Department of Economics
UNIVERSITY OF CANTERBURY
CHRISTCHURCH, NEW ZEALAND

PRE-TESTING IN A MIS-SPECIFIED REGRESSION MODEL

Judith A. Giles

Discussion Paper

No. 9013

Department of Economics, University of Canterbury
Christchurch, New Zealand

Discussion Paper No. 9013

November 1990

PRE-TESTING IN A MIS-SPECIFIED REGRESSION MODEL

Judith A. Giles

PRE-TESTING IN A MIS-SPECIFIED REGRESSION MODEL

Judith A. Giles

Department of Economics
University of Canterbury
Private Bag, Christchurch
New Zealand

Key Words and Phrases: pre-test estimation; mis-specified regression models; spherical symmetry; multivariate Student-t.

ABSTRACT

We consider the pre-test estimation of the parameters of a linear regression model after a preliminary-test for exact linear restrictions when the model is mis-specified through the omission of relevant regressors and the usual assumption of normal regression disturbances is widened to a subclass of the family of spherically symmetric errors. We derive and analyse the exact risk (under quadratic loss) of a pre-test estimator of the prediction vector and of the scale parameter.

1. INTRODUCTION

Econometricians inevitably work with false models. So, we should be investigating the properties of estimators within a mis-specified regression model, whereas traditionally, pre-test estimators have been examined within the context of the standard linear regresssion model assuming normal iid disturbances and a correctly specified design matrix. There have been exceptions. Ohtani (1983), Mittelhammer (1984), Giles (1986), and Giles and Clarke (1989) consider the effects of omitting relevant regressors or including irrelevant ones, or of proxying unobservable variables, while Giles (1991) derives the exact risk of the estimators that we consider here assuming spherically symmetric disturbances, though a correctly specified design matrix. (See also, Judge *et al.* (1985) and Miyazaki *et al.* (1986) who investigate, via Monte Carlo experiments, the effects of non-normal regression disturbances on the risks of some related estimators.)

However, departures from the standard regression assumptions are likely to occur simultaneously. Accordingly, we derive the exact risk of pre-test estimators of the prediction vector and of the error variance when the disturbances are spherically symmetric <u>and</u> we have omitted relevant regressors from the design matrix.

2. THE MODEL FRAMEWORK AND SOME PRELIMINARY RESULTS

Suppose that the process generating a (T×1) vector of observations on a dependent variable y is

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

where X and Z are (T×k) and (T×p) full rank matrices of non-stochastic variables, and β and γ are (k×1) and (p×1) vectors of unknown parameters respectively. We assume that the (T×1) vector of disturbances e is distributed according to the laws of the class of spherical compound normal distributions (see Kelker (1970) and Muirhead (1982)) with $E(e)=0$, and $E(ee')=\sigma_e^2 I_T$. This class of distributions is a subclass of the family of spherically symmetric distributions which can be expressed as a variance mixture of normal distributions. That is, we can write

$$f(e) = \int_0^\infty f_N(e)f(\tau)d\tau, \tag{2}$$

where f_N denotes a probability density function (pdf) when $e \sim N(0,\tau^2 I_T)$, $f(\tau)$ is the pdf of τ and is supported on $[0,\infty)$. So, $\sigma_e^2=E(\tau^2)$, and the errors are uncorrelated but are dependent: independence is a feature if and only if the underlying distribution is normal. Further, the marginal distribution of the errors may have fatter or thinner tails than that which would result under a normality assumption. We write $e \sim SSD_N(0, \sigma_e^2 I_T)$.

One particular example of a distribution which satisfies (2) is the multivariate Student-t (Mt) distribution. It results when τ is an inverted gamma random variate. If this distribution has a degrees of freedom parameter ν, and scale parameter σ^2, then $E(\tau^2)=\nu\sigma^2/(\nu-2)$, and normality results when $\nu=\infty$. For $\nu<\infty$ the marginal distributions have fatter tails than when $\nu=\infty$.

Now suppose that the researcher specifies the model

$$y = X\beta + u \quad ; \quad u \sim N(0, \sigma_u^2 I_T) \tag{3}$$

as the data generating process. He proceeds assuming (3) to be properly specified when in fact $u \sim SSD_N(Z\gamma, \sigma_e^2 I_T)$. Note that $\sigma_u^2 = \sigma_e^2$. In addition, we assume that the investigator has (uncertain) extraneous prior information about the parameters β which he can express as $m(<k)$ exact linearly independent restrictions $R\beta=r$, where R is an $(m \times k)$ known full rank matrix, and r is an $(m \times 1)$ vector of known non-stochastic elements.

