CANTER

9308

Department of Economics
UNIVERSITY OF CANTERBURY
CHRISTCHURCH, NEW ZEALAND

ISSN 1171-0705

COMPARING STANDARD AND ROBUST SERIAL CORRELATION TESTS IN THE PRESENCE OF GARCH ERRORS

John P. Small

Discussion Paper

No. 9308

Department of Economics, University of Canterbury
Christchurch, New Zealand

Discussion Paper No. 9308

June 1993

COMPARING STANDARD AND ROBUST SERIAL CORRELATION TESTS IN THE PRESENCE OF GARCH ERRORS

John P. Small

COMPARING STANDARD AND ROBUST SERIAL CORRELATION TESTS IN THE PRESENCE OF GARCH ERRORS

John P. Small*

Department of Economics

University of Canterbury

June 1993

Abstract

Recently, Diebold (1986) and Wooldridge (1991) have suggested procedures for ensuring that well known tests for serial independence have asymptotically reliable sizes in the presence of conditional heteroscedasticity. This paper uses a Monte Carlo experiment to compare the sizes and powers of several versions of these robust tests with their "non-robust" forms and with standard exact tests. The general conclusion is that both robust procedures lack power and are dominated by well specified exact tests. This conclusion is not altered when the assumption of normally distributed innovations is relaxed.

* This paper has benefitted from comments recieved during seminars at the Universities of Canterbury and Auckland, Monash University and the Royal Melbourne Institute of Technology. In particular, thanks are due to David Giles, Max King and Kim Sawyer.

1. Introduction

In most applications of the GARCH model (Engle (1982), Bollerslev (1986)) the conditional mean is specified prior to the modelling of the conditional variance (see Chou (1988), Hsieh (1989) or Baillie and Bollerslev (1989) for example). While this is a natural order in which to approach the modelling task, it also raises important questions about the properties of tests used to detect autocorrelation in the mean equation, when the errors follow a GARCH process. Diebold (1986) addressed this issue by using results from Milhöj (1985) to show that modified forms of the Box-Pierce (1970) and Ljung-Box (1978) tests have the correct sizes asymptotically. Wooldridge (1991) considered the standard Lagrange multiplier (LM) test for serial independence in this context and constructed a set of modified LM tests which are similarly robust.

Although Diebold (1986) presented empirical evidence, using an observed series, which supported his claim for the asymptotic size of his procedure, he did not consider the power of the resulting tests. Wooldridge (1991) conducted no empirical study of his modified LM test. Given the strong influence of the GARCH model for regression error variances, there is a need to examine the relative performance of the various methods available for specifying the mean equation in such models. This paper reports on a study designed to clarify such issues and provide some guidance for applied workers in the early stages of modelling. A recent investigation by Silvapulle and Evans (1993) addresses similar issues, concentrating on the size of tests for autocorrelation under a variety of non-normal conditional distributions.

The next section describes the models and tests used in the study. Section 3 outlines the design of a Monte Carlo experiment and is followed, in Section 4, by a discussion of our findings. We conclude with some recommendations for applied researchers.

2. The Models and Tests

The analysis of this paper is based on the residuals from a regression model, rather than an observed series. Accordingly, we specify the basic model as

$$(1) \qquad y_t = x_t'\beta + u_t \; ; \qquad t=1,\ldots,n,$$

where y_t is a scalar, x_t is a k×l vector, β is a conformable parameter vector and u_t is a random disturbance.

We want to allow u_t to exhibit serial correlation both in the mean and in the conditional variance. To achieve this we can use the framework pioneered by Weiss (1984) which simply involves appending a GARCH[1] innovation term to a standard ARMA process. Restricting attention to AR processes, we can write this as

$$(2) \qquad u_t = \sum_{k=1}^{p} \rho_k u_{t-k} + \varepsilon_t$$

where $\varepsilon_t | \psi_{t-1} \sim N(0,h_t)$ and

$$(3) \qquad h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}.$$

The model described by (2) and (3) will be referred to as the Weiss model. Clearly, no restrictions are implied by this model on the order of the autoregressive component of the error term. Similarly, the GARCH process component of u_t is completely unrestricted.

An alternative synthesis of serial correlation and GARCH has been proposed by Bera, Higgins and Lee (1992) (henceforth BHL). In this model the conditional heteroscedasticity is generated by allowing the parameters of a standard autoregression to be random. The ARCH version of this model is written as (1) with the following specification for u_t:

$$u_t = \sum_{j=1}^{p} \phi_{jt} u_{t-j} + \varepsilon_t \tag{4}$$

$$\phi_{jt} = \phi_j + \eta_{jt} \tag{5}$$

where $\eta_t = (\eta_{1t}, \eta_{2t}, \ldots, \eta_{pt})$ and $\eta_t \sim N(0, \Sigma)$ for some positive definite matrix Σ. The covariances of current and previous errors in the conditional variance equation can affect the estimate of the variance, for any given period, through the off-diagonal elements of Σ. This feature allows leverage effects to enter risk prediction (*i.e.* the **sign** of the lagged u_t's affects the conditional variance in the manner suggested by Nelson (1991)), something which cannot occur in the Weiss model. If Σ is diagonal the linear ARCH model of Engle (1982) is obtained[2]. For the purpose of comparability with the Weiss model, as well as simplicity in experimental design, the study reported here uses only diagonal forms of Σ.

Applied researchers in the empirical finance literature typically use a range of formal tests to diagnose deficiencies in the specification of the conditional mean of regression errors. The portmanteau tests of Box and Pierce (BP) (1970) and Ljung and Box (LB) (1978) are heavily used, as are standard Lagrange multiplier (LM) tests for serial independence.

