CANTER 9115

Department of Economics
UNIVERSITY OF CANTERBURY
CHRISTCHURCH, NEW ZEALAND

ISSN 1171-0705

PRE-TEST ESTIMATION IN A REGRESSION MODEL WITH A MISSPECIFIED COVARIANCE MATRIX

K. V. Albertson

Discussion Paper

No. 9115

Department of Economics, University of Canterbury
Christchurch, New Zealand

Discussion Paper No. 9115

November 1991

PRE-TEST ESTIMATION IN A REGRESSION MODEL WITH A MISSPECIFIED COVARIANCE MATRIX

K. V. Albertson

PRE-TEST ESTIMATION IN A REGRESSION MODEL WITH A MISSPECIFIED ERROR COVARIANCE MATRIX

K.V. ALBERTSON*

Department of Economics
University of Canterbury

November, 1991

Abstract

We consider the effects of incorrectly assuming a scalar error covariance matrix in a linear regression model in the context of a pre-test for linear restrictions on the coefficients. Because of this misspecification the (true) size and power of the pre-test may differ from their assumed values, distorting the pre-test estimator risk function towards that of one or other of its component estimators. The restricted and pre-test estimators may dominate the unrestricted estimator over a larger, or smaller, part of the parameter space, compared to the case with a correctly specified model.

*This paper is based on the author's Ph.D. research. The author is grateful to Professor D.E.A. Giles for his supervision and Dr J.A. Giles for her helpful comments. An earlier version of this paper was presented at the 1991 Australasian Meeting of the Econometric Society, Sydney.

1. Introduction

A pre-test situation arises in econometrics when the outcome of a statistical test determines which estimator is to be used for the parameter(s) of interest. For example, in the general linear model,

$$y = X\beta + \epsilon,$$

where y is $T \times 1$, X is $T \times K$, non-stochastic and of full rank and ϵ is assumed to be $N(0, \sigma^2 I_T)$, a test may be used to determine the validity of J exact linear restrictions described by

$$H_0 : R\beta = r \quad \text{vs} \quad H_A : R\beta \neq r.$$

Here, R is $J \times K$ and of rank J, r is $J \times 1$, and both are non-random. If, based on the outcome of the test, the restrictions are rejected, ordinary least squares (OLS) is used to estimate β. Otherwise restricted least squares (RLS) is used. The OLS and RLS estimators of β are $b = (X'X)^{-1}X'y$ and $b^* = b + (X'X)^{-1}R'(R(X'X)^{-1}R')^{-1}(r - Rb)$ respectively.

It is well known that, given the assumptions of the model, a UMPI test of these restrictions can be constructed using the statistic

$$u = \frac{(Rb-r)'(RS^{-1}R')^{-1}(Rb-r)(T-K)}{(y-Xb)'(y-Xb)\ J}, \tag{1}$$

where $S \equiv X'X$. This statistic has a central $F_{(J,T-K)}$ distribution under the null hypothesis and a non-central $F'_{(J,T-K;\lambda)}$ distribution under the alternative hypothesis, where the (numerator) non-centrality parameter is

$$\lambda \equiv \frac{(R\beta-r)'(RS^{-1}R')^{-1}(R\beta-r)}{2\sigma^2}. \tag{2}$$

A pre-test estimator (PTE) of β can be constructed using this statistic and the OLS and RLS estimators of β. This PTE can be written

$$\hat{\beta} = \begin{cases} b^* & \text{if } u < c(\alpha) \\ b & \text{if } u \geq c(\alpha), \end{cases}$$

where $c(\alpha)$ is the critical value chosen for the test when the nominal test size is α. Bancroft (1944), Brook (1976), Wallace and Ashar (1972), Wallace (1977) and Judge and Bock (1978, 1983), among others, discuss the properties of this estimator.

Recently, several authors have considered the properties of pre-test estimators in models which are misspecified in various ways: see, for example, Ohtani (1983), Mittelhammer (1984), Giles (1986) and Giles (1991*a* & *b*, 1992), among others. This paper extends the aforementioned studies by considering the consequences of failing to take into account a non-scalar error covariance matrix in the general linear model when applying a pre-test for exact linear restrictions.

The true covariance matrix of the error term is unknown and the tests and estimators applied to the model are based on the assumption that $\epsilon \sim N(0, \sigma^2 I_T)$. However, if $\epsilon \sim N(0, \Omega)$, with $\Omega \neq \sigma^2 I_T$, the assumed properties of the test statistic, and the OLS, RLS and PT estimators no longer hold. The hypothesis to be tested involves J independent linear restrictions, and the test is based on the usual statistic given in (1) above, because the researcher is unaware that the model is misspecified.[1]

In section 2 the bias and risk of the two component estimators (OLS and RLS) are stated, and results are given which are used to calculate the (true) size and power of the pre-test and the bias and risk of the PTE. Details of the specific models used to evaluate these formulae are contained in section 3 and the numerical results obtained are summarized in section 4. Section 5 concludes the paper.

2. The power of the test and properties of the estimators

Let the bias and risk (under quadratic loss) of an estimator $\hat{\xi}$ for an unknown parameter vector ξ be defined as $B(\xi) = E(\hat{\xi}) - \xi$ and $\rho(\hat{\xi}) = E\left[(\hat{\xi}-\xi)'(\hat{\xi}-\xi)\right]$ respectively.

It is well known that the OLS estimator of the β vector is not biased by the misspecification of the error covariance matrix and that its risk is given by $\rho(b) = tr(S^{-1}X'\Omega XS^{-1})$.

