Department of Economics
UNIVERSITY OF CANTERBURY
CHRISTCHURCH, NEW ZEALAND

ESTIMATION OF THE ERROR VARIANCE AFTER A PRELIMINARY-TEST OF HOMOGENEITY IN A REGRESSION MODEL WITH SPHERICALLY SYMMETRIC DISTURBANCES

JUDITH A. GILES

Discussion Paper

No. 9004

Department of Economics, University of Canterbury
Christchurch, New Zealand

Discussion Paper No. 9004

July 1990

ESTIMATION OF THE ERROR VARIANCE AFTER A PRELIMINARY-TEST OF HOMOGENEITY IN A REGRESSION MODEL WITH SPHERICALLY SYMMETRIC DISTURBANCES

JUDITH A. GILES

Department of Economics
University of Canterbury
Christchurch, New Zealand

ESTIMATION OF THE ERROR VARIANCE AFTER A PRELIMINARY-TEST OF HOMOGENEITY IN A REGRESSION MODEL WITH SPHERICALLY SYMMETRIC DISTURBANCES

Judith A. Giles*

University of Canterbury

ABSTRACT

In this paper we consider the risk (under quadratic loss) of an estimator of the error variance after a pre-test for homogeneity of the variances in the two-sample linear regression model. We investigate the effects on risk of assuming normal disturbances when in fact the error distribution is spherically symmetric. We also broaden the standard assumption that the never-pool variance estimators are based on the least squares principle. Using the special case of multivariate Student-t regression disturbances as an illustration, our results show that in some situations we should always pre-test, *even if the error variances are equal*, and we provide the optimal test critical value. The evaluations also show that using the least squares technique to form the never-pool estimators may not always be the preferred strategy.

*The author is grateful to David Giles for many helpful discussions and suggestions.

Correspondence:

Judith A. Giles
Department of Economics
University of Canterbury
Private Bag, Christchurch 1
NEW ZEALAND

1. Introduction and estimator definitions

We consider a regression model which uses two samples, with T_1 and T_2 observations ($T_1+T_2=T$). There is a common location vector, β, but possibly different error variances:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}\beta + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} , \qquad (1)$$

or more compactly, $y=X\beta+e$, where y_i is a $(T_i \times 1)$ vector of observations on the dependent variable, X_i is a non-stochastic $(T_i \times k)$ design matrix of rank k $(<T_i)$, and e_i is a $(T_i \times 1)$ vector of regression disturbances, i=1,2. We assume that $E(e_i)=0$, and that $E(e_i e_i')=\sigma^2_{e_i} I_{T_i}$. Let $\phi=\sigma^2_{e_1}/\sigma^2_{e_2}$, and

$$E(ee') = \begin{bmatrix} \sigma^2_{e_1} I_{T_1} & 0 \\ 0 & \sigma^2_{e_2} I_{T_2} \end{bmatrix} = \sigma^2_{e_2} \begin{bmatrix} \phi I_{T_1} & 0 \\ 0 & I_{T_2} \end{bmatrix} = \sigma^2_{e_2} \Sigma \ . \qquad (2)$$

Suppose that e has a non-normal distribution of the form $f(e)={}_0\int^\infty f_N(e)f(\tau)d\tau$, where $f_N(e)$ is $f(e)$ when $e\sim N(0,\tau^2\Sigma)$,and $f(\tau)$ is supported on $[0,\infty)$. Hence, $\sigma^2_{e_2}=E(\tau^2)$ and $\sigma^2_{e_1}=\phi E(\tau^2)$. So, (Muirhead (1982)) the joint distribution of the disturbances in each sample is spherically symmetric, while e has an elliptically symmetric distribution (ESD) when $\sigma^2_{e_1}\neq\sigma^2_{e_2}$, and a spherically symmetric distribution (SSD) when the variances are equal.[1,2] The normal and multivariate Student-t (Mt) distributions are well known members of this family. The latter distribution arises if $f(\tau)$ is an inverted gamma density with, say, scale parameter σ^2_2 and degrees of freedom parameter ν; we write $\tau\sim IG(\sigma^2_2,\nu)$ and $e\sim Mt\left(0,\nu\sigma^2_2/(\nu-2)\Sigma\right)$. The marginal distributions are univariate Student-t and for small values of ν they will have thicker tails than under normality. When $\nu=1$ the probability density function (pdf) is Cauchy while it is normal when $\nu=\infty$.[3]

We consider the estimation of $\sigma^2_{e_1}$ given uncertainty about whether the second sample comes from the same population as the first when the researcher wrongly assumes that $e \sim N\left(0, \sigma^2_{e_2}\Sigma\right)$. The pre-test under investigation is of

$$H_0 : \phi = 1 \quad \text{vs} \quad H_1 : \phi < 1 . \tag{3}$$

We assume for simplicity a one-sided alternative hypothesis, though the analysis can be extended to the two-sided case.[4]