The LB and BP tests are each based on sums of squared sample autocorrelations, differing only by a scaling factor. For the BP test against AR(4) errors, the statistic

$$Q(\hat{r}) = n \sum_{k=1}^{4} \hat{r}_k^2$$

is treated as being asymptotically distributed as χ_4^2 under the null hypothesis that the first four autocorrelations are jointly zero. Here $\hat{r}_k$ is the k^{th} sample autocorrelation. Ljung and Box (1978) proposed a modification to the BP statistic which was intended to provide a closer approximation to a quantity related to sums of the **true** squared autocorrelations. They suggested that treating

$$Q_m(\hat{r}) = n(n+2) \sum_{k=1}^{p} (n-k)^{-1} \hat{r}_k^2$$

as χ_4^2 under the null hypothesis would provide a more powerful test. Both of these test statistics draw on the finding of Bartlett (1946) that the p^{th} sample autocorrelation of a white noise process is asymptotically normal with zero mean and variance of $(n-p)/(n^2+2n)$. This is not true of an ARCH process, however, in which the variance of the p^{th} sample autocorrelation is shown by Milhöj (1985) to be $(1/n)(1 + \gamma_p^2 / \sigma^4)$, where γ_p^2 is the p^{th} autocovariance for the squared process and σ^4 is the unconditional fourth moment.

Because $\gamma_p^2 / \sigma^4 > 0$, the approximation of the variance of $\hat{r}_k$ by $1/n$ (as is done by Box and Pierce in constructing $Q(\hat{r})$) will systematically underestimate $var(\hat{r}_k)$, even in large samples. Furthermore, the additional factor of $(n-k)/(n+2)$ which is taken into account by the LB statistic, $Q_m(\hat{r})$, reduces the value of the assumed variance still further. We should therefore expect that the size of the LB test is more severely affected by ARCH processes than the BP test.

Diebold (1986) suggested estimating γ_p^2 / σ^4 and using the estimates to construct adjusted versions of both the BP and LB tests, denoted BPA and LBA. The test statistics for these are respectively

$$Q^a(\hat{r}) = n\sum_{k=1}^{p}\left(\frac{\hat{\sigma}^4}{\hat{\sigma}^4 + \hat{\gamma}_k^2}\right)\hat{r}_k \, , \quad \text{and}$$

$$Q_m^a(\hat{r}) = n(n+2)\sum_{k=1}^{p}\left(\frac{\hat{\sigma}^4}{\hat{\sigma}^4 + \hat{\gamma}_k^2}\right)\hat{r}_k/(n-k) \quad .$$

Exact expressions for σ^4 and γ_k^2 are available for some conditional variance specifications (see Milhöj (1985) for example) but in practice these terms must be estimated and this is the method which was used in the Monte Carlo experiment reported in the next section.

When ARCH is present but the regression disturbances are serially uncorrelated, $Q^a(\hat{r}_k)$ and $Q_m^a(\hat{r}_k)$ are each asymptotically distributed as χ_p^2. This leaves several questions open. First, how are the rejection probabilities affected by this adjustment when the null hypothesis is not true? Second, what is the effect on the true size of the tests of allowing for ARCH processes in this way when no ARCH effect is present? Third, is the adjustment also valid for GARCH processes? The first two of these questions will be addressed in the empirical study described below; we turn our attention to the third question now.

Suppose that $\varepsilon_t = \eta_t\sqrt{h_t}$, where $\eta_t \sim N(0,1)$ and

$$h_t = \alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2 + \sum_{j=1}^{p}\beta_j h_{t-j}$$

$$= \alpha_0 + \alpha(L)\varepsilon_t^2 + \beta(L)h_t$$

where $\alpha(L) = \alpha_1 L + \ldots + \alpha_q L^q$, $\beta(L) = \beta_1 L + \ldots + \beta_p L^p$ and $L^j h_t = h_{t-j}$. Then

$$h_t = \frac{\alpha_0}{1-\beta(1)} + \frac{\alpha(L)}{1-\beta(L)}\varepsilon_t^2$$

$$(6) \qquad = \alpha_0^* + \sum_{l=1}^{\infty} \delta_l \varepsilon_{t-l}^2$$

where $\alpha_0^* = \alpha_0 / 1 - \sum_{j=1}^{p} \beta_j$ and δ_l is the coefficient of L^l in the expansion of $\alpha(L)/(1-\beta(L)$.

Equation (6) shows that a GARCH process is directly equivalent to an infinite order ARCH process. Thus Milhöj's (1985) representation of the variances of the sample autocorrelations for an ARCH process also applies to GARCH models. Recalling that Milhöj's expression was employed in the Diebold (1986) standard error correction, we conclude that this procedure is similarly valid for GARCH models.

To introduce the standard LM test and Wooldridge's robust (WLM) version, we consider the following AR(4) scheme for the u_t of (1):

$$(7) \qquad u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \phi_3 u_{t-3} + \phi_4 u_{t-4} + \varepsilon_t$$

where it is assumed that the eigenvalues of the associated determinental polynomial lie within the unit circle, so that the process is stationary.

Under the null hypothesis H_0: $\phi_1 = \phi_2 = \phi_3 = \phi_4 = 0$, and assuming that all other classical assumptions are satisfied, the Best Linear Unbiased Estimator is OLS, which is also the Maximum Likelihood Estimator. The nR^2 (or Outer Product Gradient) form of the LM test statistic for this problem is n times the uncentered coefficient of determination from a regression of the residuals, $\hat{u}_t$, from OLS estimation of (1) on the X matrix, and the first four lags of $\hat{u}_t$. Under the null, this statistic is asymptotically χ^2 with 4 degrees of freedom.

Wooldridge (1991), observing that this test is invalid in a dynamic model with conditional heteroscedasticity, proposed a general methodology for constructing tests which have sizes which are asymptotically robust in such cases. To focus on the practical application of Wooldridge's ideas we assume that the variables contained in X **do not**[3] include all lagged values of y_t.

Define $\lambda_t = (\hat{u}_{t-1},\ldots,\hat{u}_{t-4})$ and $h_t = E_t\left(Var(y_t|x_t)\right)$. In this notation the standard LM test discussed above uses nR^2 from the OLS estimation of:

$$\hat{u}_t = (x_t,\lambda_t)'\gamma + \eta_t$$

where γ is a suitably dimensioned parameter vector and η_t is a random error term. Wooldridge suggests initially weighting λ_t and x_t by dividing through by $\sqrt{h_t}$ where h_t is one's prior belief about the conditional variance function. In this study the weighting procedure is omitted because we wish to compare Wooldridge's LM test (WLM) with the standard procedure on the equivalent basis of ignorance about the presence (and therefore the form) of conditional heteroscedasticity. The Wooldridge procedure involves the following steps:

(i) Extract the 4×1 vector of residuals, r_t, from the vector regression of λ_t on x_t.