The bias and risk of the RLS estimator are given by $B(b^*) = -\eta\delta$ and $\rho(b^*) = \rho(b) - tr(2S^{-1}X'\Omega XC - CX'\Omega XC) + \delta'\eta'\eta\delta$, where $\delta \equiv (R\beta - r)$ (i.e., the hypothesis error), $\eta \equiv S^{-1}R'(RS^{-1}R')^{-1}$ and $C \equiv S^{-1}R'(RS^{-1}R')^{-1}RS^{-1}$. Using this notation we can write the test statistic in (1) as a ratio of two quadratic forms in the normal random vector $\bar{\epsilon} \equiv \epsilon + X\eta\delta$:

$$u = \frac{\bar{\epsilon}'XCX'\bar{\epsilon}}{\bar{\epsilon}'M\bar{\epsilon}}\frac{(T-K)}{J},$$

where $M \equiv (I_T - XS^{-1}X')$, and $\bar{\epsilon} \sim N(X\eta\delta, \Omega)$. If $XCX'\Omega$ and $M\Omega$ are idempotent and orthogonal then u is distributed as an F random variable, under H_0 (see Searle (1982 p.356)). In general, however, $XCX'\Omega$ and $M\Omega$ are neither idempotent nor independent and u is not F-distributed under the null.

Now consider the $T \times T$ symmetric matrix $\Phi \equiv \Omega^{\frac{1}{2}}\left(XCX' - \left(\frac{Jc(\alpha)}{(T-K)}\right)M\right)\Omega^{\frac{1}{2}}$. This matrix has T real eigenvalues, denoted λ_i, $(i = 1,\ldots,T)$, and a $T \times T$ orthonormal eigenvector matrix, Υ. The eigenvector corresponding to λ_i is the i'th column of Υ, denoted Υ_i.

Following Koerts and Abrahamse (1969) it is easily shown that $Pr.(u < c(\alpha)) = Pr.(z'\Lambda z < 0) = Pr.\left(\sum_{i=1}^{T}\lambda_i\chi^{2\prime}_{(1,\theta_i)} < 0\right)$, where $\Lambda = diag\{\lambda_i\}$, $z = \Upsilon'\Omega^{(-1/2)}\bar{\epsilon}$ and the non-centrality parameters of the independent χ^2 variates are $\theta_i \equiv \frac{1}{2}\left(\Upsilon_i\Omega^{(-1/2)}X\eta\delta\right)^2$. Using these formulae we can calculate the true size and power of the test using, for example, the algorithms of Imhof (1961) or Davies (1980).

In the well specified pre-test problem the test statistic is $F'_{(J,T-K;\lambda)}$ distributed, as noted above, and the overall level of hypothesis error in the model is represented by the non-centrality parameter, λ, given in (2) above. If the error covariance matrix is misspecified the distribution of the test statistic can be represented by a weighted sum of $\chi^{2\prime}_{(1,\theta_i)}$ variables and we can define a unit measure of the level of hypothesis error in the model as

$$\theta \equiv \sum_{(i:\lambda_i\ 0)}\theta_i = \sum_{(i:\lambda_i>0)}\frac{1}{2}\left(\Upsilon_i\Omega^{(-1/2)}X\eta\delta\right)^2.$$

If $\Omega = \sigma^2 I_T$ then $\theta = \lambda$ and $\Pr.(z'\Lambda z < 0) = \Pr.(F'_{(J,T-K;\lambda)} < c(\alpha))$.

Theorem 1

Under the stated assumptions, the bias and risk functions of the PTE are

$$B(\hat{\beta}) = -\ S^{-1}R'(RS^{-1}R')^{-1}RS^{-1}X'\Omega^{(1/2)}\Upsilon\ P_3\Upsilon'\Omega^{(-1/2)}X\eta\delta \tag{3}$$

and
$$\begin{aligned}\rho(\hat{\beta}) = \operatorname{tr}(S^{-1}X'\Omega XS^{-1}) &- 2\ \operatorname{tr}\left(CX'\Omega^{(1/2)}\Upsilon\ B\ \Upsilon'\Omega^{(1/2)}XS^{-1}\right) \\ &+ 2\left(\delta'\eta'CX'\Omega^{(1/2)}\Upsilon\ P_3\Upsilon'\Omega^{(-1/2)}X\eta\delta\right) \\ &+ \operatorname{tr}\left(CX'\Omega^{(1/2)}\Upsilon\ B\ \Upsilon'\Omega^{(1/2)}XC\right) \end{aligned} \tag{4}$$

respectively, where B is a T x T matrix with ij'th element

$$B_{ij} \equiv \begin{cases} P_{3i} + \left(\Upsilon_i\Omega^{(-1/2)}X\eta\delta\right)^2 P_{5i} & ;\ \text{when } i=j \\ \left(\Upsilon_i\Omega^{(-1/2)}X\eta\delta\right)\left(\Upsilon_j\Omega^{(-1/2)}X\eta\delta\right)P_{3ij}; & \text{when } i\neq j\ , \end{cases}$$

and $\qquad P_3 \equiv \operatorname{diag}\{P_{3i}\},$

with $\qquad P_{mi} = \Pr.\left(\lambda_i\chi^{2\prime}_{(m,\theta_i)} + \sum_{j\neq i}^{T}\lambda_j\chi^{2\prime}_{(1,\theta_j)} < 0\right),\quad m = 3,5,$

and $\qquad P_{3ij} = \Pr.\left(\lambda_i\chi^{2\prime}_{(3,\theta_i)} + \lambda_j\chi^{2\prime}_{(3,\theta_j)} + \sum_{k\neq i,j}^{T}\lambda_k\chi^{2\prime}_{(1,\theta_k)} < 0\right),\ i,j = 1,\ldots,T.$

Proof

See Appendix A.