The research on this particular pre-test problem has followed the literature associated with the pooling of two normal samples. If the variances are unequal then an unbiased (never-pool) estimator of $\sigma^2_{e_1}$ is $s^2_{NL}=s^2_1$ where $s^2_i=(y_i-X_ib_i)'(y_i-X_ib_i)/v_i=e_i'M_ie_i/v_i$, $v_i=T_i-k$, $M_i=I_{T_i}-X_iS_i^{-1}X_i'$, and $b_i=S_i^{-1}X_i'y_i$, $S_i=(X_i'X_i)$, $i=1,2$. s^2_i is the usual unbiased least squares (L) estimator of $\sigma^2_{e_i}$. Conversely, if the variances are equal then the two samples may be pooled and an unbiased (always-pool) estimator of $\sigma^2_{e_1}$ is $s^2_{AL}=\left(v_1s^2_1+v_2s^2_2\right)/(v_1+v_2)$. A test statistic for homoscedasticity is $J=\left(v_1e_2'M_2e_2\right)/\left(v_2e_1'M_1e_1\right)$, and $f(J)=\phi^{-1}f\left(F_{(v_2,v_1)}\right)$, where $F_{(v_2,v_1)}$ is a central F random variate with v_2 and v_1 degrees of freedom.

The researcher tests H_0 using J and so he employs the pre-test estimator

$$s^2_P = \begin{cases} s^2_{NL} & \text{if } J > c \\ s^2_{AL} & \text{if } J \le c \end{cases} = I_{[0,c]}(J)s^2_{AL} + I_{(c,\infty)}(J)s^2_{NL} , \tag{4}$$

where $I_{[.,.]}(J)$ takes the value unity if J lies within [.,.], zero otherwise, and c is the critical value of the test corresponding to a test size of α. Assuming normal errors and the least squares based estimators s^2_{NL} and s^2_{AL}, Bancroft (1944), Toyoda and Wallace (1975), Ohtani and Toyoda (1978), and Bancroft and Han (1983) have examined the sampling properties of s^2_P.[5] We generalise both of these assumptions in this paper.

There is substantial empirical evidence to support the possibility that

There is substantial empirical evidence to support the possibility that some economic series may be generated by processes whose distributions have more kurtosis than the normal distribution (e.g., Mandelbrot (1963, 1967), Fama (1963, 1965), Blattberg and Gonedes (1974), Praetz (1972), and Praetz and Wilson (1978)). The family of distributions that we investigate is motivated by some of these studies, and by those papers which have studied linear regression models with ESD (or SSD) disturbances, including Thomas (1970), King (1979), Chmielewski (1981), Judge, Miyazaki and Yancey (1985), Ullah and Zinde-Walsh (1985), Zinde-Walsh and Ullah (1987), and Giles (1990b). In particular, Chmielewski (1981) shows that f(J) holds for all members of the elliptically symmetric family[6], and Giles (1990b) derives 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 when the pre-test is for the validity of a set of exact linear restrictions on the coefficient vector. She finds that when estimating the conditional forecast of y the widening of the error distribution assumption has little impact on the qualitative properties of the risk function of the predictor pre-test estimator, though there are quantitative effects. In contrast, there can be a substantial impact on the risk functions of the estimators of the error variance. Specifically, she shows that pre-testing, with a critical value of unity, is the preferable strategy when the disturbances are Mt with small ν.[7]

In this paper we extend the aforementioned pre-testing literature first by deriving the risk functions of a family of pre-test estimators of $\sigma^2_{e_1}$ after a pre-test for H_0, when the researcher wrongly assumes that the sub-sample spherically symmetric disturbances are normally distributed. Our second extension relates to the component estimators under investigation. To date the research in this area has only considered the pre-test estimator based on s^2_{NL} and s^2_{AL} while, within the linear regression model framework, two other never-pool estimators of $\sigma^2_{e_i}$ are commonly used (assuming normality): the maximum likelihood (ML) estimator and the minimum mean squared error (M) estimator. Let these estimators be denoted by s^2_{iML} and s^2_{iM} respectively. They differ from s^2_{iL} by the divisor used, this being T_i for the ML estimators and (v_i+2) for the M estimators. These estimators are members of the family

$$S^2_i = (e_i'M_ie_i)/(T_i+\mu). \tag{5}$$

We can generate s^2_{iL}, s^2_{iML}, and s^2_{iM} by setting μ to $-k$, 0, and $(-k+2)$, respectively. Let $S^2_N=S^2_1$ be the family of never-pool estimators of $\sigma^2_{e_1}$.

In the spirit of s^2_{AL} we can conceive of two alternative always-pool estimators s^2_{AML} and s^2_{AM}, which have as their components the sample ML and M estimators, s^2_{iML} and s^2_{iM}, respectively. That is, $s^2_{AML}=\left(T_1s^2_{1ML}+T_2s^2_{2ML}\right)/T$ and $s^2_{AM}=\left((v_1+2)s^2_{1M}+(v_2+2)s^2_{2M}\right)/(v_1+v_2+4)$. s^2_{AL}, s^2_{AML}, and s^2_{AM} are always-pool estimators of the form

$$S^2_A = \left[(T_1+\mu)\left(e_1'M_1e_1/(T_1+\mu)\right)+(T_2+\mu)\left(e_2'M_2e_2/(T_2+\mu)\right)\right]/(T+2\mu). \tag{6}$$