(ii) Define $\xi_t \equiv u_t r_t$ and extract the 4×1 vector of residuals, v_t, from the vector autoregression of ξ_t on $\xi_{t-1},\ldots,\xi_{t-G}$.

(iii) Treat nR^2 from the regression of ι on v_t as asymptotically χ^2_4 under the H_0 (where ι is a vector of ones).

The number of lags in the VAR of step (ii) is arbitrary and will clearly affect the power of the test. Wooldridge recommends the use of "one or two (times) the integer part of $\sqrt[4]{n}$ ". Throughout this study four lags were used in this vector autoregression.

Despite the relatively rare use of exact tests against AR alternatives in the empirical finance literature, it was decided to include two such tests in this study. The fourth-order analogue of the Durbin-Watson (1950) test (denoted DW_4) was derived by Wallis (1972). The test statistic is defined as

$$d_4 = \frac{u' A_4 u}{u' u}$$

where $A_4 = A_m \otimes I_4$, A_m is a tri-diagonal mxm matrix with all non-zero off diagonal entries being -1, one's in the north-west and south-east corners and two's for the remaining diagonal elements. The dimension of A_m is one quarter of the sample size and $\otimes$ denotes the Kronecker product. Exact critical values for the Wallis test can be easily computed, using the algorithm by Davies (1980), for example.

The only other exact test used in the study is a fourth-order generalisation of the s(0.75) test of King (1985). The test statistic for this test is given by

$$s_4(0.75) = \frac{u' Q u}{u' u}$$

where $Q = \Sigma - \Sigma X\left(X'\Sigma X\right)^{-1}X'\Sigma$, and Σ is the inverse of the theoretical covariance matrix of a simple AR(4) process assuming that $\rho_4 = 0.75$.

3. Experimental Design

To study the effect of conditional variance mis-specification on the size and power of the group of tests outlined above, a Monte Carlo study was conducted. All the work described below was conducted using 2000 replications which was found to produce reliable size figures for the exact tests used prior to the addition of conditional heteroscedasticity[4]. We used the SHAZAM (1993) package on a Vax 6340 computer. All psuedo random numbers were generated with a seed of 123.

Five design matrices were used, each of which contained an intercept and one other regressor. The non-constant regressors for the first four design matrices were based on the AR(1) process

(8) $x_t = \lambda x_{t-1} + \varepsilon_t$; $t = 1,\ldots,n$; $\varepsilon_t \sim N(0,1)$.

The matrices, denoted X1, X2, X3 and X4, were constructed using λ values of 0, 0.8, 1.0 and 1.02 respectively. These regressors are the same as those used by Engle, Hendry and Trumble (1985) and Lee and King (1993) and were incorporated in larger matrices by Giles, Giles and Wong (1993). The final data matrix (X5) contained the first two vectors from Watson's (1955) matrix[5] which was shown by Watson to produce the least efficient OLS estimates within the class of orthogonal matrices.

The study was conducted with a sample size of 60, using the following basic model for the conditional mean:

(9) $y_t = x_t'\beta + u_t$

(10) $u_t = \rho_4 u_{t-4} + \varepsilon_t$; $t=1,\ldots,n$; $\varepsilon_t \sim N(0,1)$.

The power of fourth order variants of the DW and s(0.75) tests were evaluated, along with those of the BP and LB tests, their Diebold (1986) adjusted versions (BPA and LBA) and the LM and WLM tests. In each case two nominal sizes were used, namely 1% and 5%. All of the asymptotic tests were conducted against the general AR(4) alternative of (7).

For each design matrix the power of each test was evaluated at ten values of ρ_4 in the range [0,0.9] thus establishing benchmark power functions in correctly specified models. Conditional heteroscedasticity was then introduced into the model using both the Weiss and BHL specifications introduced above. The conditional variance function

(11) $$h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

was used with α_0 set to $1-\alpha_1$ and the following parameter sets for (α_1,β_1)

ARCH Models: (0.2,0) (0.4,0) (0.6,0) (0.8,0);

GARCH Models: (0.2,0.2) (0.2,0.4) (0.2,0.6) (0.2,0.8).

These parameter sets allow for a range of GARCH models which include some important cases in which the unconditional fourth moment of the disturbances is not finite. For the ARCH(1) model $3\alpha_1^2$ must be less than unity for the existence of the unconditional fourth moment of u_t, a condition which is violated for $\alpha_1 \geq 0.577$. In the GARCH(1,1) case the existence condition is $3\alpha_1^2 + 2\alpha_1\beta_1 + \beta_1^2 < 1$ and the GARCH parameter sets used here include one pairing (0.2,0.8) which violates this condition.

In addition to this basic format the experiment was repeated with design matrix X4 using innovations drawn from a conditional Student-t distribution with 4 degrees of freedom. The purpose of this variation was to assess the dependence of the main results on the conditional normality which is employed elsewhere.

4. Experimental Results

Table 1 shows that the power function of the DW_4 test is reasonably robust to the addition of ARCH(1) innovations. This is particularly true of the endpoints of the curve; mid-range power is slightly reduced by ARCH. Similar effects were found for the $s_4(0.75)$ test across different data sets and for both the Weiss and BHL models.

Turning our attention to the BP and LB portmanteau tests, we can see from Figure 1a that the power curves of both the BP and LB tests are less steep in the presence of ARCH, but eventually converge to unity as ρ_4 increases. Figure 1b shows that these results also apply to AR-GARCH models and with other design matrices. The following general conclusion is supported by all of the cases considered in this study: the major effect of conditional variance mis-specification on the BP and LB tests is a significant increase in their true sizes. Power effects are negligible for very strong autocorrelation (*i.e.*, for $\rho_4 \geq 0.9$), but can be larger or smaller than the correctly specified power for moderate values of ρ_4.