As $c(\alpha) \rightarrow 0\ (\infty)$, (3) and (4) collapse to the risk and bias of the unrestricted (restricted) estimators. These formulae can also be shown to be equivalent to those for the PTE in a correctly specified model (e.g., Judge and Bock (1978)); that is when $\Omega = \sigma^2 I_T$.

As these expressions are complicated it is difficult to determine the effects of the misspecification without numerical evaluations. This has been done using the SHAZAM package (White *et al.* (1990)) and Davies' (1980) algorithm on a Vax 6340 computer.

3. The models

We consider the effects of the misspecification of the error covariance matrix on the risks of the OLS, RLS and PT estimators in a number of models. Several different data sets are used in each model as the associated formulae are data dependent. The data are described in Appendix B.

3.1 Autoregressive Errors

We consider testing the significance of one or more of the regressors in the quarterly regression model

$$y_t = x_t\beta + u_t; \qquad t = 1,\ldots,T,$$

where x_t is the t'th row of the matrix X, and u_t is generated either by an AR(1) process, $(1 - \rho_1 L)u_t = \epsilon_t$, or an AR(4) process, $(1 - \rho_4 L^4)u_t = \epsilon_t$, where $-1 < \rho_i < 1$, $(i = 1,4)$, $\epsilon \sim N(0, \sigma^2 I_T)$ and $L^j u_t = u_{t-j}$. The exclusion restrictions are written in the usual way as H_0: $R\beta = r$ vs H_A: $R\beta \neq r$.

Bounds on the true size of a test for restrictions in the general linear model with an ARMA process in the error term are calculated by Kiviet (1980), while the effects of an AR(1) process in the errors on the true size of a Chow test are considered by Consiglieri (1981) and Giles and Scott (1991). It has been argued that an AR(4) process may also be common in regressions using quarterly data (e.g., Thomas and Wallis (1971) and King (1989)).

3.2 Moving Average Errors

The situation we consider is as above, except that the error term is generated by the MA(1) process $u_t = (1 + \tau L)\epsilon_t$, where $-1 < \tau < 1$ and $\epsilon \sim N(0, \sigma^2 I_T)$. The effect of such a process on the true size of a Chow test is also considered by Giles and Scott (1991).

3.3 Heteroscedastic Errors

Model (a) We consider the application of the Chow (1960) test for structural change in the model

$$y \equiv \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix}\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} \equiv X\beta + \epsilon,$$

where $\epsilon \sim N\left(0,\ \sigma_1^2 \begin{bmatrix} I_{T_1} & 0 \\ 0 & \varphi I_{T_2} \end{bmatrix} \right)$, $T_1 = T_2 = T/2$.

The parameter φ measures the degree of heteroscedasticity in the errors. In this case the null and alternative hypotheses are

$$H_0\colon \beta_1 = \beta_2 \quad \text{vs} \quad H_A\colon \beta_1 \neq \beta_2.$$

The properties of the Chow test in the presence of heteroscedasticity of this form are considered by Toyoda (1974) and Schmidt and Sickles (1977), among others. Giles and Lieberman (1991b) calculate bounds on the true size of the test in this case.

In our study, the other five models consider the case of testing the significance of one or more of the regressors in the model

$$y_t = x_t\beta + \epsilon_t, \qquad t = 1,\dots,T,$$

where x_t is the t'th row of the matrix X and $var(\epsilon_t)$ is some function of a variable z_t. The exclusion restrictions are written as H_0: $R\beta = r$ vs H_A: $R\beta \neq r$. The five models differ in the functions which are used to form the error covariance matrix. If γ is defined as a constant parameter, tr as a linear trend variable and $\bar{x}$ as one of the columns of the X matrix, the five different functional forms considered are;

<u>Model (b)</u> $\quad var(\epsilon) = \sigma_1^2 \begin{bmatrix} I_{T_1} & 0 \\ 0 & \varphi I_{T_2} \end{bmatrix}$, $T_1 = T_2 = T/2$.

<u>Model (c)</u> $\quad var(\epsilon_t) = f_c(z_t) = \sigma^2 \bar{x}_t^{\gamma}$,

<u>Model (d)</u> $\quad var(\epsilon_t) = f_d(z_t) = \sigma^2 \exp(\gamma\bar{x}_t)$,

<u>Model (e)</u> $\quad var(\epsilon_t) = f_e(z_t) = \sigma^2 tr_t^{\gamma}$,

<u>Model (f)</u> $\quad var(\epsilon_t) = f_f(z_t) = \sigma^2 \exp(\gamma tr_t)$.

For models (c) to (f) the measure of heteroscedasticity, φ, is defined as

$$\varphi = \frac{f_i\left(\max_{t=1..T}(z_t) \right)}{f_i\left(\min_{t=1..T}(z_t) \right)}, \quad i = c, d, e \text{ or } f.$$

A value of φ of greater than unity implies that the error variance is increasing as z_t increases, while a φ value of less than unity implies that the error variance is decreasing as z_t increases.

4. Numerical Results

The nominal test size is fixed at 5% in the discussion which follows. There is no size correction applied to the test as we wish to determine the consequences of assuming that the error term is well behaved when in fact it may not be. For the purposes of this discussion the term "power" refers to the size uncorrected power of the test.