We obtain s^2_{AL}, s^2_{AML}, and s^2_{AM} by setting μ to $-k$, 0, and $(-k+2)$. Clearly s^2_{AML} is not a ML estimator, nor does s^2_{AM} possess the M property, even if the errors are normal. Of course, when the errors are non-normal then even the sample estimators are not the ML or M estimators, though the researcher proceeds assuming that they possess these properties.[8]

So, given (5) and (6), the pre-test estimator we investigate is

$$S^2_P = \begin{cases} S^2_N & \text{if } J > c \\ S^2_A & \text{if } J \le c \end{cases} = I_{[0,c]}(J)S^2_A + I_{(c,\infty)}(J)S^2_N \ . \tag{7}$$

In the next section we derive and discuss the risk functions of S^2_N, S^2_A, and S^2_P. To illustrate our results we consider exact evaluations of the risks for the special case of Mt regression disturbances in Section 3. The final section is devoted to concluding remarks.

2. The risk functions

Theorem 1 gives the risk functions of the estimators, where we define the risk of an estimator $\bar{s}^2$ of $\sigma^2_{e_1}$ as $\rho(\sigma^2_{e_1},\bar{s}^2)=E(\bar{s}^2-\sigma^2_{e_1})^2=E\left(\bar{s}^2-\phi E(\tau^2)\right)^2$.

Theorem 1. If e~ESD with E(e)=0, $E(ee')=\sigma^2_{e_2}\Sigma$ and the pre-test is of H_0 in (3), then

$$\rho(\sigma^2_{e_1},S^2_N) = \phi^2\left(v_1(v_1+2)E(\tau^4)+\left(E(\tau^2)\right)^2(T_1+\mu)(T_1+\mu-2v_1)\right)/(T_1+\mu)^2, \tag{8}$$

$$\begin{aligned}\rho(\sigma^2_{e_1},S^2_A) = &\left(\phi^2\left[v_1(v_1+2)E(\tau^4)+\left(E(\tau^2)\right)^2(T+2\mu)(T+2\mu-2v_1)\right]\right.\\ &\left.+ 2v_2\phi\left[v_1E(\tau^4)-(T+2\mu)\left(E(\tau^2)\right)^2\right]+v_2(v_2+2)E(\tau^4)\right)/(T+2\mu)^2,\end{aligned} \tag{9}$$

$$\begin{aligned}\rho(\sigma^2_{e_1},S^2_p) = &\left(\phi^2\left[v_1(v_1+2)E(\tau^4)\left((T+2\mu)^2-(T_2+\mu)(2T_1+T_2+3\mu)Q_{04}\right)\right.\right.\\ &\left.+ (T+2\mu)(T_1+\mu)\left(E(\tau^2)\right)^2\left((T_1+\mu-2v_1)(T+2\mu)+2(T_2+\mu)v_1Q_{02}\right)\right]\\ &+ 2(T_1+\mu)^2\phi\left[v_1v_2E(\tau^4)Q_{22}-v_2(T+2\mu)\left(E(\tau^2)\right)^2Q_{20}\right]\\ &\left.+ v_2(v_2+2)(T_1+\mu)^2E(\tau^4)Q_{40}\right)/\left((T_1+\mu)(T+2\mu)\right)^2,\end{aligned} \tag{10}$$

where $Q_{ij} = \Pr.\left[F_{(v_2+i,v_1+j)}\le\left(v_2(v_1+j)c\phi\right)/\left(v_1(v_2+i)\right)\right]$, $i,j = 0,1,\ldots$.

Note that Q_{ij} does not depend on τ.

Proof. See the appendix.

Corollary 1. If $e\sim Mt\left(0,\nu\sigma^2_2/(\nu-2)\Sigma\right)$ then, for $\nu>4$

$$\rho_{Mt}(\sigma^2_{e_1},S^2_N) = \phi^2\nu^2\sigma^4_2\left(2v_1(v_1+\nu-2)+(k+\mu)^2(\nu-4)\right)/\left((T_1+\mu)^2(\nu-4)(\nu-2)^2\right), \tag{11}$$

$$\begin{aligned}\rho_{Mt}(\sigma^2_{e_1},S^2_A) \doteq \nu^2\sigma^4_2&\left(\phi^2\left[(\nu-4)\left(v_2+2(k+\mu)\right)^2+2v_1(v_1+\nu-2)\right]\right.\\ &+ 2v_2\phi\left[2v_1-(\nu-4)\left(v_2+2(k+\mu)\right)\right]+v_2(v_2+2)(\nu-2)\Big)/\end{aligned}$$

$$\left((T+2\mu)^2(\nu-2)^2(\nu-4)\right), \tag{12}$$

$$\rho_{Mt}(\sigma^2_{e_1}, S^2_p) = \nu^2\sigma^4_2\Big(\phi^2\Big[(T+2\mu)^2\Big((\nu-4)(k+\mu)^2+2v_1(v_1+\nu-2)\Big)$$

$$- v_1(v_1+2)(\nu-2)(2T_1+T_2+3\mu)(T_2+\mu)Q_{04}+2v_1(T_1+\mu)(T+2\mu)$$

$$.(T_2+\mu)(\nu-4)Q_{02}\Big]+2(T_1+\mu)^2\phi v_2\Big[v_1(\nu-2)Q_{22}-(T+2\mu)(\nu-4)Q_{20}\Big]$$

$$+ v_2(v_2+2)(T_1+\mu)^2(\nu-2)Q_{40}\Big)\Big/\Big((T_1+\mu)^2(T+2\mu)^2(\nu-4)(\nu-2)^2\Big). \tag{13}$$