In the light of the above finding, one might expect that a correction which ensures that the size of a BP or LB test is robust to GARCH (such as that of Diebold (1986)) would be a major advantage in the models studied here. Table 2 gives power values for the LB and LBA tests under ARCH innovations with the X2 matrix. This table shows that for any substantial degrees of ARCH ($\alpha_1 \geq 0.4$) the sizes of the adjusted tests are much closer to their nominal levels, relative to the standard (unadjusted tests). Also of interest in this table are the substantial power differences found between the Weiss and BHL models at moderate to large ρ_4 values (*ceteris paribus*), the causes of which are unknown.

Figure 2 graphs the powers of the BP and the LB tests with their adjusted versions under specific ARCH (Figure 2a) and GARCH (Figure 2b) models. These graphs clearly show that the cost (in power terms) of obtaining a reasonably robust size by using the Diebold adjustment can be very high in all but the extreme regions of the (ρ_4) parameter space. This conclusion is reached regardless of the data or the choice of Weiss or BHL models. Furthermore, it is clearly evident from Figure 2 that all four versions of the portmanteau test are markedly inferior to the standard fourth order DW test[6]. This reinforces the view of Geweke (1988) that "the properties of "Q" are terrible in almost all econometric work", and suggests that efforts to devise powerful exact tests against general AR(p) alternatives could be of major benefit to applied finance researchers.

We now consider the standard LM test and its (size) robust counterpart, the WLM test. Figure 3 shows the power function of each of these tests in four different models. First, in Figure 3a, ARCH innovations are used. In this case the size of the LM test is increased by ARCH, while the main effect on the WLM test is a reduction in power for moderately large ρ_4. More striking, however, is the very substantial difference between the power curves of the LM and WLM tests for a given model specification. As noted above, and confirmed in this graph, the WLM test has a true size which is somewhat lower than its nominal level. This does not adequately account for the very modest slope of the test's power curve, however. A very similar story is told by Figure 3b with respect to the inter-test comparison. In this case, however, the effect on the LM test of stronger conditional heteroscedasticity is reversed, with larger β_1 values tending to **increase** the size of the test. For all design matrices used, the direction of size distortion of the LM test was found to be upwards in ARCH models and downwards in GARCH model.

Subject to the caveat that we have only considered models with two regressors, the results appear to be relatively independent of the data. The broad findings remain valid for all of the X matrices used[7].

To assess the dependence of the above findings on the assumption of conditional normality in the error term a limited investigation was conducted using the X1 and X4 matrices with the Weiss model. These data can be thought of as bounding design matrices described by (8). For this section of the study the conditional distribution of the ε_t of (10) and (11) was assumed to be Student-t with 4 degrees of freedom.

The conclusions of the main study with respect to the exact AR(4) tests remain valid with conditionally t_4 errors. An example of this is shown in Figure 4a where the DW test can be seen to maintain its assigned significance level and suffer only very minor losses in power in the presence of strong ARCH. We also found that the Diebold size adjustment is more successful when the underlying distribution has heavier tails, but that the power curve is less steeply sloped in this case.

The LM test, which has severe size problems even without ARCH, suffers very badly from the relaxation of the conditional normality assumption. The power functions of the LM test for different degrees of ARCH, however, are always steeper than those of the WLM test (Figure 4b).

In summary, the relaxation of the assumption of conditional normality exacerbates the size problems of the BP, LB and LM tests but does not change their qualities relative to the proposed "robust" versions. These adjusted tests, while generally achieving their aim of lowering true size, remain markedly inferior by the criterion of power function slope. The exact tests stand out as being the ideal choice under the distributional assumption adopted here.

8.6 Conclusion

In this paper we have substantially clarified several issues related to the specification of the conditional mean of a regression model when the errors are conditionally heteroscedastic. It has been shown that the well known exact tests for simple autoregressive processes are outstandingly robust to the presence of GARCH effects. This conclusion is, of course, subject to the usual assumption that the alternative model is indeed a simple AR process of the appropriate order. We have also shown that the very frequently used BP and LB tests have sizes which are substantially greater than their nominal levels when conditional heteroscedasticity is present but that despite this increased size they are still less powerful than the exact tests for virtually all degrees of autocorrelation. The LM test for AR(4) errors grossly over rejects the null model even without GARCH, which makes the problem worse.

The two existing methods for correcting size distortion in the BP, LB and LM tests, due to Diebold (1986) and Wooldridge (1991) are generally successful in their stated aims, although Wooldridge's procedure tends to over correct in the moderate sample size used here. The power curves of these "robust" tests are much less steep than those of the standard tests, however, which raises serious doubts about the advisability of their use.

Finally, these conclusions do not depend on the assumption of conditional normality, having been also found using the thicker tailed student t distribution with 4 degrees of freedom. It is, of course, possible that a skewed distribution may alter some of the conclusions.

REFERENCES

Baillie, R.T., and T Bollerslev (1989), The Message in Daily Exchange Rates: A Conditional Variance Tale, *Journal of Business and Economic Statistics*, 7, 297-305.

Bartlett, M.S., (1946), On the theoretical specification of sampling properties of autocorrelated time series, *Journal of the Royal Statistical Society, Series B*, 8, 27-41.

Bera, A.K., M.L. Higgins and S.Lee, (1992), Interaction between autocorrelation and conditional heteroscedasticity: a random coefficient approach, *Journal of Business and Economic Statistics*, 10, 133-142.

Bera, A.K., and M.L. Higgins, (1993), A survey of ARCH models: properties, estimation and testing, forthcoming in *Journal of Economic Surveys.*

Bollerslev, T., (1986), Generalised autoregressive conditional heteroscedasticity, *Journal of Econometrics*, 31, 307-327.

Box, G.E.P. and D.A. Pierce (1970), Distribution of the residual autocorrelations in ARIMA time series models, *Journal of the American Statistical Association*, 65, 1509-1526.