Typical OLS, RLS and PTE risk functions for a regression model that is correctly specified (i.e., $\varphi = 1$, $\rho_1, \rho_4, \tau = 0$) are shown in Figures 1, 3 and 5, Appendix C.[2] Quantitatively, the presence of an autoregressive, moving average or heteroscedastic process in the error term may increase, or slightly reduce, the risks of the estimators for each value of θ. Qualitatively, the misspecification introduces a bias in the pre-test power function and changes the relative dominance of the three estimators.

The effect of the misspecification on the true size and power of the pre-test depends on a number of different factors, including the number of regressors in the model, the particular characteristics of the regressors and the form of the true error covariance matrix. For example, if the significance of a (group of) trended regressors is being tested, the true size and power of the test increase with increasing τ, in the case of MA(1) errors, or with increasing ρ_1,[3] in the case of AR(1) errors. The converse may occur if the regressors are not trended. As the value of ρ_1 or τ, decreases below zero the opposite effect is observed with the (true) size and power of the test falling for most regressor sets.[4]

In general, if the errors are generated by an AR(4) process, the power of the pre-test is reduced if the absolute value of ρ_4 is close to unity. An exception to this is the case of testing the joint significance of a set of seasonal dummy variables where the true size and power of the test increase with increasing values of ρ_4. The opposite effect is observed as ρ_4 decreases below zero. Figure 2 illustrates the effects of a downward distortion in the power function on the PTE risk. Comparing Figure 2 with Figure 1 we see that the PTE risk is closer to the RLS risk at each level of θ than is assumed to be the case.

When the errors are heteroscedastic, an increase in the degree of heteroscedasticity appears likely to increase the true size and power of the test for small values of θ, and to reduce the power for large values of θ, if there are three or more regressors. In models with less than three regressors, there is no consistent pattern.

When the errors are generated by an AR(1) process, any increase in the value of ρ_1 generally has the effect of decreasing the range of θ over which it is preferable to pre-test rather than to simply ignore prior information and estimate using OLS. With some regressor sets OLS may strictly dominate both the RLS and the PT estimators. In this case the imposition of valid restrictions serves to increase the estimator risk.[5]

Conversely, if ρ_1 decreases the range of θ over which it is preferable to pre-test generally increases, making the use of the PTE more attractive relative to OLS. An example of this is shown in Figure 4. Comparing Figure 4 with Figure 3 we see that the PTE dominates OLS over a larger part of the θ space than is assumed to be the case and also that there exits a range of θ over which the PTE dominates both of its component estimators. This is in contrast to the usual result, that in a correctly specified model the PTE of β is never the minimum risk estimator and will have higher risk than both of its two component estimators over some part of the θ range.[6] When the errors are formed by a moving average process, increasing values of τ appear to have little or no effect on the relative dominance of the estimators.

In the models with heteroscedastic errors, increasing levels of heteroscedasticity are more likely to increase the range of θ over which the PTE risk is lower than the OLS risk than to reduce it, particularly if the regressor variables are trended. There are instances, however, when the converse occurs. In some cases the RLS and PT estimators may become completely dominated by OLS as shown in Figure 6. However, no general result is apparent as the effect of the misspecification varies with both the type of heteroscedasticity and the form of the regressor variables.

Other things being equal, an increase in the sample size leads to a reduction in estimator risk for each value of θ. However such an increase may not alleviate the distortions introduced to the models as a result of the misspecification. The (size uncorrected) power of the test may, in fact, be reduced by such an increase if the misspecification is severe and θ is close to zero. Also, the range of θ over which the PTE has a lower risk than the OLS estimator may be further reduced or increased by an increase in sample size.

In general the effects of increasing the sample size are ambiguous, particularly with real or non-trended data. This may be because the additional data points may change the characteristics of the regressor set and all of the models considered are very sensitive to the form of the data.

5. Conclusion

The practical implications of the misspecification vary depending on the type of misspecification and the regressors. Because of this, little can be offered by way of a general prescription. Some points, however, can be made.

When the errors are generated by an AR or MA process with positive coefficients, and the regressors whose coefficients are included in the restrictions are trended, the PTE may be strictly dominated by OLS, in which case it is better to ignore the prior information. Even if the PTE is not strictly dominated, the θ range over which the risk of the PTE is lower than OLS is generally reduced (compared to the correctly specified model). The regret associated with using OLS rather than the PTE in that θ range is reduced also, as the pre-test power function is likely to be distorted upwards and the test will tend to over-reject valid restrictions. Hence, although in practice the degree of distortion is unknown, it may be preferable to ignore the prior information rather than pre-test if an autocorrelation problem is suspected.

Conversely, if the errors are generated by an AR or MA process with negative coefficients, or if the regressors whose coefficients are included in the restrictions are not trended, the power of the pre-test is likely to be reduced by the misspecification and the PTE will dominate OLS over a greater portion of the θ range compared to the correctly specified model. Therefore it is likely to be preferable to pre-test rather than ignore the prior information in this case.

On the basis of these results, it appears advisable to test for such processes before any testing of linear restrictions is carried out. If the linear restrictions involve the coefficients of trended regressors, it may be wise to choose a critical value such that the test has a higher power against a positive AR or MA process than against a negative process, as the costs of failing to correct for a positive process are the greater of the two. The converse is true if the restrictions involve the coefficients of non-trended regressors. However, it should be noted that there are further implications associated with such multiple pre-testing (eg, see King and Giles (1984) and Giles and Lieberman (1991*a*)).

