\tau_1^2/\tau_2^2$.

Under the assumption that $\varepsilon_1 \sim N(0, \tau_1^2I_{T_1})$ and $\varepsilon_2 \sim N(0, \tau_2^2I_{T_2})$ and using, for example, Lemma 1 of Clarke *et al.* (1987) we have

$$E_N\left[(\varepsilon_2' M_2 \varepsilon_2/\tau_2^2) I_{[0,c*]}(.)\right] = v_2 P_{20} ,$$

$$E_N\left[(\varepsilon_1' M_1 \varepsilon_1/\tau_1^2) I_{[0,c*]}(.)\right] = v_1 P_{02}$$

where

$$P_{ij} = \Pr.\left[F_{(v_2+i, v_1+j)} \leq \left(c^* v_2(v_1+j)\right)/\left(v_1(v_2+i)\right)\right]$$

$$= I_x\left[\tfrac{1}{2}(v_2+i);\tfrac{1}{2}(v_1+j)\right]$$

and $x = c\tau_1^2 v_2/(v_1\tau_2^2 + cv_2\tau_1^2)$. $I_x(.;.)$ is the incomplete beta function. So,

$$E(s_P^2) = \left\{(v_1+v_2)E(\tau_1^2) + v_2E(\tau_2^2P_{20}) - v_2E(\tau_1^2P_{02})\right\}/(v_1+v_2). \tag{A.3}$$

Following similar steps,

$$E(s_P^2) = \Big\{(v_1+v_2)^2(v_1+2)E(\tau_1^4) + v_1v_2(v_2+2)E(\tau_2^4P_{40})$$

$$- v_2(2v_1+v_2)(v_1+2)E(\tau_1^4P_{04}) + 2v_1^2v_2E(\tau_1^2\tau_2^2P_{22})\Big\}/\left(v_1(v_1+v_2)^2\right). \tag{A.4}$$

Substituting (A.2) and (A.3) into (A.1) we have

$$R(s_P^2) = \Big\{(v_1+v_2)^2(v_1+2)E(\tau_1^4) + v_1v_2(v_2+2)E(\tau_2^4P_{40})$$

$$- v_2(2v_1+v_2)(v_1+2)E(\tau_1^4P_{04}) + 2v_1^2v_2E(\tau_1^2\tau_2^2P_{22})$$

$$- 2E(\tau_1^2)v_1(v_1+v_2)\left[(v_1+v_2)E(\tau_1^2) + v_2E(\tau_2^2P_{20}) - v_2E(\tau_1^2P_{02})\right]$$

$$+ v_1(v_1+v_2)^2\left(E(\tau_1^2)\right)^2\Big\}/\left(v_1(v_1+v_2)^2\right). \tag{A.5}$$

To evaluate (A.5) under the inverted gamma assumptions we require the following lemma:

Lemma:

$$E(\tau_1^a\tau_2^b P_{ij}) = (\nu_1^2\sigma_1^2)^{(a+b)/2} P_{abij}$$

where

$$P_{abij} = \frac{\Gamma\left(\dfrac{\nu_1+\nu_2-a-b}{2}\right)}{\Gamma\left(\dfrac{\nu_1}{2}\right)\Gamma\left(\dfrac{\nu_2}{2}\right)} \, 2^{-(a+b)/2}\left(\frac{\nu_2-2}{\phi}\right)^{\nu_2/2}$$

$$\times\ v_1^{(\nu_2-b)/2}(cv_2)^{(\nu_1-a)/2}(\nu_1-2)^{(\nu_1-a-b)/2}$$

$$\times \int_0^\infty t_1^{(\nu_2-b)/2-1}(1-t_1)^{(\nu_1-a)/b-1}\Big[v_1 t_1(\nu_2-2)/\phi$$

$$+\ (\nu_1-2)cv_2(1-t_1)\Big]^{(a+b-\nu_1-\nu_2)/2} I_{t_1}\left(\tfrac{1}{2}(v_2+i);\tfrac{1}{2}(v_1+j)\right)dt_1.$$

Proof.