Chou, R.Y. (1988), Volatility Persistence and Stock Valuations: Some Empirical Evidence Using GARCH, *Journal of Applied Econometrics*, 3, 279-294.

Davies, R.B., (1980), The distribution of a linear combination of chi square random variables, *Applied Statistics*, 29, 323-333.

Diebold, F.X., (1986), Testing for serial correlation in the presence of ARCH, *Proceedings from the American Statistical Association, Business and Economic Statistics Section*, 323-328.

Diebold, F.X, and M, Nerlove, (1989), The dynamics of exchange rate volatility: a multivariate latent factor ARCH model, *Journal of Applied Econometrics*, 4, 1-21.

Durbin, J. and G.S. Watson, (1950), Testing for serial correlation in least squares regression I, *Biometrika*, 37, 409-428.

Engle, R.F.,(1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, 50, 987-1008.

Engle, R.F., D.F. Hendry and D. Trumble, (1985), Small sample properties of ARCH estimators and tests, *Canadian Journal of Economics*, 18, 66-93.

Geweke, J., (1988), Diagnostics for the diagnostics, *Econometric Reviews*, 7, 59-62.

Giles, D.E.A., J.A. Giles and J.K. Wong, (1993), Testing for ARCH-GARCH errors in a mis-specified regression, *Computational Statistics*, 8, 109-126.

Hsieh, D.A., (1989), Modelling Heteroscedasticity in Daily Foreign Exchange Rates, *Journal of Business and Economic Statistics*, 7, 307-317.

King, M.L., (1985), A point optimal test for autoregressive disturbances, *Journal of Econometrics*, 27, 21-37.

Lee, J.H.H., and M.L. King, (1993), A locally most mean powerful based score test for ARCH and GARCH regression disturbances, *Journal of Business and Economic Statistics*, 11, 17-27.

Ljung, G.M. and G.E.P. Box, (1978), On a measure of lack of fit in time series models, *Biometrika*, 65, 297-303.

Milhöj, A., (1985), The moment structure of ARCH processes, *Scandanavian Journal of Statistics*, 12, 281-292.

Nelson, D.B., (1991), Conditional heteroscedasticity in asset return: a new approach, *Econometrica*, 59, 347-370.

SHAZAM Econometrics Computer Program (Version 7.0), Users Reference Manual, (1993), McGraw-Hill, New York.

Silvapulle, P., and M.A. Evans (1993), Testing for serial correlation in the presence of conditional heteroscedasticity, Mimeo, La Trobe University, Melbourne.

Small, J.P., (1993), The exact powers of some autocorrelation tests when the disturbances are heteroscedastic, forthcoming in *Journal of Econometrics*.

Wallis, K.F., (1972), Testing for fourth order autocorrelation in quarterly regression equations, *Econometrica*, 40, 617-636.

Watson, G.S., (1955), Serial correlation in regression Analysis I, *Biometrika*, 42, 327-341.

Weiss, A.A., (1984), ARMA models with ARCH errors, *Journal of Time Series Analysis*, 5, 129-143.

Wooldridge, J.M., (1991), On the application of robust regression based diagnostics to models of conditional means and conditional variances, *Journal of Econometrics*, 47, 5-46.

Footnotes

[1] In fact Weiss (1984) used ARCH processes only. The extension to GARCH is trivial.

[2] A GARCH model can also be given a random coefficient interpretation (Bera and Higgins (1993)).

[3] The majority of Wooldridge's paper is based on models with completely specified dynamics, in which x_t contains the entire past history of y_t.

[4] This was verified by using the standard exact techniques for evaluating rejection probabilities for these tests as in Small (1993), for example.

[5] The columns of this matrix are given by a_1, $(a_2+a_T)/\sqrt{2}$, where $a_1,\ldots,a_T$, are the eigenvectors corresponding to the eigenvalues of A, arranged in increasing order, and A is a tri-diagonal matrix with all off diagonal elements being -1, the first and last leading diagonal elements being 1's and all other leading diagonal elements being 2's. Note that a_1 is a constant.

[6] The $s_4(0.75)$ test performs just as well as the DW_4 test in these models.

[7] Further detailed evidence is available from the author.

TABLE 1
Power of DW_4 Test with ARCH Errors
Data Matrix X1

	Weiss Model		BHL Model	
ρ4	1% Size	5% Size	1% Size	5% Size
α1=0.0				
0.0	0.006	0.049	0.011	0.050
0.3	0.416	0.677	0.413	0.676
0.5	0.889	0.967	0.895	0.972
0.7	0.994	0.999	0.999	1.000
0.9	1.000	1.000	1.000	1.000
α1=0.2				
0.0	0.006	0.054	0.011	0.047
0.3	0.405	0.669	0.388	0.645
0.5	0.895	0.966	0.883	0.961
0.7	0.995	0.999	0.995	0.998
0.9	1.000	1.000	1.000	1.000
α1=0.4				
0.0	0.008	0.052	0.010	0.047
0.3	0.394	0.661	0.371	0.631
0.5	0.891	0.965	0.861	0.951
0.7	0.994	1.000	0.992	0.999
0.9	1.000	1.000	1.000	1.000
α1=0.6				
0.0	0.010	0.053	0.013	0.047
0.3	0.385	0.659	0.332	0.604
0.5	0.890	0.963	0.824	0.934
0.7	0.993	1.000	0.985	0.997
0.9	1.000	1.000	1.000	1.000
α1=0.8				
0.0	0.015	0.053	0.018	0.057
0.3	0.378	0.648	0.321	0.563
0.5	0.883	0.961	0.781	0.922
0.7	0.992	0.999	0.973	0.993
0.9	1.000	1.000	0.995	0.999