If the errors are possibly heteroscedastic, there is no general prescription as, although increasing heteroscedasticity may increase the pre-test size, the power of the pre-test will be reduced in models with more than two regressors. Also, as we have seen, the PTE may be strictly dominated by OLS, in which case the prior information should be ignored. However, given that the true error covariance matrix is unknown, the effect of the misspecification on a given model cannot be determined.

Because of the effects of the misspecification on the (true) size and power of the test, any attempt to apply an "optimal" critical value, such as is suggested by Brook (1976), will not necessarily lead to an "optimal" pre-test risk. This is analogous to the findings of Giles, Lieberman and Giles (1990), who generalise Brook's result to the case where the model is misspecified by the exclusion of relevant regressors, and Wong and Giles (1991) who consider the problem of determining the optimal critical value for a pre-test in a model with Multivariate Student-t disturbances.

Footnotes

1 Other pre-testing papers which consider a possible non-scalar error covariance matrix include Greenberg (1980) and Mandy (1984), who consider the problem of testing for a heteroscedastic process in the error term, and Fomby and Guilkey (1978), King and Giles (1984), and Giles and Lieberman (1991a), among others, who consider the problem of testing for the presence of serial correlation in the error term. The problem considered in this paper differs from their problems in that the pre-test is applied, not to determine whether or not the error term in the model is well behaved, but to determine the validity of linear restrictions.

2 Although the risk functions are data dependent, there are no significant qualitative differences in these functions for the majority of the different regressor sets considered in this study if the model is well specified. This is not the case if the model is misspecified.

3 A similar result is found by Consiglieri (1981) and Giles and Scott (1991) in the case of the size of the Chow test, and this is consistent with Kiviet's (1980) results.

4 This is also consistent with Giles and Scott's (1991) results.

5 A similar result is obtained by Giles and Giles (1991) in the case of estimating the regression scale parameter if a sufficiently asymmetric loss function is used. It is also partially analogous to Mittelhammer's (1984) result that in a model misspecified by the exclusion of relevant regressors the RLS and PT estimators may be dominated by OLS.

6 Another example of the possible dominance of a PTE is given in Ohtani (1983), who shows that, in a model in which a relevant, unobservable, variable has been replaced by a proxy variable, there may exist a region in which the PTE risk is lower than the risks of its two component estimators. Other examples are given in Giles (1991*a* & *b*), Giles and Giles (1991), and Wong and Giles (1991).

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.

Brook, R.J. (1976), On the use of a regret function to set significance points in prior tests of estimation, *Journal of the American Statistical Association* 71, 126-131.

Chow, G.C. (1960), Tests of equality between sets of coefficients in two linear regressions, *Econometrica* 28, 591-605.

Consiglieri, I. (1981), The Chow test with serially correlated errors, *Rivista Internazionale di Scienze Sociali* 89, 125-137.

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

Fomby T.B. and D.K. Guilkey (1978), On Choosing the optimal level of significance for the Durbin-Watson test and the Bayesian alternative, *Journal of Econometrics* 8, 203-213.

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

Giles, D.E.A. & O. Lieberman (1991a), Some Properties of the Durbin-Watson test after a preliminary t-test, *Journal of Statistical Computation and Simulation*, forthcoming.

Giles, D.E.A. and O. Lieberman (1991b), Bounds on the effects of heteroscedasticity on the Chow test for structural change, Discussion Paper 9101, Department of Economics, University of Canterbury.

Giles, D.E.A., O. Lieberman and J.A. Giles (1990), The optimal size of a preliminary test of linear restrictions in a mis-specified regression model, Discussion Paper 9006, Department of Economics, University of Canterbury.

Giles, D.E.A. and M. Scott, Some consequences of using the Chow test in the context of autocorrelated disturbances, Discussion Paper 9112, Department of Economics, University of Canterbury.

Giles, J.A. (1991a), Pre-testing in a mis-specified regression model, *Communications in Statistics* A 20, 3221-3238.

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

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*, forthcoming.

Giles, J.A. and D.E.A. Giles (1991), Preliminary-test estimation of the regression scale parameter when the loss function is asymmetric, Discussion Paper 9104, Department of Economics, University of Canterbury.

Greenberg, E. (1980), Finite sample moments of a preliminary-test estimator in the case of possible heteroscedasticity, *Econometrica* 48, 1805-1813.

Imhof, P. (1961), Computing the distribution of quadratic forms in normal random variables, *Biometrika* 48, 419-426.

Judge, G.G. and M.E. Bock (1978), *The Statistical Implications of Pre-test and Stein Rule Estimators in Econometrics,* Amsterdam, North-Holland.

Judge, G.G. and M.E. Bock (1983), Biased estimation, in Z. Griliches and M.D. Intriligator (eds.), *Handbook of Econometrics,* Amsterdam, North-Holland, 599-649.

King, M.L. (1989), Testing for fourth-order autocorrelation in regression disturbances when first order autocorrelation is present, *Journal of Econometrics* 41, 285-301.

King, M.L. and D.E.A. Giles (1984), Autocorrelation pre-testing in the linear model: Estimation, testing and prediction, *Journal of Econometrics*, 25, 35-48.

Kiviet, J.F. (1980), Effects of ARMA errors on tests for regression coefficients: Comments on Vinod's article; improved and additional results, *Journal of the American Statistical Association* 75, 353-358.

Koerts, J. and A.P.J. Abrahamse (1969), *On the Theory and Application of the General Linear Model*, Rotterdam, Rotterdam University Press.