Under the assumptions of (3) the unrestricted and the restricted least squares (and maximum likelihood) estimators of β are respectively, $b=S^{-1}X'y$ and $b^*=b+S^{-1}R'[RS^{-1}R']^{-1}(r-Rb)$, where $S=(X'X)$. Note that b and b^* are the MLE's under the spherical assumption assuming that $Z\gamma=0$, but that for model (1) this holds only if X and Z are orthogonal. Similarly, under (3) the unrestricted least squares estimator of σ_e^2 is $\tilde{\sigma}_e^2=(y-Xb)'(y-Xb)/v$ and the restricted least squares estimator of σ_e^2 is $\sigma_e^{*2}=(y-Xb^*)'(y-Xb^*)/(v+m)$, where $v=(T-k)$.[1]

---

1. The estimators $\tilde{\sigma}_e^2$ and σ_e^{*2} are the least squares estimators of σ_e^2 for the wider assumption of $e \sim SSD_N$. This is not so for the usual maximum likelihood (ML) or the minimum mean squared error (M) estimators though, for the problem examined here, the researcher would proceed using the usual ML or M estimators. Giles (1990) extends the results presented here to these other cases.

The researcher, uncertain of the validity of the restrictions, undertakes a pre-test of $H_0:\delta=0$ versus $H_1:\delta\neq0$, where $\delta=R\beta-r$ represents an (m×1) hypothesis error vector, using the traditional Wald (and Lagrange Multiplier) test statistic $\ell=(Rb-r)'[RS^{-1}R']^{-1}(Rb-r)v/m(y-Xb)'(y-Xb)$. When $Z\gamma=0$ and $e\sim SSD_N(0,\sigma_e^2I_T)$ then $\ell\sim F_{(m,v)}$ under H_0 (see King(1979)). However, this property no longer holds if the design matrix is mis-specified; then $f(\ell)$ depends not only on m, v and the degree of mis-specification but it depends also on the variance mixing distribution. This is shown by Theorem 1.

Theorem 1. Under the above assumptions,

$$f(\ell) = \sum_{r=0}^{\infty}\sum_{s=0}^{\infty} \frac{\left(\theta_n^r/r!\right)\left(\theta_d^s/s!\right)m^{m/2+r}\, v^{v/2+s}\, \ell^{m/2+r-1}}{B\left(\frac{m}{2}+r;\frac{v}{2}+s\right)\left(v+m\ell\right)^{(m+v)/2+r+s}}$$

$$\times \int_0^{\infty} e^{-(\theta_n+\theta_d)/\tau^2}\left(\tau^2\right)^{-(r+s)} f(\tau)d\tau \;, \qquad (4)$$

where $\theta_n=(\Lambda+\delta)'[RS^{-1}R']^{-1}(\Lambda+\delta)/2$, $\theta_d=\gamma'Z'MZ\gamma/2$, $M=I-X(X'X)^{-1}X'$, $\Lambda=RS^{-1}X'Z\gamma$, and $B(.;.)$ is the beta function.

Proof. $f(\ell)=\int_0^{\infty} f_N(\ell)f(\tau)d(\tau)$. Now, $f_N(\ell)$ is a doubly non-central F density with m and v degrees of freedom and non-centrality parameters $\lambda_{n\tau}=\theta_n/\tau^2$ and $\lambda_{d\tau}=\theta_d/\tau^2$ (see Ohtani (1983) or Mittelhammer (1984)), and so (4) follows directly.

Clearly, if H_0 is true (4) is still not the density function of a central F random variate (even if the errors are normally distributed). So, the classical test is invalid if we omit relevant regressors.

3. THE RISK FUNCTIONS OF ESTIMATORS OF E(y)

We consider risk under quadratic loss which for an estimator ϕ^* of ϕ is given by $\rho\left(\phi,\phi^*\right)=E\left[\left(\phi-\phi^*\right)'\left(\phi-\phi^*\right)\right]=tr\left[cov(\phi^*)+bias(\phi^*)bias(\phi^*)'\right]$, which is the trace of the matrix mean squared error.[2] Our interest in this section lies with the estimation of the prediction vector E(y) after the pre-test of H_0 within the mis-specified framework of model (3). We consider this quantity rather than β so that our results are independent of the design matrix. In terms of the β space this is equivalent to assuming orthonormal regressors.

The pre-test estimator of E(y) is $X\hat{b} = I_{[0,c]}(\ell)Xb^* + I_{(c,\infty)}(\ell)Xb$, where c is the critical value of the test associated with a (nominal) significance level of $\alpha\%$, and $I_{[.,.]}$ is an indicator function which is unity if ℓ lies within the subscripted range and zero otherwise. The risks of Xb, Xb* and $X\hat{b}$ are given in Theorem 2.

Theorem 2. Under the above assumptions,

$$\rho\left(E(y),Xb\right) = kE(\tau^2) + 2\theta_d \tag{5}$$

$$\rho\left(E(y),Xb^*\right) = (k-m)E(\tau^2) + 2(\theta_d+\theta_n) \tag{6}$$

$$\rho\left(E(y),X\hat{b}\right) = kE(\tau^2) + 2\theta_d + \int_0^\infty\left(2\theta_n\left[2P_{20}-P_{40}\right]-m\tau^2P_{20}\right)f(\tau)d\tau \tag{7}$$

where

$$P_{ij} = Pr.\left[F''_{(m+i,v+j;\lambda_{n\tau},\lambda_{d\tau})}<\left(cm(v+j)\right)/\left(v(m+i)\right)\right],\quad i,j=0,1,2,\ldots .$$

Proof. See Giles (1990). The proof is similar in form to that given for Theorem 2 of Giles (1991).