$$E(\tau_1^a\tau_2^b P_{ij}) = \int_0^\infty\int_0^\infty \tau_1^a\tau_2^b I_x\left[\tfrac{1}{2}(v_2+i);\tfrac{1}{2}(v_1+j)\right]$$

$$\times\ f(\tau_1)f(\tau_2)d\tau_1 d\tau_2\ .$$

Substituting in $f(\tau_1)$ and $f(\tau_2)$ we have

$$E(\tau_1^a\tau_2^b P_{ij}) = \frac{4}{\Gamma\left(\dfrac{\nu_1}{2}\right)\Gamma\left(\dfrac{\nu_2}{2}\right)}\left(\frac{\nu_1\sigma_1^2}{2}\right)^{\nu_1/2}\left(\frac{\nu_2\sigma_2^2}{2}\right)^{\nu_2/2}$$

$$\times \int_0^\infty\int_0^\infty \tau_1^{a-\nu_1-1}\tau_2^{b-\nu_2-1}\exp\left[-(\nu_1\sigma_1^2/\tau_1^2+\nu_2\sigma_2^2/\tau_2^2)/2\right]$$

$$\times\ I_x\left[\tfrac{1}{2}(v_2+i);\tfrac{1}{2}(v_1+j)\right]d\tau_1 d\tau_2\ .$$

Note that $I_x(.)$ does not depend on τ_1 or τ_2. Now, using the following changes of variables:

1. $z_1 = \tau_1^2,\ z_2 = \tau_2^2$

2. $t_1 = cv_2z_1/(cv_2z_1+v_1z_2)$, $t_2 = z_2$

3. $s = p/(2t_2)$, $t_1 = t_1$

where $p = \nu_2\sigma_2^2 + \nu_1\sigma_1^2 cv_2(1-t_1)/(v_1t_1)$, it follows that

$$E(\tau_1^a\tau_2^b P_{ij}) = \frac{\Gamma\left(\frac{\nu_1+\nu_2-a-b}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \left(\frac{\nu_1\sigma_1^2}{2}\right)^{(a+b)/2} \left(\frac{\nu_2-2}{\phi}\right)^{\nu_2/2}$$

$$\times\ v_1^{(\nu_2-b)/2}(cv_2)^{(\nu_1-a)/2}(\nu_1-2)^{(\nu_1-a-b)/2}$$

$$\times \int_0^1 t_1^{(v_2-b)/2-1}(1-t_1)^{(\nu_1-a)/b-1}\Big[v_1t_1(\nu_2-2)/\phi$$

$$+ (\nu_1-2)cv_2(1-t_1)\Big]^{(a+b-\nu_1-\nu_2)/2} I_{t_1}\left(\tfrac{1}{2}(v_2+i);\tfrac{1}{2}(v_1+j)\right)dt_1$$

$$= (\nu_1\sigma_1^2)^{(a+b)/2}P_{abij}\ .$$

Using this Lemma repeatedly in (A.5) gives the desired result. #

Proof of Theorem 2.

Now,

$$R(s_P^2) = E\left[s_A^2 I_{[0,c]}(J)+s_N^2 I_{(c,\infty)}(J)-E(\tau_1^2)\right]^2$$

$$= E\left[\left(s_A^2-E(\tau_1^2)\right)^2 I_{[0,c]}(J) + \left(s_N^2-E(\tau_1^2)\right)^2 I_{(c,\infty)}(J)\right]\ .$$

Let $q_1 = \varepsilon_1' M_1\varepsilon_1$ and $q_2 = \varepsilon_2' M_2\varepsilon_2$ so that

$$s_A^2 = (q_1+q_2)/(v_1+v_2),\ s_N^2 = q_1/v_1;\ J = v_1q_2/(v_2q_1) \text{ and}$$

$$R(s_P^2) = E\left\{\left[(q_1+q_2)/(v_1+v_2)-E(\tau_1^2)\right]^2 I\left[q_2 \le cv_2q_1/v_1\right]\right.$$

$$+ \left(q_1/v_1 - E(\tau_1^2)\right)^2 \left(1 - I\left[q_2 \le cv_2 q_1/v_1\right]\right)\Bigg\}$$

$$= \int_0^\infty\int_0^\infty E_N\left[\left((q_1+q_2)/(v_1+v_2) - E(\tau_1^2)\right)^2 I\left[q_2 \le cv_2q_1/v_1\right]\right.$$

$$\left. + \left(q_1/v_1 - E(\tau_1^2)\right)^2 \left(1 - I\left[q_2 \le cv_2 q_1/v_1\right]\right)\right\} f(\tau_1)f(\tau_2)d\tau_1 d\tau_2$$