Proof. $e\sim Mt\left(0,\nu\sigma^2_2/(\nu-2)\Sigma\right)$ when $\tau\sim IG(\sigma^2_2,\nu)$. Then, $f(\tau)=\left[2/\Gamma(\nu/2)\right] .(\nu\sigma^2_2/2)^{\nu/2}\tau^{-(\nu+1)}e^{-\nu\sigma^2_2/2\tau^2}$, so, $E(\tau^2)=\nu\sigma^2_2/(\nu-2)$, $E(\tau^4)=\nu^2\sigma^4_2/\left((\nu-2)(\nu-4)\right)$. Using these expressions appropriately in Theorem 1 yields Corollary 1. #

Corollary 2. If $e\sim N(0,\sigma^2_2\Sigma)$ then $\sigma^2_{e_2}=\sigma^2_2$, $\sigma^2_{e_1}=\sigma^2_1$ (say), and

$$\rho_N(\sigma^2_1, S^2_N) = \phi^2\sigma^4_2\left(2v_1+(k+\mu)^2\right)/(T_1+\mu)^2, \tag{14}$$

$$\rho_N(\sigma^2_1, S^2_A) = \sigma^4_2\Big(\phi^2\Big[\Big(v_2+2(k+\mu)\Big)^2+2v_1\Big]-2\phi v_2\Big(v_2+2(k+\mu)\Big)$$

$$+ v_2(v_2+2)\Big)/(T+2\mu)^2, \tag{15}$$

$$\rho_N(\sigma^2_1, S^2_P) = \sigma^4_2\Big(\phi^2\Big[(T+2\mu)^2\Big((k+\mu)^2+2v_1\Big)-v_1(v_1+2)(T_2+\mu)(2T_1+T_2+3\mu)Q_{04}$$

$$+ 2v_1(T_1+\mu)(T_2+\mu)(T+2\mu)Q_{02}\Big]+2(T_1+\mu)^2\phi v_2\Big[v_1Q_{22}-(T+2\mu)Q_{20}\Big]$$

$$+ v_2(v_2+2)(T_1+\mu)^2Q_{40}\Big)\Big/\Big((T_1+\mu)^2(T+2\mu)^2\Big). \tag{16}$$

Proof. This corollary follows from Corollary 1 as $e\sim N(0,\sigma^2_2\Sigma)$ when $\nu=\infty$. #

Remarks.

(i) The risk expressions of Bancroft (1944) (allowing for the change in H_1) and of Toyoda and Wallace (1975) follow from Corollary 2 by setting $\mu=-k$.

(ii) If $\alpha=0$, $c=\infty$, then $Q_{ij}=1$ so we never reject H_0. Then, $\rho(\sigma^2_{e_1},S^2_P)=\rho(\sigma^2_{e_1},S^2_A)$. Conversely, if $\alpha=1$, $c=0$, then $Q_{ij}=0$ so that we reject H_0. Then, $\rho(\sigma^2_{e_1},S^2_P)=\rho(\sigma^2_{e_1},S^2_N)$.

(iii) $\lim\limits_{\phi\to 0}\left[\rho(\sigma^2_{e_1},S^2_P)\right]=\lim\limits_{\phi\to 0}\left[\rho(\sigma^2_{e_1},S^2_N)\right]=0$ while $\lim\limits_{\phi\to 0}\left[\rho(\sigma^2_{e_1},S^2_A)\right]=\left(v_2(v_2+2)E(\tau^4)\right)/(T+2\mu)^2>0$. Intuitively it is better to ignore the prior information when it is very false, and pre-testing leads us to follow the correct strategy of ignoring the second sample when estimating $\sigma^2_{e_1}$.

(iv) If $\phi=1$, that is, the error variances are equal, then the sign of

$$\rho(\sigma^2_{e_1},S^2_A|\phi=1)-\rho(\sigma^2_{e_1},S^2_N|\phi=1)=\Big\{E(\tau^4)\Big[(T_1+\mu)^2(v_1+v_2)(v_1+v_2+2)$$
$$-v_1(v_1+2)(T+2\mu)^2+\left(E(\tau^2)\right)^2(T+2\mu)(T_1+\mu)\Big[(T_2+\mu)\left(v_1-(k+\mu)\right)$$
$$-(T_1+\mu)\left(v_2-(k+\mu)\right)\Big]\Big]\Big\}/\left((T+2\mu)(T_1+\mu)\right)^2 \tag{17}$$

is negative if $\{.\}<0$, so that imposing valid prior information produces a risk gain. The sign of (17) is not obvious. If we are employing the L components then (17) is equal to $\left[-2E(\tau^4)v_2/\left(v_1(v_1+v_2)\right)\right]$ which is negative for all v_1, v_2, $f(\tau)$. However, the sign is still ambiguous if we are using the ML or the M components. We will return to this feature in Section 3.