Mandy, D.M. (1984), The moments of a pre-test estimator under possible heteroscedasticity, *Journal of Econometrics*, 25, 29-33.

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

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

Schmidt, P. and R. Sickles (1977), Some further evidence on the use of the Chow test under heteroskedasticity, *Econometrica* 45 1293-1298.

Searle, S.R. (1982) *Matrix Algebra Useful for Statistics*, New York, John Wiley and Sons.

Thomas, J.J. and K.F. Wallis (1971), Seasonal variation in regression analysis, *Journal of the Royal Statistical Society* A 134, 57-72.

Toyoda, T. (1974), Use of the Chow test under heteroscedasticity, *Econometrica* 42, 601-608.

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

Wallace, T.D. and V.G. Ashar (1972), Sequential methods in model construction *Review of Economics and Statistics* 54, 172-178.

White, K.J., S.D. Wong, D. Whistler and S.A. Haun (1990), *SHAZAM Econometrics Computer Package: User's Reference Manual, Version 6.2*, New York, McGraw-Hill.

Wong, J.K. and J.A. Giles (1991), Optimal critical values of a preliminary test for linear restrictions in a regression model with multivariate Student-t disturbances, Discussion Paper 9114, Department of Economics, University of Canterbury.

APPENDIX A

To prove Theorem 1 we require the following lemma.

Lemma

Let $\Psi(x)$ be an indicator function such that $\Psi(x) = \begin{cases} 0 \text{ if } x \geq 0 \\ 1 \text{ if } x < 0 \end{cases}$.

Then, using the notation of section 2,

a) $E[z'\Psi(z'\Lambda z)] = P_3\Upsilon'\Omega^{(-1/2)}X\eta\delta$,

b) $E[zz'\Psi(z'\Lambda z)] = B$.

Proof

a) Consider the i'th element of the T × 1 vector $E[z\Psi(z'\Lambda z)]$,

$$E[z_i\Psi(z'\Lambda z)] = E\left[z_i\Psi\left(\lambda_i z_i^2 + \sum_{j\neq i}^{T} \lambda_j z_j^2\right)\right].$$

$\mathrm{Covar}(z) = I_T$ and therefore z_i is independent of z_j, $i \neq j$. Hence

$$E[z_i\Psi(z'\Lambda z)] = \underset{j\neq i}{E}\left[\underset{i}{E}\left[z_i\ \Psi\left(\lambda_i z_i^2 + \sum_{j\neq i}^{T}\lambda_j z_j^2\right)\right]\right]$$

$$= \underset{j\neq i}{E}\left[\underset{i}{E}\left[\Upsilon_i'\Omega^{(-1/2)}X\eta\delta\Psi\left(\lambda_i\chi^{2\prime}_{(3,\theta_i)} + \sum_{j\neq i}^{T}\lambda_j z_j^2\right)\right]\right],$$

by Lemma 2 of Judge and Bock (1978, p.320).

Now, because $E[\Psi(x)] = \Pr(x < 0)$ by definition,

$$E[z_i\Psi(z'\Lambda z)] = P_{3i}\Upsilon_i'\Omega^{(-1/2)}X\eta\delta, \qquad i = 1,\ldots,T,$$

therefore $\qquad E[z\Psi(z'\Lambda z)] = P_3\Upsilon'\Omega^{(-1/2)}X\eta\delta.$

b) Consider the ij'th element of the $T \times T$ matrix $E[zz'\Psi(z'\Lambda z)]$.

It is straightforward to show that when $i = j$,

$$E[z_i z_j \Psi(z'\Lambda z)] = P_{31} + \left(\Upsilon_i'\Omega^{(-1/2)}X\eta\delta\right)^2 P_{51},$$

by Lemma 2 of Judge and Bock (1978, p.320), and that, when $i \neq j$,

$$E[z_i z_j \Psi(z'\Lambda z)] = \left(\Upsilon_i'\Omega^{(-1/2)}X\eta\delta\right)\left(\Upsilon_j'\Omega^{(-1/2)}X\eta\delta\right)P_{31j},$$

by Lemma 1 of Judge and Bock (1978, p.320). □

Proof of Theorem 1

The PTE can be written $\hat{\beta} = \begin{cases} b & \text{if} \quad u \geq c(\alpha) \\ b + S^{-1}R'(RS^{-1}R')^{-1}(r-Rb) & \text{if} \quad u < c(\alpha) \end{cases}$.

Because $u < c(\alpha) \Leftrightarrow z'\Lambda z < 0$, we can write this as

$$\begin{aligned} \hat{\beta} &= b + S^{-1}R'(RS^{-1}R')^{-1}(r-Rb)\Psi(z'\Lambda z). \\ &= \beta + S^{-1}X'\epsilon + \eta(r-R\beta-RS^{-1}X'\epsilon)\Psi(z'\Lambda z) \\ &= \beta + S^{-1}X'\epsilon - CX'\Omega^{(1/2)}\Upsilon z\Psi(z'\Lambda z). \end{aligned}$$

The bias of the PTE is

$$\begin{aligned} B(\hat{\beta}) = E(\hat{\beta}) - \beta &= S^{-1}X' \; E(\epsilon) - CX'\Omega^{(1/2)}\Upsilon E[z\Psi(z'\Lambda z)] \\ &= - S^{-1}R'(RS^{-1}R')^{-1}RS^{-1}X'\Omega^{(1/2)}\Upsilon P_3\Upsilon'\Omega^{(-1/2)}X\eta\delta, \end{aligned}$$

by the Lemma given above.