$$= \int_0^\infty\int_0^\infty E_{N_{q_1}}\left[E_{N_{q_2}}\left((q_1+q_2)/(v_1+v_2) - E(\tau_1^2)\right)^2 I\left[q_2 \le cv_2q_1/v_1\right]\right.$$

$$\left. + \left(q_1/v_1 - E(\tau_1^2)\right)^2 \left(1 - I(q_2 \le cv_2 q_1/v_1)\right)\right]$$

$$\times f(\tau_1)f(\tau_2)d\tau_1 d\tau_2$$

$$= \int_0^\infty\int_0^\infty E_{N_{q_1}}\left\{\int_0^x \left((q_1+q_2)/(v_1+v_2) - E(\tau_1^2)\right)^2 f_N(q_2)dq_2\right.$$

$$\left. + \left(q_1/v_1 - E(\tau_1^2)\right)^2 \left(1 - \int_0^x f_N(q_2)dq_2\right)\right\} f(\tau_1)f(\tau_2)d\tau_1 d\tau_2$$

where $x = cv_2q_1/v_1$ and $f_N(q_2)$ is the density function of a Chi-squared random variable with v_2 degrees of freedom.

So,

$$\frac{\partial R(s_P^2)}{\partial c} = \int_0^\infty\int_0^\infty E_{N_{q_1}}\left\{\frac{\partial x}{\partial c}\left[\frac{\partial}{\partial x}\int_0^x \left((q_1+q_2)/(v_1+v_2) - E(\tau_1^2)\right)^2 f_N(q_2)dq_2\right.\right.$$

$$\left.\left. - \left(q_1/v_1 - E(\tau_1^2)\right)^2 \frac{\partial}{\partial x}\int_0^x f_N(q_2)dq_2\right]\right\}$$

$$\times f(\tau_1)f(\tau_2)d\tau_1 d\tau_2$$

$$= \int_0^\infty\int_0^\infty E_{N_{q_1}}\left\{\left(\frac{v_2q_1}{v_1}\right) f_N\left(\frac{cv_2q_1}{v_1}\right)\right.$$

$$\times \left[\left((q_1+cv_2q_1/v_1)/(v_1+v_2) - E(\tau_1^2)\right)^2\right.$$

$$- \left(q_1/v_1 - E(\tau_1)\right)^2\Big]\Big\} f(\tau_1)f(\tau_2)d\tau_1 d\tau_2 \ .$$

A sufficient condition for this derivative to be zero is c=0, $c \to \infty$, and

$$\left[\left((q_1+cv_2q_1/v_1)/(v_1+v_2)-E(\tau_1^2)\right)^2 - \left((q_1/v_1)-E(\tau_1^2)\right)^2\right] = 0. \qquad \text{(A.5)}$$

This will hold if c=1 or $1-2v_1(v_1+v_2)\left((q_1/v_1)-E(\tau_1^2)\right)^2/(q_1v_2) = 0$. It does not appear possible to sign the second partial derivative with respect to c and so it was not possible to analytically confirm whether c=1 corresponds to a minimum or a maximum. Our numerical evaluations suggested a minimum.

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.

Bancroft, T.A. and Han, C-P. (1983). A note on pooling variances. *Journal of the American Statistical Association*, 78, 981-983.

Chmielewski, M.A. (1981). Invariant tests for the equality of K scale parameters under spherical symmetry. *Journal of Statistical Planning and Inference*, 5, 341-346.

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

Giles, J.A. (1990). Preliminary test estimation of a mis-specified linear model with spherically symmetric disturbances. Ph.D. thesis, University of Canterbury.

Giles, J.A. (1992). Estimation of the error variance after a preliminary test of homogeneity in a regression model with spherically symmetric disturbances. *Journal of Econometrics*, 53, 345-361.

Giles, J.A. and D.E.A. Giles (1993). Pre-test estimation and testing in econometrics: Recent developments. *Journal of Economic Surveys*, forthcoming.

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

Ohtani, K. (1990). Testing equality of scale parameters between two linear regressions when error terms have multivariate distributions and two error terms are mutually independent. Mimeo., Faculty of Economics, Kobe University.

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

Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1986). *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, New York.