(v) The risk functions of S^2_N and S^2_A have two intersections with respect to ϕ. Let these be ϕ_1 and ϕ_2. Their values are

$$\phi_i = \Big[v_2(T_1+\mu)^2\Big[v_1E(\tau^4)-(T+2\mu)\left(E(\tau^2)\right)^2\Big]\pm(T_1+\mu)\Big\{v_1v_2\Big(E(\tau^4)^2$$
$$\cdot\Big[v_1v_2(T_1+\mu)^2+(v_1+2)(v_2+2)(T_2+\mu)(2T_1+T_2+3\mu)\Big]+v_2^2(T_1+\mu)^2(T+2\mu)^2\left(E(\tau^2)\right)^4$$
$$-2v_1v_2(T_1+\mu)(T+2\mu)E(\tau^4)\left(E(\tau^2)\right)^2\Big[v_2(T+2\mu)+2(T_2+\mu)\Big]\Big\}^{\frac{1}{2}}\Big]/\Big[v_1(v_1+2)E(\tau^4)$$
$$\cdot(T_2+\mu)(2T_1+T_2+3\mu)-2v_1(T_1+\mu)(T+2\mu)(T_2+\mu)\left(E(\tau^2)\right)^2\Big], \tag{18}$$
$$= \omega\pm\kappa \quad , \quad i = 1,2.$$

to discern the signs of ϕ_1 and ϕ_2 from (18), though our numerical evaluations with Mt regression disturbances suggest that there are two possibilities. We comment on this in the next section.

(vi) Bancroft (1944) and Toyoda and Wallace (1975) show that there is a ϕ-range over which it is preferable to pre-test rather than to always-pool or to never-pool the two samples when $e \sim N(0, \sigma_2^2 \Sigma)$, using the usual L components. They find that there is a family of pre-test estimators, with $c \in (0,2)$, which strictly dominate first, the never-pool estimator for all ϕ and secondly, the always-pool estimator for a wide range of ϕ. It is only in the neighbourhood of $\phi=1$ that the risk of s_{AL}^2 is smaller than that of s_{PL}^2. Ohtani and Toyoda (1978) prove that of this family of dominating estimators the pre-test estimator with c=1 has the smallest risk. The following theorem extends this result to the case that we are investigating.

Theorem 2. The pre-test risk function has a minimum when $c^* = \left(v_1(T_2+\mu)\right) / \left(v_2(T_1+\mu)\right)$.

Proof. See the appendix.[9]

So, in particular, $c_L^*=1$, $c_{ML}^*=(v_1T_2)/(v_2T_1)$, and $c_M^* = \left(v_1(v_2+2)\right) / \left(v_2(v_1+2)\right)$.

3. Numerical evaluations of the risk functions

Given the complexities of the risk expressions, it is useful to evaluate them numerically, which we have done, assuming Mt errors, for the L, the ML and the M component estimators for various values of ν, α, v_1, v_2, and k, as functions of ϕ. We evaluate the risks relative to σ_2^2 and so, define the relative risk of an estimator $\bar{s}^2$ of $\sigma_{e_1}^2$ as $R_{\bar{s}^2} = \rho(\sigma_{e_1}^2, \bar{s}^2)/\sigma_2^4$.[10] Giles (1990a) details the range of the values of the arguments considered and the computer programs employed. Figures 1 and 2 depict typical results for the L components when $v_1=16$, $v_2=8$, k=3, $\nu=5$ and $\nu=\infty$. For this example $c^*=1$

details the range of the values of the arguments considered and the computer programs employed. Figures 1 and 2 depict typical results for the L components when $v_1=16$, $v_2=8$, $k=3$, $\nu=5$ and $\nu=\infty$. For this example $c^*=1$ corresponds to $\alpha=47.3\%$. The risk functions for the ML and the M components are qualitative similar, though there are quantitative differences which we mention below.

Remarks.

(i) If $e \sim Mt\left(0, \nu\sigma_2^2/(\nu-2)\Sigma\right)$ and we are using the ML or the M components then the risk difference (17) is negative for all possible values of ν.[11] So, when the error variances are equal it is always preferable to pool the samples, rather than to ignore the prior information.

(ii) The numerical evaluations suggest that there are two possible values of ϕ_{1j} and ϕ_{2j}, j=L,ML,M. First, $0<\phi_{1j}<1$, $\phi_{2j}<0$ and secondly, $0<\phi_{1j}<1$, $\phi_{2j}>1$. Thus, there exists one feasible intersection, $\phi_{1j}\in(0,1)$. So, the never-pool estimator dominates the always-pool estimator when $0<\phi<\phi_{1j}$. Alternatively, the always-pool estimator has smaller risk than the never-pool estimator when $\phi_{1j}<\phi\le1$. For this ϕ-range the gain in sampling variance from the extra degrees of freedom when pooling the samples outweighs the bias from pooling the (unequal) variances. These conclusions accord with those found by Toyoda and Wallace (1975).