TABLE 2
Power of LB and LBA Tests with ARCH Errors
Data Matrix X2

	Weiss Model		BHL Model	
ρ4	LB	LBA	LB	LBA
α1=0.0				
0.0	0.060	0.069	0.059	0.070
0.3	0.398	0.215	0.401	0.221
0.5	0.831	0.558	0.834	0.564
0.7	0.982	0.854	0.987	0.861
0.9	1.000	0.959	1.000	0.964
α1=0.2				
0.0	0.079	0.070	0.077	0.079
0.3	0.410	0.220	0.391	0.225
0.5	0.837	0.582	0.827	0.554
0.7	0.986	0.874	0.986	0.884
0.9	0.999	0.967	1.000	0.983
α1=0.4				
0.0	0.108	0.073	0.110	0.076
0.3	0.429	0.223	0.409	0.230
0.5	0.848	0.593	0.837	0.546
0.7	0.987	0.899	0.978	0.871
0.9	0.999	0.971	0.997	0.974
α1=0.6				
0.0	0.145	0.078	0.156	0.084
0.3	0.462	0.232	0.442	0.223
0.5	0.859	0.597	0.818	0.522
0.7	0.986	0.905	0.968	0.832
0.9	0.999	0.978	0.987	0.925
α1=0.8				
0.0	0.197	0.083	0.211	0.102
0.3	0.511	0.237	0.463	0.228
0.5	0.868	0.602	0.805	0.492
0.7	0.983	0.906	0.945	0.760
0.9	0.998	0.978	0.956	0.826

TABLE 3
Power of BP and BPA Tests under GARCH
Data Matrix X4; Weiss Model

	Normal Errors		Student t Errors	
$\rho 4$	BP(5)[1]	BPA(5)	BP(5)	BPA(5)
$\beta 1=0.2$				
0.0	0.071	0.062	0.120	0.060
0.3	0.360	0.188	0.411	0.188
0.5	0.801	0.520	0.816	0.521
0.7	0.979	0.870	0.979	0.876
0.9	0.999	0.979	1.000	0.976
$\beta 1=0.4$				
0.0	0.078	0.050	0.152	0.048
0.3	0.370	0.169	0.427	0.196
0.5	0.796	0.502	0.814	0.513
0.7	0.979	0.874	0.976	0.866
0.9	0.999	0.982	0.999	0.977
$\beta 1=0.6$				
0.0	0.097	0.032	0.176	0.041
0.3	0.378	0.141	0.434	0.172
0.5	0.798	0.468	0.787	0.469
0.7	0.974	0.871	0.960	0.842
0.9	0.999	0.984	0.995	0.969
$\beta 1=0.8$				
0.0	0.139	0.017	0.255	0.055
0.3	0.382	0.099	0.444	0.140
0.5	0.760	0.393	0.706	0.354
0.7	0.954	0.792	0.891	0.648
0.9	0.995	0.964	0.960	0.863