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

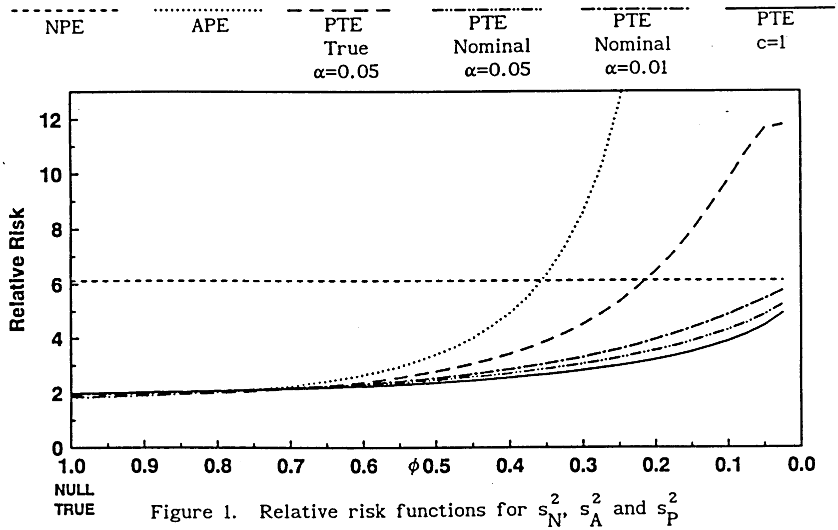

Figure 1. Relative risk functions for s_N^2, s_A^2 and s_P^2

$v_1 = v_2 = 30$, $\nu_1 = 5$, $\nu_2 = 10$

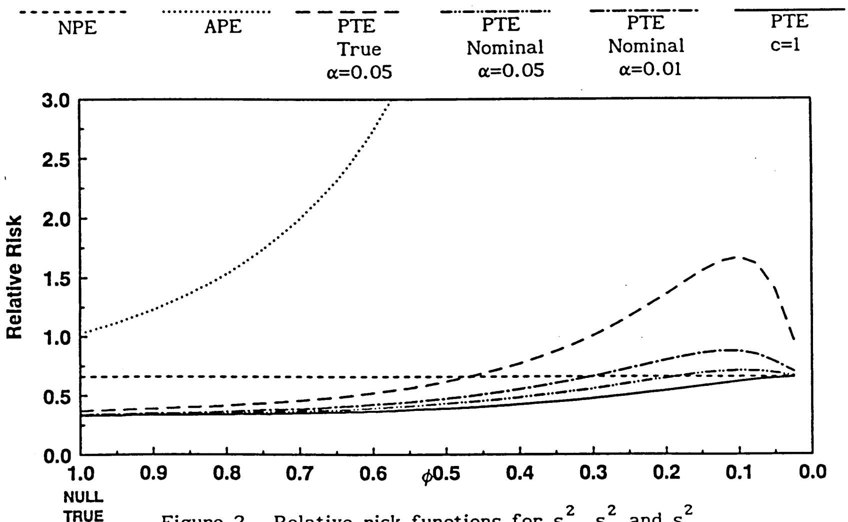

Figure 2. Relative risk functions for s_N^2, s_A^2 and s_P^2

$v_1 = v_2 = 30$, $\nu_1 = 10$, $\nu_2 = 5$

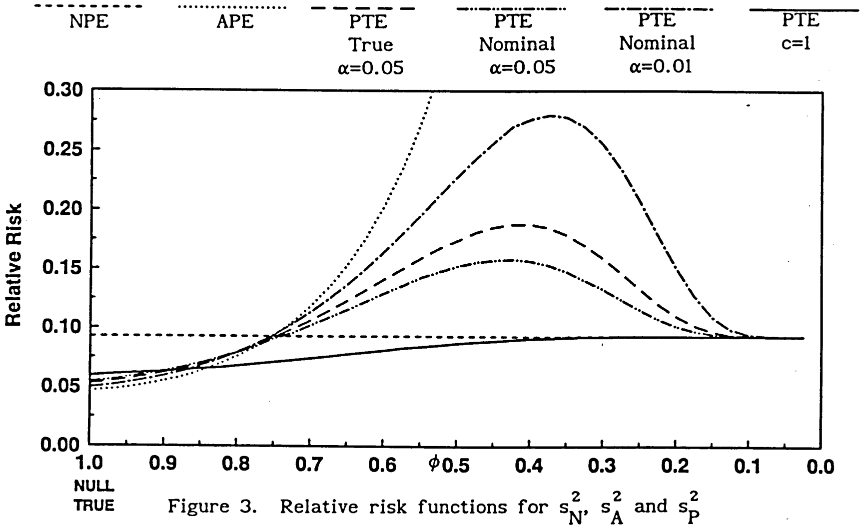

Figure 3. Relative risk functions for s_N^2, s_A^2 and s_P^2

$v_1 = v_2 = 30,\ \nu_1 = \nu_2 = 100$

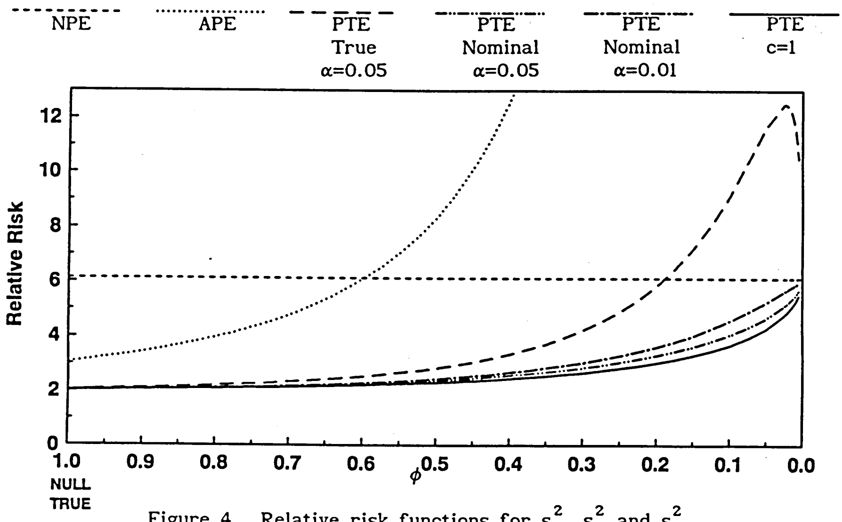

Figure 4. Relative risk functions for s_N^2, s_A^2 and s_P^2

$v_1 = v_2 = 30,\ \nu_1 = \nu_2 = 5$

FOOTNOTES

* The authors are grateful to David Giles for his helpful comments and suggestions.

1. For simplicity we assume a one-sided alternative hypothesis. It is straightforward, though tedious, to extend our analysis to the two-sided case.

2. See Giles and Giles (1993) for a survey of this literature.

3. Note that although the risks under normality are a special case of our results, those of Giles (1992) are not.

4. For example, Giles (1992) also considers the case where the mixing distribution is gamma.

LIST OF DISCUSSION PAPERS*

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 Estimator, 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, by Judith A. Giles.
No. 9001	The Durbin-Watson Test for Autocorrelation in Nonlinear Models, by Kenneth J. White.
No. 9002	Determinants of Aggregate Demand for Cigarettes in New Zealand, by Robin Harrison and Jane Chetwyd.
No. 9003	Unemployment Duration and the Measurement of Unemployment, by Manimay Sengupta.
No. 9004	Estimation of the Error Variance After a Preliminary-Test of Homogeneity in a Regression Model with Spherically Symmetric Disturbances, by Judith A. Giles.
No. 9005	An Expository Note on the Composite Commodity Theorem, by Michael Carter.
No. 9006	The Optimal Size of a Preliminary Test of Linear Restrictions in a Mis-specified Regression Model, by David E. A. Giles, Offer Lieberman, and Judith A. Giles.