Our numerical evaluations also suggest that $\phi_{1ML}<\phi_{1M}<\phi_{1L}$ if $v_2\le v_1$, while the inequalities are reversed if $v_1<v_2$. Further, ϕ_{1j} decreases as ν increases, j=L,ML,M. This implies if we assume normality when in fact $e \sim Mt\left(0, \nu\sigma_2^2/(\nu-2)\Sigma\right)$, $\nu<\infty$, that there is then a ϕ-range over which we would incorrectly choose to pool the samples.

(iii) For relatively small ν the pre-test estimator can strictly dominate *both* of its component estimators. In such cases, it is always preferable to pre-test, and given Theorem 2, to use $c=c^*$. For these values of ν the pre-test estimator has smaller variability than either of its component estimators.[12]

(iv) Of the L, ML, and the M estimators, the numerical results suggest, if one adopted a pre-test strategy and a crude minimax risk criterion, that for normal disturbance terms the preferred estimator is s^2_{PM} for $\alpha=0.01$ and s^2_{PL} for $\alpha\geq0.05$. However, if ν is relatively small then it is preferable to use the ML component estimators. So, given our previous discussion, for small ν we should pre-test using the ML components and a critical value of $(v_1T_2)/(v_2T_1)$.

4. Concluding remarks

In this paper we have examined the risk properties of estimators of the disturbance variance, after a preliminary test of homogeneity, when the joint distribution of the unobservable errors in each sample is SSD but it is assumed to be normal. We have considered the usual least squares estimators of the error variance and we have also investigated the risks of the never-pool, the always-pool and the pre-test estimators whose components are the usual never-pool maximum likelihood and the minimum mean squared error estimators assuming a correctly specified error distribution. Of course, under the investigated specification error these estimators do not possess their desired properties.

Nevertheless, our results suggest that these estimators are preferred to the usual least squares estimator when ν is small. Then the ML pretest estimator which uses $c=c^*$ strictly dominates: it is never preferable to always-pool the samples without testing the validity of the null hypothesis, nor is it optimal to ignore the prior information.

We should recall that the results discussed here apply to a one-sided alternative hypothesis. It remains for future research to consider the two-sided case. We also need to investigate the sensitivity of the results to the particular form of non-normality considered. Whether they will extend to the situation of non-normal but identically, independently distributed disturbances is not clear.

July, 1990

Appendix

Proof of Theorem 1

$S_N^2=\phi e^{*\prime}M_1^*e^*/(T_1+\mu)$ where M_1^* is a (T×T) idempotent matrix partitioned as $M_1^*=\begin{bmatrix} M_1 & 0 \\ 0 & 0 \end{bmatrix}$, $M_1=X_1(X_1'X_1)^{-1}X_1'$, $r(M_1)=v_1$, and $e^{*\prime}=[e_1'/\sqrt{\phi} \quad e_2']$. Now, $\rho(\sigma_{e_1}^2,S_N^2) = E(S_N^4)-2\phi E(\tau^2)E(S_N^2)+\phi^2\left(E(\tau^2)\right)^2$ and $E(S_N^2)=\int_0^\infty E_N(S_N^2)f(\tau)d\tau$ where $E_N(.)=E(.)$ when $e\sim N(0,\tau^2\Sigma)$. So, then $e^*\sim N(0,\tau^2I_T)$ and $e^{*\prime}M_1^*e^*/\tau^2\sim\chi_{v_1}^2$, which gives $E_N(e^{*\prime}M_1^*e^*/\tau^2)=v_1$, $E_N(e^{*\prime}M_1^*e^*/\tau^2)^2=v_1(v_1+2)$. Using these results $\rho(\sigma_{e_1}^2,S_N^2)$ follows directly.

Similarly, $S_A^2=(\phi e^{*\prime}M_1^*e^*+e^{*\prime}M_2^*e^*)/(T+2\mu)$, where M_2^* is a (T×T) idempotent matrix partitioned as $M_2^*=\begin{bmatrix} 0 & 0 \\ 0 & M_2 \end{bmatrix}$, $M_2=X_2(X_2'X_2)^{-1}X_2'$, $r(M_2)=v_2$. Now, when $e\sim N(0,\tau^2\Sigma)$ so that $e^*\sim N(0,\tau^2I_T)$ it is straightforward to show that the quadratic forms $(e^{*\prime}M_2^*e^*/\tau^2)$ and $(e^{*\prime}M_1^*e^*/\tau^2)$ are independent. Further, $(e^{*\prime}M_2^*e^*/\tau^2)\sim\chi_{v_2}^2$. So, using the moments of a χ^2 random variate and the fact that we can write $E(.)=\int_0^\infty E_N(.)f(\tau)d\tau$ the $\rho(\sigma_{e_1}^2,S_A^2)$ follows in the same manner as it did for S_N^2.