---

2. So, we require the existence of the first two moments. This implies that our results are inapplicable, in particular, to distributions with infinite variance, such as the Cauchy distribution.

The risk functions depend on the hypothesis error through δ (and hence θ_n), on the specification error through Λ, and so on θ_n and θ_d, and on the variance mixing distribution. Equations (5)-(7) collapse to the expressions derived by Giles (1991) when $Z\gamma=0$, and to those derived by Mittelhammer (1984) when $e \sim N(0,\sigma^2 I_T)$.

Comparing the risk functions we find that regardless of $f(\tau)$ the results discussed by Mittelhammer (1984) for normal errors qualitatively carry over to the wider case of SSD_N errors. First, when the model is mis-specified the use of prior information (even if it is correct) does not guarantee a reduction in the risk of estimating E(y). This arises as $\theta_n \neq 0$ when H_0 is true unless θ_d is simultaneously zero or X and Z are orthogonal. Let $\theta_{n0}=\Lambda'[RS^{-1}R']^{-1}\Lambda/2$, be the value of θ_n under the null. Then, if $\theta_{n0}>mE(\tau^2)/2$, Xb^* has greater risk than Xb even though $\delta=0$. Similarily, $X\hat{b}$ may have higher risk than Xb even if H_0 is true.

Secondly, for a given value of θ_n the risk functions of Xb, Xb^*, and $X\hat{b}$ are unbounded as $\theta_d \to \infty$. However, $\rho\left(E(y),Xb\right)-\rho\left(E(y),Xb^*\right)$ is bounded and equal to $mE(\tau^2)-2\theta_n$, while $\rho\left(E(y),Xb^*\right)-\rho\left(E(y),X\hat{b}\right)=0$ when $\theta_d=\infty$, for any fixed value of θ_n. Thirdly, $\rho\left(E(y),Xb\right)=\rho\left(E(y),X\hat{b}\right)$ within the bounds $mE(\tau^2)/4 \leq \theta_n \leq mE(\tau^2)/2$, while $\rho\left(E(y),Xb\right)=\rho\left(E(y),Xb^*\right)$ when $\theta_n=mE(\tau^2)/2$. These values of θ_n are independent of θ_d.

To illustrate the risk functions we have numerically evaluated them for the special case of Mt disturbances. Then,

$$\rho_{Mt}\left(E(y),Xb\right) = \sigma^2\left(k\nu + 2\lambda_d(\nu-2)\right)/(\nu-2) \tag{8}$$

$$\rho_{Mt}\left(E(y),Xb^*\right) = \sigma^2\left((k-m)\nu + 2(\lambda_d+\lambda_n)(\nu-2)\right)/(\nu-2) \tag{9}$$

$$\rho_{Mt}\left(E(y),X\hat{b}\right) = \sigma^2\left(k\nu-m\nu P_{201}+2\lambda_d(\nu-2)+2\lambda_n(\nu-2)(2P_{202} - P_{402})\right)/(\nu-2) \qquad (10)$$

where

$$P_{ijn} = \sum_{r=0}^{\infty}\sum_{s=0}^{\infty} \frac{\left(2\lambda_n/\nu\right)^r\left(2\lambda_d/\nu\right)^s\ \Gamma\left(\frac{\nu}{2}+r+s+n-2\right)}{r!s!\ \left(1+2(\lambda_n+\lambda_d)/\nu\right)^{\frac{\nu}{2}+r+s+n-2}\ \Gamma\left(\frac{\nu}{2}+n-2\right)}$$

$$\times\ I_x\left(\ \frac{1}{2}(m+i)+r;\ \frac{1}{2}(v+j)+s\right) \quad , \qquad i,j,n,=0,1,2,\ldots \quad ,$$

$\lambda_n=\theta_n/\sigma^2$, $\lambda_d=\theta_d/\sigma^2$, and $I_x(.;.)$ is the incomplete beta function with $x=cm/(v+cm)$.

We have considered v=10, 16, 20, 30; k=4, 5; m=1, 3; α=0.01, 0.05, 0.30, 0.50, 0.75; ν=5, 10, 100, 1000, ∞; $\lambda_n\in[0,20]$; and $\lambda_d\in[0,20]$. The FORTRAN computer programs were executed on a VAX 6230 computer, and subroutines from Press *et al.* (1986) and Davies (1980) were used to undertake the evaluations. Figures 1 to 4 present some typical results. There, we consider risk relative to σ^2 and parameterise with respect to λ_n and λ_d rather than with respect to θ_n and θ_d to eliminate the scale parameter σ^2. Equivalently, the figures represent the risks of the estimators when σ^2=1. So, we define the relative risk of an estimator $X\bar{b}$ of $E(y)$ as $R(X\bar{b})=\rho\left(E(y),X\bar{b}\right)/\sigma^2$. Full details of the results are given in Giles (1990), or are available on request.