No. 9007	Inflation, Unemployment and Macroeconomic Policy in New Zealand: A Public Choice Analysis, by David J. Smyth and Alan E. Woodfield.
No. 9008	Inflation — Unemployment Choices in New Zealand and the Median Voter Theorem, by David J. Smyth and Alan E. Woodfield.
No. 9009	The Power of the Durbin-Watson Test when the Errors are Heteroscedastic, by David E. A. Giles and John P. Small.
No. 9010	The Exact Distribution of a Least Squares Regression Coefficient Estimator After a Preliminary t-Test, by David E. A. Giles and Virendra K. Srivastava.
No. 9011	Testing Linear Restrictions on Coefficients in a Linear Regression Model with Proxy variables and Spherically Symmetric Disturbances, by Kazuhiro Ohtani and Judith A. Giles.
No. 9012	Some Consequences of Applying the Goldfeld-Quandt Test to Mis-Specified Regression Models, by David E. A. Giles and Guy N. Saxton.
No. 9013	Pre-testing in a Mis-specified Regression Model, by Judith A. Giles.
No. 9014	Two Results in Balanced-Growth Educational Policy, by Alan E. Woodfield.
No. 9101	Bounds on the Effect of Heteroscedasticity on the Chow Test for Structural Change, by David Giles and Offer Lieberman.
No. 9102	The Optimal Size of a Preliminary Test for Linear Restrictions when Estimating the Regression Scale Parameter, by Judith A. Giles and Offer Lieberman.
No. 9103	Some Properties of the Durbin-Watson Test After a Preliminary t-Test, by David Giles and Offer Lieberman.
No. 9104	Preliminary-Test Estimation of the Regression Scale Parameter when the Loss Function is Asymmetric, by Judith A. Giles and David E. A. Giles.
No. 9105	On an Index of Poverty, by Manimay Sengupta and Prasanta K. Pattanaik.
No. 9106	Cartels May Be Good For You, by Michael Carter and Julian Wright.
No. 9107	L_p-Norm Consistencies of Nonparametric Estimates of Regression, Heteroskedasticity and Variance of Regression Estimate when Distribution of Regression is Known, by Radhey S. Singh.