The risk of the PTE is

$$\begin{aligned} \rho(\hat{\beta}) &= E[(\hat{\beta}-\beta)'(\hat{\beta}-\beta)] \\ &= E\big[\epsilon'XS^{-1}S^{-1}X'\epsilon - 2\epsilon'XS^{-1}CX'\Omega^{(1/2)}\Upsilon z\Psi(z'\Lambda z) \\ &\qquad + z'\Upsilon'\Omega^{(1/2)}XC\Psi(z'\Lambda z)CX'\Omega^{(1/2)}\Upsilon z\big]. \end{aligned}$$

Now, because $\epsilon = \Omega^{(1/2)}\Upsilon z - X\eta\delta$,

$$\begin{aligned} \rho(\hat{\beta}) &= \mathrm{tr}\left(S^{-1}X'\Omega XS^{-1}\right) - 2\mathrm{tr}\left(CX'\Omega^{(1/2)}\Upsilon B\Upsilon'\Omega^{(1/2)}XS^{-1}\right) \\ &\quad + 2\delta'\eta'CX'\Omega^{(1/2)}\Upsilon P_3\Upsilon'\Omega^{(-1/2)}X\eta\delta \\ &\quad + \mathrm{tr}\left(CX'\Omega^{(1/2)}\Upsilon B\Upsilon'\Omega^{(1/2)}XC\right) \end{aligned}$$

by the Lemma given above. □

Appendix B

The Regressor Data

As the bias, risk and power functions are data dependent, a number of real and artificial regressor series have been chosen to evaluate the test and estimator properties. The artificial regressor series are,

a) Random variables formed through the application of an AR(1) process (autocorrelation coefficient = 0.5) to standard normal data,

b) Log normal data, based on the standard normal distribution,

c) Exponentially trended data, (16.182% increase per period),

d) Standard normal data,

e) Linearly trended data,

f) Uniformly [0,1] distributed data.

For the heteroscedastic models the real regressor series chosen were,

g) Australian real GDP, (annual, 1960-89),

h) Australian money supply, (quarterly, 1960q1-89q4),

i) New Zealand real GDP index, (quarterly 1972q2-90q2)*,

j) Australian$ - U.S.$ spot rate, (quarterly, 1960q1-90q1).

For the autoregressive and moving average models the real regressor series chosen were,

k) Australian CPI, (quarterly 1960q1-90q2),

l) Australian real retail trade, (quarterly 1960q1-90q1),

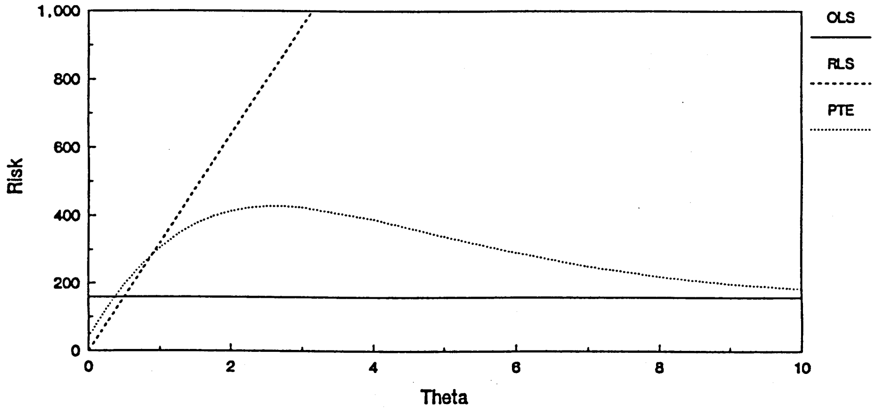
m) Australian trade balance, (quarterly 1960q1-90q2).

All real regressor series taken from the N.Z. Statistics Dept. I.N.F.O.S. database, source O.E.C.D., except *, source N.Z. Statistics Dept.

FIGURE 1 : Risk Functions

Correctly Specified Model

Regressors are a constant and Australian C.P.I.



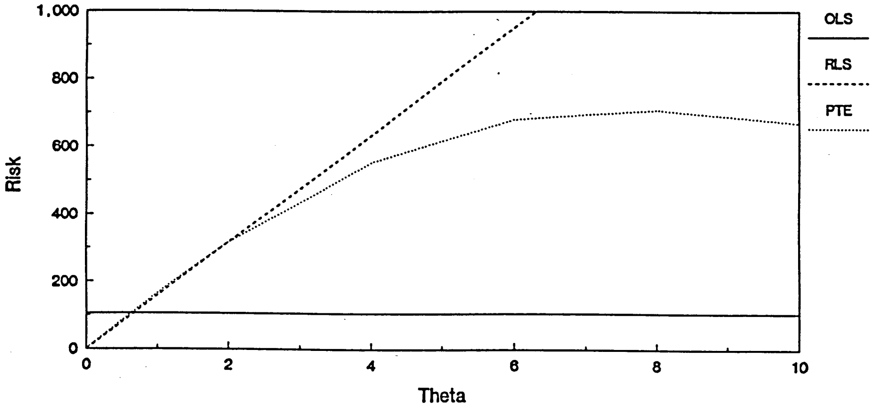
One restriction
alpha=5%

Sample Size=20

FIGURE 2 : Risk Functions

MA(1) errors; tau = -0.9

Regressors are a constant and Australian C.P.I.