These figures illustrate the features discussed above. They also show first, that it is never preferable to pre-test. Pre-testing can be the worst strategy. Secondly, they show that an increase in the degree of mis-specification of the design matrix causes an upward shift of the estimator risk functions and thirdly, they show that the effect of changes in the value of ν when variables have been omitted

FIGURE 1 : Relative risk functions for Xb, Xb*, and $X\hat{b}$ when T=30, k=5, m=3, ν=5, and λ_d=0.

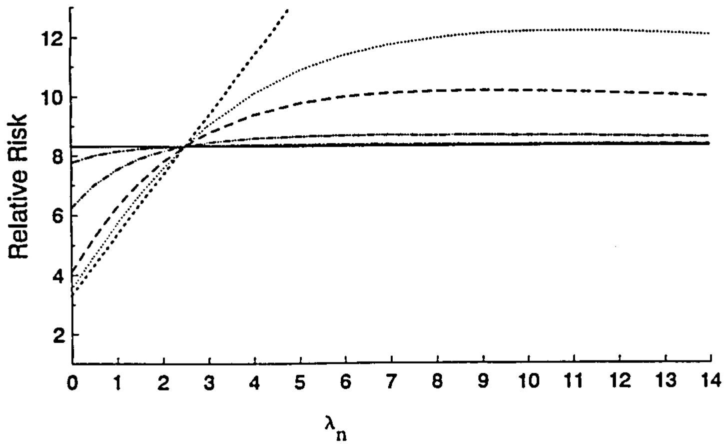


FIGURE 2 : Relative risk functions for Xb, Xb*, and $X\hat{b}$ when T=30, k=5, m=3, ν=5, and λ_d=10.

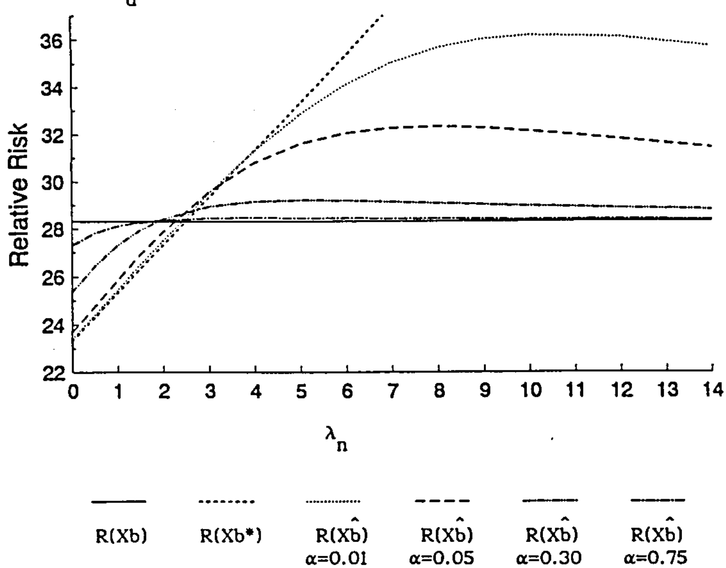


FIGURE 3 : Relative risk functions for Xb, Xb^*, and $X\hat{b}$ when $T=30$, $k=5$, $m=3$, $\nu=\infty$, and $\lambda_d=0$.

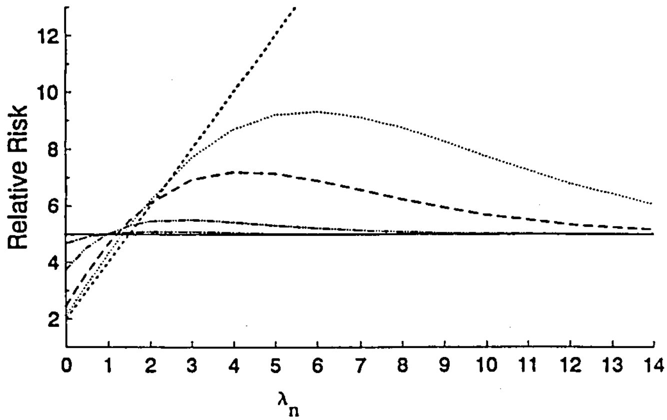
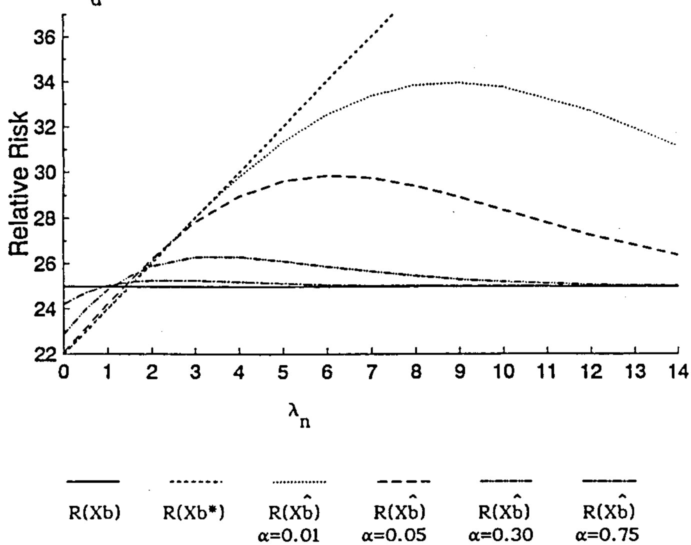