Finally, to establish $\rho(\sigma_{e_1}^2,S_P^2)$ we write

$$S_P^2 = \left(\phi(T+2\mu)(e^{*\prime}M_1^*e^*)+\left[(T_1+\mu)(e^{*\prime}M_2^*e^*)-\phi(T_2+\mu)(e^{*\prime}M_1^*e^*)\right]\right.$$
$$\left.\cdot I_{[0,c\phi]}\left((v_1e^{*\prime}M_2^*e^*)/(v_2e^{*\prime}M_1^*e^*)\right)\right)/\left((T_1+\mu)(T+2\mu)\right),$$

and note that

$$\rho(\sigma_{e_1}^2,S_P^2) = E(S_p^4)-2\phi E(\tau^2)E(S_P^2)+\phi^2\left(E(\tau^2)\right)^2. \tag{A.1}$$

Now, $E(.)=\int_0^\infty E_N(.)f(\tau)d\tau$ and so we require $E_N(S_P^2)$ and $E_N(S_P^4)$. Using Lemma 1 of Clarke *et al.* (1987) we have

$$E_N(S_P^2) = \left(\phi v_1\tau^2(T+2\mu)+v_2\tau^2(T_1+\mu)Q_{20}-v_1\phi(T_2+\mu)\tau^2Q_{02}\right)/\left((T_1+\mu)(T+2\mu)\right),$$

and

$$E_N(S_P^4) = \tau^4\Big(\phi^2v_1(v_1+2)(T+2\mu)^2-\phi^2v_1(v_1+2)(T_2+\mu)(2T_1+T_2+3\mu)Q_{04}$$
$$+ 2\phi v_1v_2(T_1+\mu)^2Q_{22}+v_2(v_2+2)(T_1+\mu)^2Q_{40}\Big)/\left((T_1+\mu)(T+2\mu)\right)^2.$$

To complete the proof we integrate these expressions with respect to τ to give $E(S_P^2)$ and $E(S_P^4)$, then substitute these into (A.1). #

Proof of Theorem 2

$$\rho(\sigma^2_{e_1},S_P^2) = E\left[\left(S_A^2-\phi E(\tau^2)\right)^2 I_{[0,c]}(J)+\left(S_N^2-\phi E(\tau^2)\right)^2 I_{(c,\infty)}(J)\right]$$

$$= \int_0^\infty \tau^4 E_N\Big\{\Big[\Big((\phi e^{*\prime} M_1^* e^*/\tau^2)+(e^{*\prime} M_2^* e^*/\tau^2)\Big)/(T+2\mu)$$
$$- \phi E(\tau^2)/\tau^2\Big]^2 I\left((e^{*\prime} M_2^* e^*/\tau^2)\le g\right)+\Big[(\phi e^{*\prime} M_1^* e^*/\tau^2)/(T_1+\mu)$$
$$- \phi E(\tau^2)/\tau^2\Big]^2\Big[1-I\left((e^{*\prime} M_2^* e^*/\tau^2)\le g\right)\Big]\Big\}f(\tau)d\tau\ , \qquad \text{(A.2)}$$

where $g=c\phi v_2(e^{*\prime} M_1^* e^*/\tau^2)/v_1$. Given (A.2) the remainder of this proof follows the approach outlined in Giles (1990b) for a similar theorem. #

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 C-P Han, 1983, A note on pooling variances, Journal of the American Statistical Association 78, 981-983.

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.

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

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.

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

Giles, J.A., 1990b, Pre-testing for linear restrictions in a regression model with spherically symmetric disturbances, mimeo., University of Canterbury.

Judge, G.G., S. Miyazaki and T.A. Yancey, 1985, Minimax estimators for the location vectors of spherically symmetric densities, Econometric Theory 1, 509-417.

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.

Mandelbrot, B.B., 1963, The variation of certain speculative prices, Journal of Business 36, 394-419.

Mandelbrot, B.B., 1967, The variation of some other speculative prices, Journal of Business 40, 393-413.

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

Ohtani, K., 1987, Some small sample properties of a pre-test estimator of the disturbance variance in a misspecified linear regression, Journal of the Japan Statistical Society 17, 81-89.

Ohtani, K. and T. Toyoda, 1978, Minimax regret critical values for a preliminary test in pooling variance, 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.

Praetz, P.D. and E.J.G. Wilson, 1978, The distribution of stock market returns: 1858-1973, Australian Journal of Management 3, 79-90.

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 of variance 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, 1985, Estimation and testing in a regression model with spherically symmetric errors, Economics Letters 17, 127-132.

Yancey, T.A., G.G. Judge and D.M. Mandy, 1983, The sampling performance of pre-test estimators of the scale parameter under squared error loss, Economics Letters 12, 181-186.

Zinde-Walsh, V. and A. Ullah, 1987, On robustness of tests of linear restrictions in regression models with elliptical error distributions, in I.B. MacNeill and G.J. Umphrey (eds), Time series and econometric modelling (D. Reidel, Boston).

Footnotes

[1]A discussion of this family of distributions is beyond our scope. See, for example, Kelker (1970) and Muirhead (1982).

[2]We could assume that each sample is generated by a different variance mixing distribution. If the mixing distributions are independent then we can easily extend our analysis. However, it is unclear how we would proceed if they are dependent.