1. BP(5) refers to the BP test with a 5% nominal size. The other columns are similarly designated.

Figure 1a

Power of BP and LB Tests with AR(4)-ARCH(1) Errors
Data Matrix X2; Weiss Model

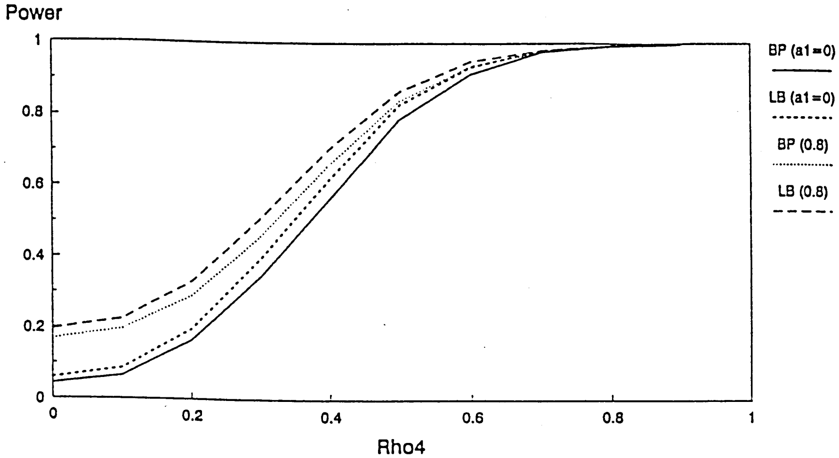


Figure 1b

Power of BP and LB Tests with AR(4)-GARCH(1,1) Errors
Data Matrix X4; BHL Model

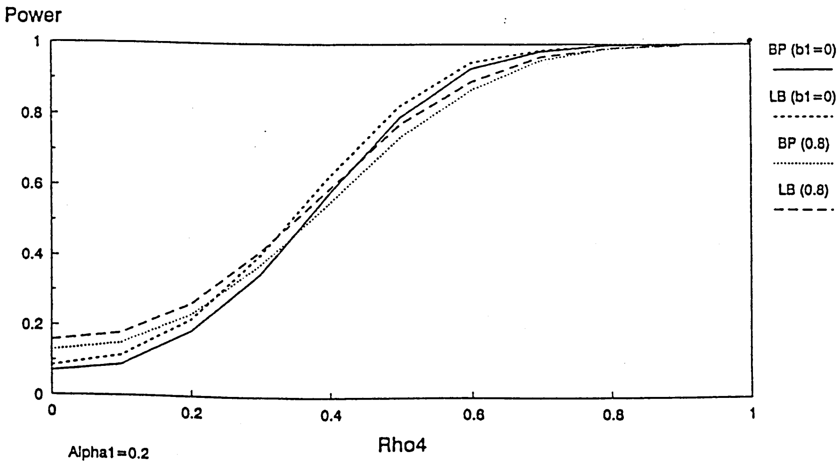


Figure 2a

Power of Several Tests with AR(4)-ARCH(1) Errors

Data Matrix X1; Weiss Model

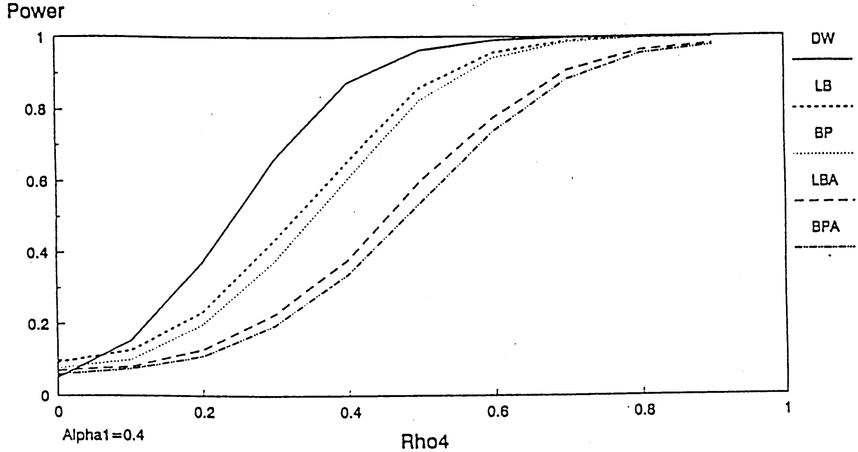


Figure 2b

Power of Several Tests with AR(4)-GARCH(1,1) Errors

Data Matrix X1; BHL Model

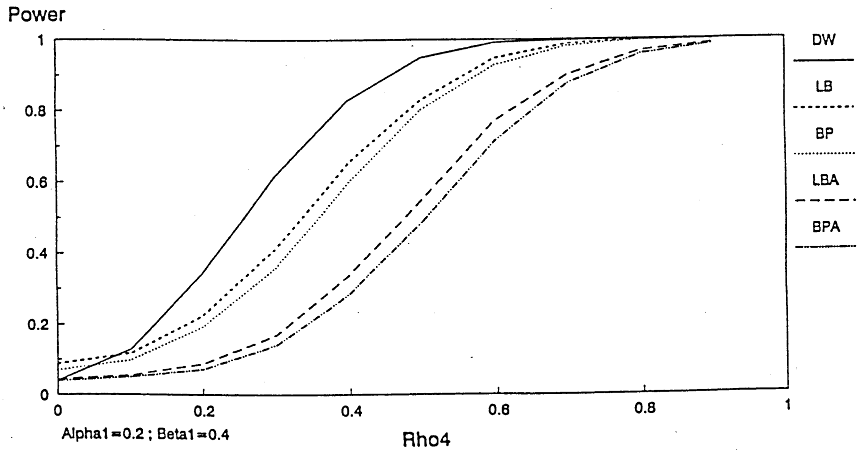
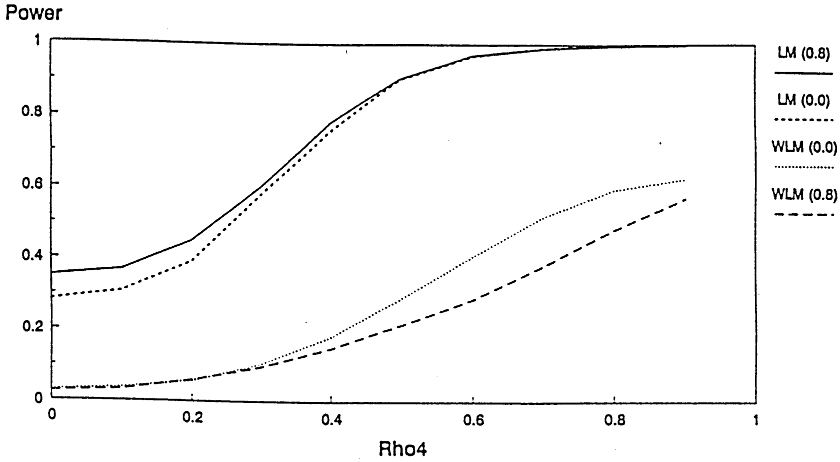


Figure 3a

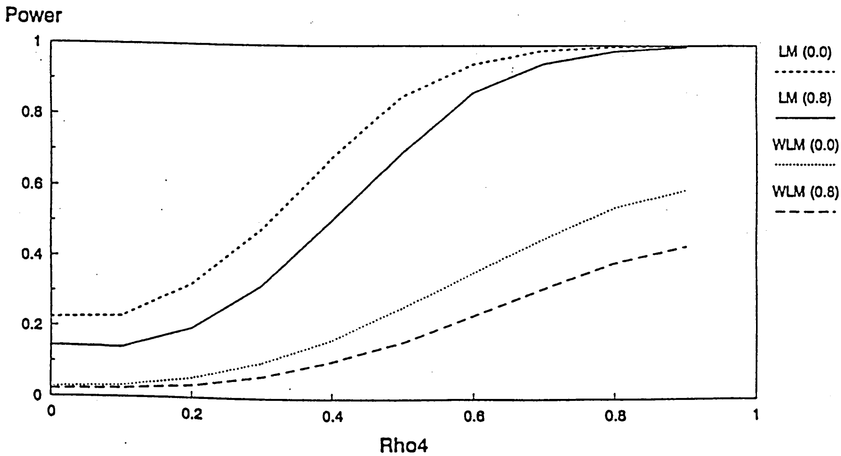
Power of LM and WLM Tests with AR(4)-ARCH(1) Errors
Data Matrix X3; Weiss Model