FIGURE 4 : Relative risk functions for Xb, Xb^*, and $X\hat{b}$ when $T=30$, $k=5$, $m=3$, $\nu=\infty$, and $\lambda_d=10$.



from the design matrix are the same as those discussed by Giles (1991) for the properly specified case. In particular, if the marginal distribution of the errors has fatter tails than under normality then the risk functions shift upwards, there is an increase in the range over which we prefer Xb* to Xb, and there is a decrease in the rate at which the risk of the pre-test estimator approaches that of the unrestricted estimator.

4. THE RISK FUNCTIONS OF ESTIMATORS OF σ_e^2

In this section we consider the risk, under quadratic loss, of a pre-test estimator of σ_e^2, $\hat{\sigma}_e^2$, whose component estimators are $\tilde{\sigma}_e^2$ and σ_e^{*2} when model (3) is fitted to the data. We define $\hat{\sigma}_e^2$ as $\hat{\sigma}_e^2 = I_{[0,c]}(\ell)\sigma_e^{*2} + I_{(c,\infty)}(\ell)\tilde{\sigma}_e^2$. The risks of $\tilde{\sigma}_e^2$, σ_e^{*2}, and $\hat{\sigma}_e^2$ are given in Theorem 3.

Theorem 3. Under the above assumptions,

$$\rho\left(\sigma_e^2, \tilde{\sigma}_e^2\right) = \left(v(v+2)E(\tau^4)-v^2\left[E(\tau^2)\right]^2+4\theta_d\left[\theta_d+2E(\tau^2)\right]\right)/v^2 \quad (11)$$

$$\rho\left(\sigma_e^2, \sigma_e^{*2}\right) = \left((v+m)(v+m+2)E(\tau^4)-(v+m)^2\left[E(\tau^2)\right]^2+4(\theta_n+\theta_d)\right.$$
$$\left.\times\left[(\theta_n+\theta_d)+2E(\tau^2)\right]\right)/(v+m)^2 \quad (12)$$

$$\rho\left(\sigma_e^2, \hat{\sigma}_e^2\right) = \left\{\int_0^\infty \left[(v+m)^2\left(v(v+2)\tau^4+4(v+2)\tau^2\theta_d+4\theta_d^2\right)\right.\right.$$
$$+v^2(v+m)^2\left(E(\tau^2)\right)^2-2v(v+m)^2E(\tau^2)(v\tau^2+2\theta_d)-m(2v+m)$$
$$\times\left(v(v+2)\tau^4P_{04}+4(v+2)\theta_d\tau^2P_{06}+4\theta_d^2P_{08}\right)+v^2\left(m(m+2)\tau^4P_{40}\right.$$
$$\left.+4(m+2)\theta_n\tau^2P_{60}+4\theta_n^2P_{80}\right)+2v^2\left(mv\tau^4P_{22}+2m\theta_d\tau^2P_{24}+2v\theta_n\tau^2P_{42}\right.$$

$$+4\theta_n\theta_d P_{44}\Big)+2mv(v+m)E(\tau^2)\Big(v\tau^2P_{02}+2\theta_d P_{04}\Big)-2v^2(v+m)E(\tau^2)$$

$$\times\Big(m\tau^2P_{20}+2\theta_n P_{40}\Big)\Big]f(\tau)d\tau\Big\}/\Big(v^2(v+m)^2\Big)\ . \qquad (13)$$

Proof. See Giles (1990). The proof is similar in form to that given for Theorem 3 of Giles (1991).

Equations (11)-(13) collapse to the risk functions derived by Giles (1991) when $Z\gamma=0$. For any θ_d, $\rho(\sigma_e^2,\tilde{\sigma}_e^2)$ is independent of θ_n, and so it is bounded as $\theta_n\to\infty$, but it is unbounded as $\theta_d\to\infty$. Similarly, $\rho(\sigma_e^2,\hat{\sigma}_e^2)$ is bounded (by $\rho(\sigma_e^2,\tilde{\sigma}_e^2)$) as $\theta_n\to\infty$ (given θ_d), but it is unbounded as $\theta_d\to\infty$ (given θ_n). In contrast, $\rho(\sigma_e^2,\sigma_e^{*2})$ is unbounded as either $\theta_n\to\infty$ or $\theta_d\to\infty$. These results concur with those given by Giles and Clarke (1989) for the omitted variables case with normal errors and when using the maximum likelihood component estimators of σ_e^2.