[3]In this paper we require $\nu > 4$ and so, in particular, our results exclude the Cauchy case.

[4]Our alternative hypothesis is consistent with the majority of the existing literature.

[5]Yancey *et al.* (1983) and Ohtani (1987) consider a related, though not identical, pre-test problem to that investigated here. They also assume normality.

[6]This result is also implied by King (1979).

[7]Giles (1990b) uses the least squares unrestricted and restricted estimators of the error variance. Giles (1990a) derives the critical values which minimise the pre-test risk function when one uses the usual maximum likelihood or the minimum mean squared error (assuming normality) component estimators. Then, the optimal critical values are zero and $v/(v+2)$ respectively, where $v=T-k$.

[8]One may then reasonably ask why we do not consider the L, ML, and M pooled estimators assuming that the error variances are equal and that the disturbances are normally distributed. These estimators should be more efficient than the pooled estimators that we consider as they incorporate the information that β is common to both samples. However, even under a normality assumption , the non-null distribution of $e'Me$ is not clear, so we do not proceed along this path. One could presumably conceive of many other always-pool estimators to that which we investigate.

[9]The form of this proof is not the same as that used by Ohtani and Toyoda (1978). Further, this proof can be easily extended to allow for omitted regressors. c^* remains unchanged. See Giles (1990a).

[10]We lose no generality in considering relative risk, and the results could equally be interpreted as the risk functions when $\sigma_2^2=1$.

[11]We noted in point (iv) of Section 2 that (17) is always negative for the L components but we could not algebraically sign (17) for the ML or the M components.

[12]This result is also found by Giles (1990b) when estimating the error variance after a pre-test for exact linear restrictions.

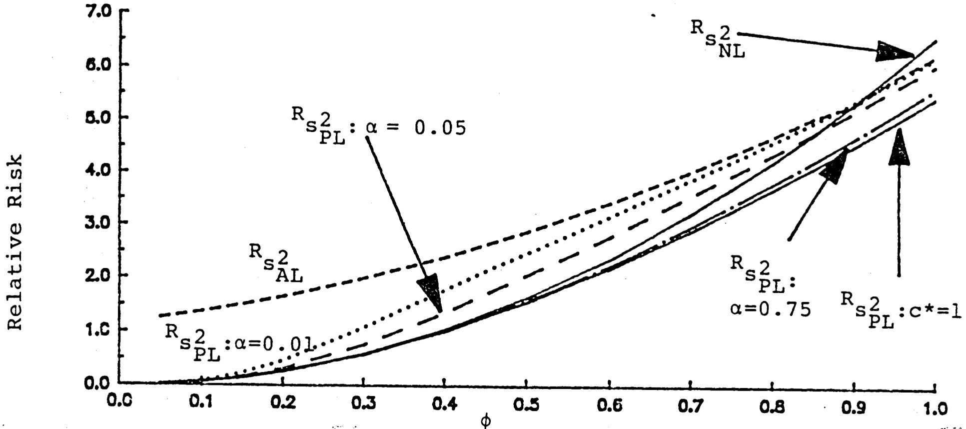

FIGURE 1. Relative risk functions for s^2_{NL}, s^2_{AL}, and s^2_{PL} when $e \sim Mt(0,\nu\sigma_2^2/(\nu-2)\Sigma)$, $v_1 = 16$, $v_2 = 8$, $k = 3$, $\nu = 5$.

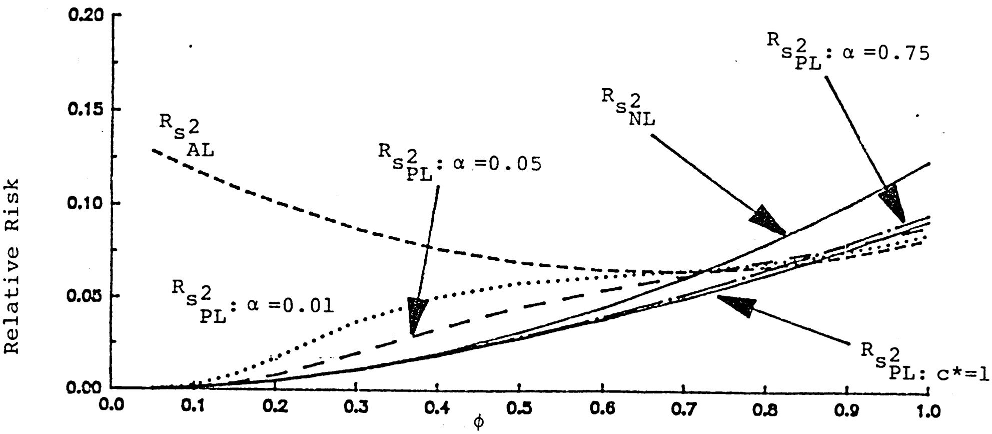

FIGURE 2. Relative risk functions for s^2_{NL}, s^2_{AL}, and s^2_{PL} when $e \sim N(0,\sigma_2^2\Sigma)$, $v_1 = 16$, $v_2 = 8$, $k = 3$.

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 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.

* 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.