Bracketted Figure in LM (0.0) Is ARCH Parameter

Figure 3b

Power of LM and WLM Tests with AR(4)-GARCH(1,1) Errors
Data Matrix X5; BHL Model



Bracketted Figure In LM (0.0) Is GARCH Parameter

Figure 4a

Power of DW Test with Student t Errors and ARCH
Data Matrix X1; Weiss Model

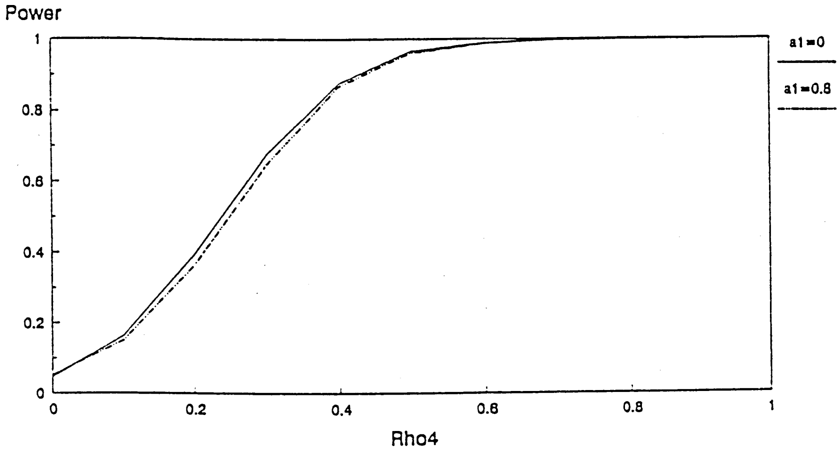
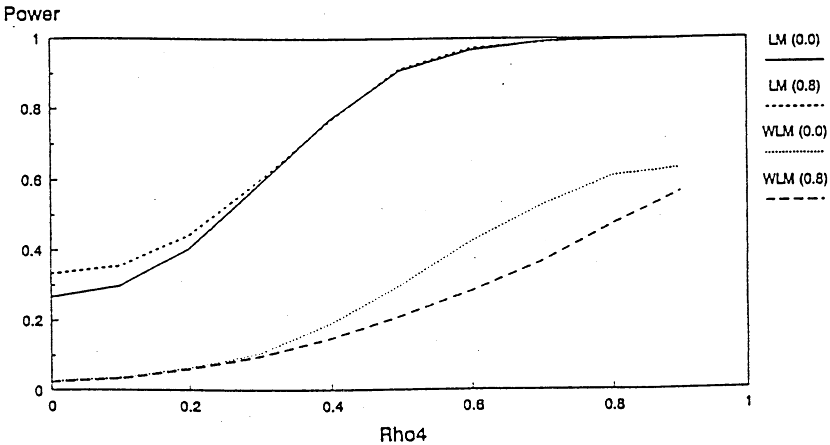


Figure 4b

Power of LM & WLM Tests with Student t Errors and ARCH
Data Matrix X1; Weiss Model



Brackettad Figure in LM (0.0) is ARCH Parameter

LIST OF DISCUSSION PAPERS*

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

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

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

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

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

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

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

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

No. 8909 Pre-testing for Linear Restrictions in a Regression Model with Spherically Symmetric Disturbances, by Judith A. Giles.

No. 9001 The Durbin-Watson Test for Autocorrelation in Nonlinear Models, by Kenneth J. White.

No. 9002 Determinants of Aggregate Demand for Cigarettes in New Zealand, by Robin Harrison and Jane Chetwyd.

No. 9003 Unemployment Duration and the Measurement of Unemployment, by Manimay Sengupta.

No. 9004 Estimation of the Error Variance After a Preliminary-Test of Homogeneity in a Regression Model with Spherically Symmetric Disturbances, by Judith A. Giles.

No. 9005 An Expository Note on the Composite Commodity Theorem, by Michael Carter.

No. 9006 The Optimal Size of a Preliminary Test of Linear Restrictions in a Mis-specified Regression Model, by David E. A. Giles, Offer Lieberman, and Judith A. Giles.

No. 9007 Inflation, Unemployment and Macroeconomic Policy in New Zealand: A Public Choice Analysis, by David J. Smyth and Alan E. Woodfield.

No. 9008 Inflation — Unemployment Choices in New Zealand and the Median Voter Theorem, by David J. Smyth and Alan E. Woodfield.

No. 9009 The Power of the Durbin-Watson Test when the Errors are Heteroscedastic, by David E. A. Giles and John P. Small.

No. 9010 The Exact Distribution of a Least Squares Regression Coefficient Estimator After a Preliminary t-Test, by David E. A. Giles and Virendra K. Srivastava.

No. 9011 Testing Linear Restrictions on Coefficients in a Linear Regression Model with Proxy variables and Spherically Symmetric Disturbances, by Kazuhiro Ohtani and Judith A. Giles.

No. 9012 Some Consequences of Applying the Goldfeld-Quandt Test to Mis-Specified Regression Models, by David E. A. Giles and Guy N. Saxton.

No. 9013 Pre-testing in a Mis-specified Regression Model, by Judith A. Giles.

No. 9014 Two Results in Balanced-Growth Educational Policy, by Alan E. Woodfield.

No. 9101 Bounds on the Effect of Heteroscedasticity on the Chow Test for Structural Change, by David Giles and Offer Lieberman.

No. 9102 The Optimal Size of a Preliminary Test for Linear Restrictions when Estimating the Regression Scale Parameter, by Judith A. Giles and Offer Lieberman.

No. 9103 Some Properties of the Durbin-Watson Test After a Preliminary t-Test, by David Giles and Offer Lieberman.

No. 9104 Preliminary-Test Estimation of the Regression Scale Parameter when the Loss Function is Asymmetric, by Judith A. Giles and David E. A. Giles.

No. 9105 On an Index of Poverty, by Manimay Sengupta and Prasanta K. Pattanaik.

No. 9106 Cartels May Be Good For You, by Michael Carter and Julian Wright.

No. 9107 Lp-Norm Consistencies of Nonparametric Estimates of Regression, Heteroskedasticity and Variance of Regression Estimate when Distribution of Regression is Known, by Radhey S. Singh.

(Continued on next page)

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

No. 9305 The Exact Powers of Some Autocorrelation Tests When Relevant Regressors are Omitted, by J. P. Small, D. E. Giles and K. J. White.

No. 9306 The Exact Risks of Some Pre-Test and Stein-Type Regression Estimators Under Balanced Loss*, by J. A. Giles, D. E. A. Giles, and K. Ohtani.

No. 9307 The Risk Behavior of a Pre-Test Estimator in a Linear Regression Model with Possible Heteroscedasticity under the Linex Loss Function, by K. Ohtani, D. E. A. Giles and J. A. Giles.

No. 9308 Comparing Standard and Robust Serial Correlation Tests in the Presence of Garch Errors, by John P. Small.

* Copies of these Discussion Papers may be obtained for $4 (including postage, price changes occasionally) each by writing to the Secretary, Department of Economics, University of Canterbury, Christchurch, New Zealand. A list of the Discussion Papers prior to 1989 is available on request.