The risk differences $\left(\rho(\sigma_e^2,\tilde{\sigma}_e^2)-\rho(\sigma_e^2,\sigma_e^{*2})\right)$, $\left(\rho(\sigma_e^2,\tilde{\sigma}_e^2)-\rho(\sigma_e^2,\hat{\sigma}_e^2)\right)$, $\left(\rho(\sigma_e^2,\sigma_e^{*2})-\rho(\sigma_e^2,\hat{\sigma}_e^2)\right)$ are unbounded as $\theta_d\to\infty$, given θ_n. For a given value of θ_d, as $\theta_n\to\infty$ the differences are unbounded except for $\left(\rho(\sigma_e^2,\tilde{\sigma}_e^2)-\rho(\sigma_e^2,\hat{\sigma}_e^2)\right)$ which is bounded and is equal to zero. The results given here as $\theta_d\to\infty$ contrast with those we observed in the previous section for estimating E(y).

$\rho(\sigma_e^2,\hat{\sigma}_e^2)$ has a minimum when c=0, 1 or ∞. Giles (1991) shows this to be the case when the design matrix is properly specified, and her proof extends easily to the mis-specified model. So, $\hat{\sigma}_e^2$ can dominate both $\tilde{\sigma}_e^2$ and σ_e^{*2} over some or all of the θ_n range.

As in the previous section, we have numerically evaluated the risk functions using the same values of the arguments as discussed there, and the case when the critical value is unity, when the regression disturbances are Mt. Then,

$$\rho_{Mt}\left(\sigma_e^2,\ \tilde{\sigma}_e^2\right) = \sigma^4\left(2v\nu^2(v+\nu-2)+4\lambda_d(\nu-2)(\nu-4)\left[\lambda_d(\nu-2)+2\nu\right]\right)/\left(v^2(\nu-2)^2(\nu-4)\right) \tag{14}$$

$$\rho_{Mt}\left(\sigma_e^2,\ \sigma_e^{*2}\right) = 2\sigma^4\left(\nu^2(v+m)(v+m+\nu-2)+2(\lambda_n+\lambda_d)(\nu-2)(\nu-4)\left[(\lambda_n+\lambda_d)(\nu-2)+2\nu\right]\right)/\left((\nu-2)^2(\nu-4)(v+m)^2\right) \tag{15}$$

$$\begin{aligned}\rho_{Mt}\left(\sigma_e^2,\ \hat{\sigma}_e^2\right) = \sigma^4\Big\{&2v(v+m)^2\nu^2(v+\nu-2)+2\lambda_d(\nu-2)(\nu-4)(v+m)^2\left[(\nu-2)\lambda_d+2\nu\right]\\ &-2v(v+m)\nu(\nu-4)\left[mv\nu\left(P_{201}-P_{021}\right)-2m\lambda_d(\nu-2)P_{042}+2v\lambda_n(\nu-2)P_{402}\right]\\ &-m(m+2v)(\nu-2)\left(v(v+2)\nu^2P_{040}+4(v+2)\lambda_d\nu(\nu-4)P_{061}+4\lambda_d^2(\nu-2)(\nu-4)P_{082}\right)\\ &+v^2(\nu-2)\left(m(m+2)\nu^2P_{400}+4(m+2)\lambda_n\nu(\nu-4)P_{601}+4\lambda_n^2(\nu-2)(\nu-4)P_{802}\right)\\ &+2v^2(\nu-2)\Big(mv\nu^2P_{220}+2m\lambda_d\nu(\nu-4)P_{241}+2v\lambda_n\nu(\nu-4)P_{421}+4\lambda_n\lambda_d(\nu-2)\\ &\times(\nu-4)P_{442}\Big)\Big\}/\left((\nu-2)^2(\nu-4)v^2(v+m)^2\right).\end{aligned} \tag{16}$$

Figures 5 to 8 illustrate a typical case. We have again considered risk relative to the scale parameter, and parameterise with respect to λ_n and λ_d. Here we define the relative risk of an estimator $\bar{\sigma}_e^2$ of σ_e^2 as $R(\bar{\sigma}_e^2)=\rho(\sigma_e^2,\bar{\sigma}_e^2)/\sigma^4$. These figures highlight the features discussed so far. They also show that in many situations it is better to use the unrestricted estimator or the pre-test estimator, even if the restrictions are valid. We recall that when the model is mis-specified this case is somewhat more complicated as θ_n is no longer zero when H_0 is true unless θ_d is simultaneously zero or X and Z are orthogonal.

FIGURE 5 : Relative risk functions for $\tilde{\sigma}_e^2$, σ_e^{*2}, and $\hat{\sigma}_e^2$ when T=20, k=4, m=3, ν=5, and λ_d=0.