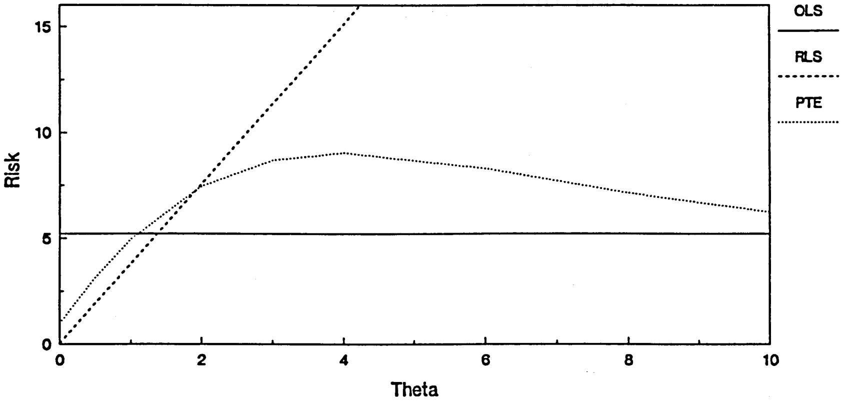


One restriction
alpha=5%

Sample Size=20

FIGURE 3 : Risk Functions

Correctly Specified Model
Regressors are const, trend & Aust. Retail Trade

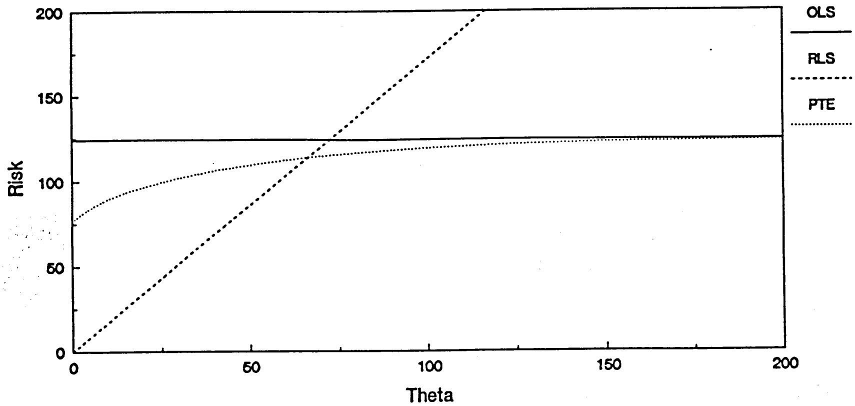


Joint test of linear trend and Retail Trade
alpha=5%

Sample Size=20

FIGURE 4 : Risk Functions

AR(1) errors; rho1 = −0.9
Regressors are const, trend & Aust. Retail Trade



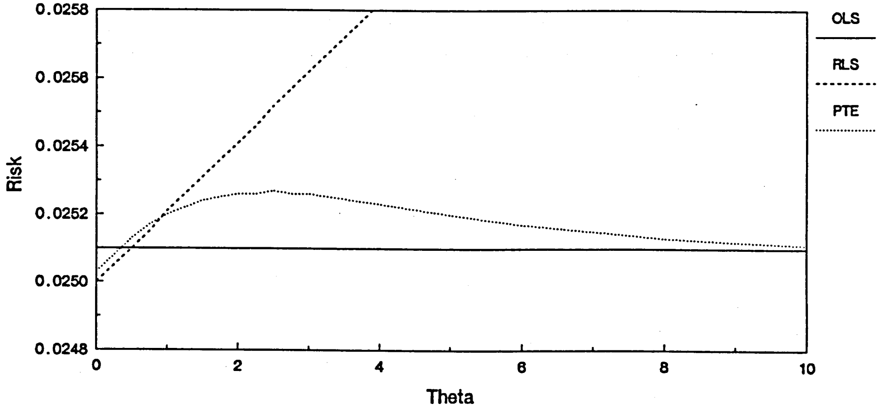
Joint test of linear trend and Retail Trade
alpha=5%

Sample Size=20

FIGURE 5 : Risk Functions

Correctly Specified Model

Regressors are a constant and a normal r.v.



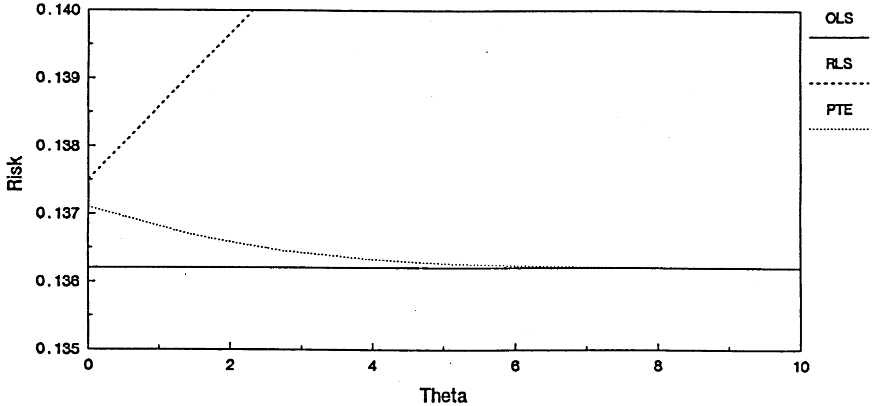
One restriction
alpha=5%

Sample Size=40

FIGURE 6 : Risk Functions

Heteroscedastic Errors (Model B); psi = 10.0

Regressors are a constant and a normal r.v.



One restriction
alpha=5%

Sample Size=40

LIST OF DISCUSSION PAPERS*

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.

(Continued on next page)

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.

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 Covariance Matrix, by K. V. Albertson.