(Continued on next page)

No. 9108 Optimal Telecommunications Tariffs and the CCITT, by Michael Carter and Julian Wright.

No. 9109 Price Indices : Systems Estimation and Tests, by David Giles and Ewen McCann.

No. 9110 The Limiting Power of Point Optimal Autocorrelation Tests, by John P. Small.

No. 9111 The Exact Power of Some Autocorrelation Tests When the Disturbances are Heteroscedastic, by John P. Small.

No. 9112 Some Consequences of Using the Chow Test in the Context of Autocorrelated Disturbances, by David Giles and Murray Scott.

No. 9113 The Exact Distribution of R^2 when the Disturbances are Autocorrelated, by Mark L. Carrodus and David E. A. Giles.

No. 9114 Optimal Critical Values of a Preliminary Test for Linear Restrictions in a Regression Model with Multivariate Student-t Disturbances, by Jason K. Wong and Judith A. Giles.

No. 9115 Pre-Test Estimation in a Regression Model with a Misspecified Error Covariance Matrix, by K. V. Albertson.

No. 9116 Estimation of the Scale Parameter After a Pre-test for Homogeneity in a Mis-specified Regression Model, by Judith A. Giles.

No. 9201 Testing for Arch-Garch Errors in a Mis-specified Regression, by David E. A. Giles, Judith A. Giles, and Jason K. Wong.

No. 9202 Quasi Rational Consumer Demand — Some Positive and Normative Surprises, by John Fountain.

No. 9203 Pre-test Estimation and Testing in Econometrics: Recent Developments, by Judith A. Giles and David E. A. Giles.

No. 9204 Optimal Immigration in a Model of Education and Growth, by K-L. Shea and A. E. Woodfield.

No. 9205 Optimal Capital Requirements for Admission of Business Immigrants in the Long Run, by K-L. Shea and A. E. Woodfield.

No. 9206 Causality, Unit Roots and Export-Led Growth: The New Zealand Experience, by David E. A. Giles, Judith A. Giles and Ewen McCann.

No. 9207 The Sampling Performance of Inequality Restricted and Pre-Test Estimators in a Mis-specified Linear Model, by Alan T. K. Wan.

No. 9208 Testing and Estimation with Seasonal Autoregressive Mis-specification, by John P. Small.

No. 9209 A Bargaining Experiment, by Michael Carter and Mark Sunderland.

No. 9210 Pre-Test Estimation in Regression Under Absolute Error Loss, by David E. A. Giles.

No. 9211 Estimation of the Regression Scale After a Pre-Test for Homoscedasticity Under Linex Loss, by Judith A. Giles and David E. A. Giles.

No. 9301 Assessing Starmer's Evidence for New Theories of Choice: A Subjectivist's Comment, by John Fountain.

No. 9302 Preliminary-Test Estimation in a Dynamnic Linear Model, by David E. A. Giles and Matthew C. Cunneen.

No. 9303 Fans, Frames and Risk Aversion: How Robust is the Common Consequence Effect? by John Fountain and Michael McCosker.

No. 9304 Pre-test Estimation of the Regression Scale Parameter with Multivariate Student-t Errors and Independent Sub-Samples, by Juston Z. Anderson and Judith A. Giles

* 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. A list of the Discussion Papers prior to 1989 is available on request.