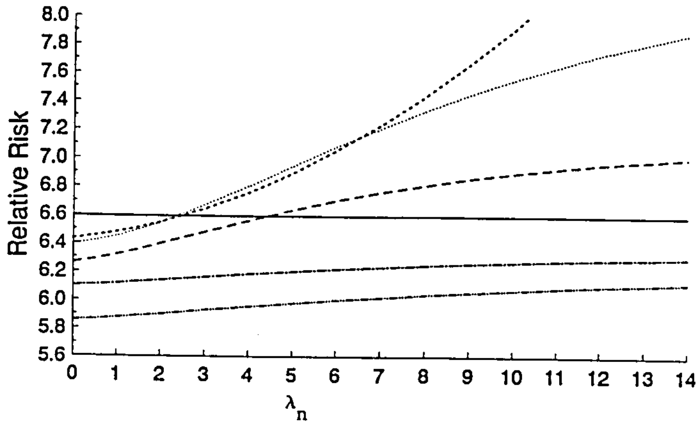
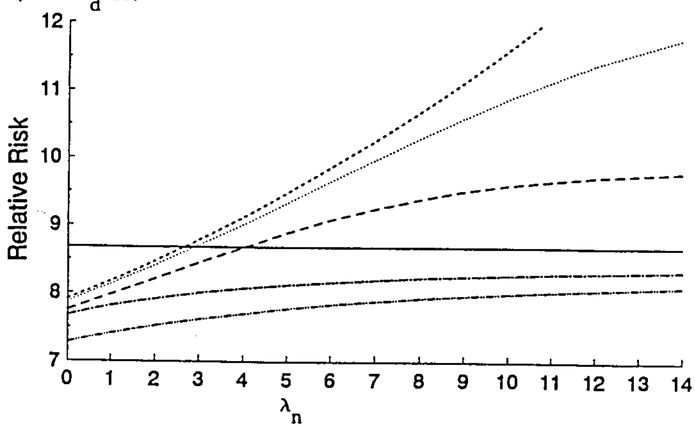


FIGURE 6 : Relative risk functions for $\tilde{\sigma}_e^2$, σ_e^{*2}, and $\hat{\sigma}_e^2$ when T=20, k=4, m=3, ν=5, and λ_d=10.



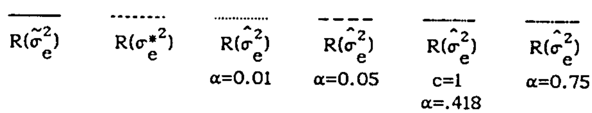


FIGURE 7 : Relative risk functions for $\tilde{\sigma}_e^2$, σ_e^{*2}, and $\hat{\sigma}_e^2$ when $T=20$, $k=4$, $m=3$, $\nu=\infty$, and $\lambda_d=0$.

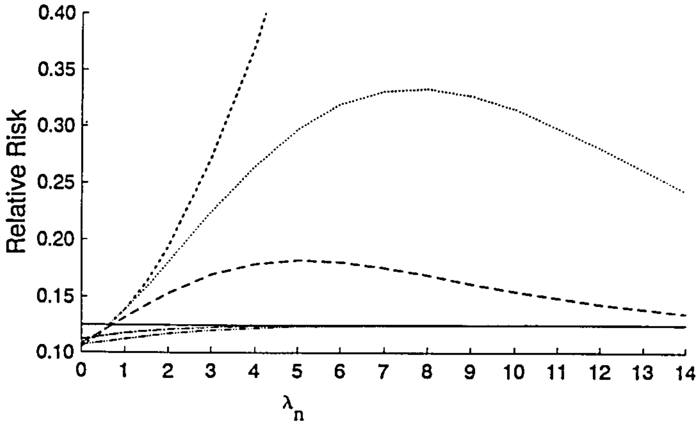
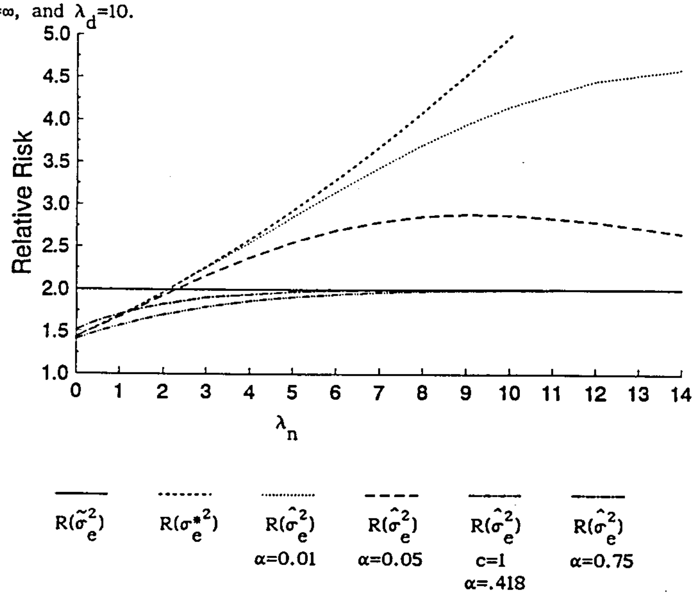


FIGURE 8 : Relative risk functions for $\tilde{\sigma}_e^2$, σ_e^{*2}, and $\hat{\sigma}_e^2$ when $T=20$, $k=4$, $m=3$, $\nu=\infty$, and $\lambda_d=10$.



Our numerical evaluations suggest that when $e \sim Mt\left(0, \nu\sigma^2/(\nu-2)I_T\right)$ it is typically better to always pre-test using c=1 for all feasible ν when $\lambda_d > 0$, rather than to impose even valid restrictions. This finding concurs with that of Giles (1991) when the design matrix is correctly specified for small values of ν, say $\nu < 15$, and for all values of ν when m=1. She found, however, that for higher values of ν and m>1 there exists a range (though sometimes a very small range) in the neighbourhood of the null over which the restricted estimator has smaller risk.

We also see from the figures that an increase in λ_d, *ceteris paribus*, shifts the risk functions upwards - there is a risk penalty for mis-specifying the model; it increases the maximum regret of $R(\hat{\sigma}_e^2)$ from that of $R(\tilde{\sigma}_e^2)$; it decreases the rate at which $R(\hat{\sigma}_e^2) \to R(\tilde{\sigma}_e^2)$; and it increases the λ_n range over which we prefer pre-testing (for all α's) to imposing the restrictions. When $\lambda_d = 0$ and α is small, say 1%, there is a region over which pre-testing has the highest risk. Once we admit that λ_d can be non-zero then this range decreases and in most cases even pre-testing with this test size is preferable to imposing the restrictions without testing their validity.

5. CONCLUDING REMARKS

In this paper we have investigated the risk under quadratic loss of estimators of the prediction vector and of the error variance in a model which may have variables omitted from the design matrix and whose distribution of the errors may be wider than the usual normality assumption, after a pre-test for exact linear restrictions. We find that the mis-specification of the distribution of the regression disturbances has little impact on the qualitative properties of the risk functions of the estimators of the prediction vector, and that the results of Mittelhammer (1984) assuming normal errors carry over to the broader problem that we investigate.

For this estimation problem we have not considered the question of the choice of an optimal test size. Giles *et al.* (1990) show in the normal errors case that the optimal critical values suggested by Brook (1976), for instance, are no longer valid if we exclude relevant regressors. They find that then the optimal critical value, according to the minimax regret criterion used by Brook, depends on the degree of mis-specification. Further, Giles (1991) shows that the Brook critical values are no longer valid if the errors are SSD_N as opposed to simply normally distributed even if we have not omitted regressors. The question of the optimal critical value when we have mis-specified both the design matrix and the error distribution is a topic of current research.

Our analysis has shown that mis-specifying the model can have a substantial impact on the risk functions of the investigated estimators of the error variance. For instance, if the errors are Mt then imposing the linear restrictions, even if they are valid, is rarely the optimal strategy, whether or not the design matrix is mis-specified. We also find that whether or not the design matrix is mis-specified it is generally better to pre-test, and if using the least squares component estimators it is best to use a critical value of unity. Generally, if the design matrix is mis-specified this pre-test estimator strictly dominates the other estimators investigated. Then the choice of the optimal test size is obvious. However, the problem of the choice of test size remains for those cases where we have no strictly dominating estimator.

ACKNOWLEDGEMENT

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

REFERENCES

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

Giles, D.E.A., (1986). Preliminary-test estimation in mis-specified regressions. *Economics Letters*, 21, 325-328.

Giles, D.E.A. and Clarke, J.A., (1989). Preliminary-test estimation of the scale parameter in a mis-specified regression model. *Economics Letters*, 30, 201-205.

Giles, D.E.A., Lieberman, O., and Giles, J.A., (1990). The optimal size of a preliminary test of linear restrictions in a mis-specified regression model. Discussion Paper No. 9006, Department of Economics, University of Canterbury.

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., (1991). Pre-testing for linear restrictions in a regression model with spherically symmetric disturbances. *Journal of Econometrics*, forthcoming.

Judge, G.G., Miyazaki, S., and Yancey, T.A., (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.

Mittelhammer, R.C., (1984). Restricted least squares, pre-test, OLS and Stein rule estimators: Risk comparisons under model misspecification, *Journal of Econometrics*, 25, 151-164.

Miyazaki, S., Judge, G.G., and Yancey, T.A., (1986). Estimation of location parameters under nonnormal errors and quadratic loss. *Journal of Business and Economic Statistics*, 4, 263-268.

Muirhead, R.J., (1982). *Aspects of multivariate statistical theory*. New York: John Wiley & Sons.

Ohtani, K., (1983). Preliminary test predictor in the linear regression model including a proxy variable. *Journal of the Japan Statistical Society*, 13, 11-19.

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

LIST OF DISCUSSION PAPERS*

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.

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.

(Continued on back cover)

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.

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.

* 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 1985 is available on request.