MONASH

W.P. no. 10/89

MONASH

UNIVERSITY

AUSTRALIA

ROBUSTNESS AND SIZE OF TESTS OF AUTOCORRELATION

AND HETEROSCEDASTICITY TO NON-NORMALITY

Merran Evans

Working Paper No. 10/89

October 1989

DEPARTMENT OF ECONOMETRICS

M O N A S H
U N I V E R S I T Y

ANCORA · IMPARO

AUSTRALIA

ROBUSTNESS AND SIZE OF TESTS OF AUTOCORRELATION AND HETEROSCEDASTICITY TO NON-NORMALITY

Merran Evans

Working Paper No. 10/89

October 1989

DEPARTMENT OF ECONOMETRICS

ISSN 1032-3813

ISBN 0 86746 968 4

ROBUSTNESS AND SIZE OF TESTS OF AUTOCORRELATION AND HETEROSCEDASTICITY TO NON-NORMALITY

Merran Evans

Working Paper No. 10/89

October 1989

DEPARTMENT OF ECONOMETRICS, FACULTY OF ECONOMICS AND POLITICS

MONASH UNIVERSITY, CLAYTON, VICTORIA 3168, AUSTRALIA.

ROBUSTNESS OF SIZE OF TESTS OF AUTOCORRELATION AND HETEROSCEDASTICITY TO NON-NORMALITY*

Keywords: Autocorrelation, heteroscedasticity, robustness, normality.

A comprehensive empirical examination is made of the sensitivity of tests of disturbance covariance in the linear regression model to non-normal disturbance behaviour. Tests of autocorrelation appear to be quite robust, except for extreme non-normality, but tests for heteroscedasticity are highly susceptible to kurtosis.

Author: (Dr) Merran Evans
Dept. of Econometrics, Monash University, Clayton, 3168, Australia.

*The research for this paper was supported by the Monash University Special Research Fund under grant ECP2/87.

1. Introduction

Parametric tests of specification of the linear regression model generally assume normally distributed disturbances. A growing number of econometricians are questioning this assumption, on which the validity of standard hypothesis tests and confidence intervals is based, and are concerned with the effect of non-normality, particularly in small samples which are characteristic of econometric analysis. Tests on regression coefficients appear to be reasonably robust to non-normal disturbances, but whether this is so for tests of disturbance behaviour is less clear. Studies, so far, are not comprehensive and their results are sometimes in conflict.

This study attempts to evaluate thoroughly the effect of assuming normality on the size of parametric tests of the disturbance covariance matrix, by including a wider range of tests, alternative distributions and regressors than in previous empirical studies. A Monte Carlo comparison of actual and nominal sizes of various test statistics for first- and fourth-order autocorrelation and heteroscedasticity is made. Particular emphasis is on more recent tests, which are based on small sample properties, rather than on an asymptotic justification.

2. Background

Consider the linear regression model

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

where y is $n \times 1$, X is $n \times k$ and non-stochastic, β is $k \times 1$, and $u \sim (0, \sigma^2 I_n)$ is $n \times 1$. $\hat{\beta} = (X'X)^{-1}X'y$ is the Ordinary Least Squares (OLS) estimate of β, $\hat{u} = (I - X(X'X)^{-1}X')y = My$ the OLS residual vector, and $s^2 = \hat{u}'\hat{u}/(n-k)$, the estimate of the disturbance variance σ^2. $\hat{\beta}$ and s^2 are unbiased and consistent and $\hat{\beta}$ is the best linear unbiased estimator. For normally distributed disturbances: $\hat{\beta}$ is the maximum likelihood estimator, and hence efficient; s^2 is asymptotically efficient; $\hat{\beta}$ has a normal, and $(n-k)s^2/\sigma^2$ a χ^2, distribution; and classical methods of statistical inference are valid.

Evidence is accumulating which suggests that the assumption of normally distributed disturbances is often inconsistent with the behaviour of many economic variables. The generation process may follow laws other than normal, the disturbances may be a mixture of distributions, or the occurence of a few outliers also can produce 'fat' tails. Non-symmetric distributions can occur in the modelling of frontier production functions and residual security returns. 'Leptokurtic', 'fat-' or 'long-tailed' distributions, characterised by a high kurtosis measure, are relevant to the study of financial and commodity markets. Symmetric stable laws, which have no finite moments, have been used to describe stock market prices, although mixtures of distributions have been preferred more recently.

With non-normal disturbances, the size of the t and F tests of regression coefficients appears to be reasonable (Box and Watson (1962), Pearson and Please (1975)), though some adjusments for kurtosis may be warranted, and these tests are asymptotically valid. These tests have the correct size for spherically symmetric disturbance distributions (Kariya and Eaton (1977), Kariya (1977) and King (1979)), and for normal regressors (Cavanagh and Rothenberg(1985)).

The t and F tests are invalid if serially correlated or heteroscedastic disturbances are ignored. Optimal power properties of tests for such disturbance behaviour, which hold under normality, also hold for any spherically symmetric disturbances (King(1979)), but the only spherically symmetric distribution with independent components is the normal distribution. Empirical examination of tests of correlated disturbances under non-normality include the very limited study of the Durbin Watson (DW) and Geary runs test by Gatswirth and Selwyn (1980), who found the DW test reasonably robust. Bartels and Goodhew (1981) used extreme error distributions, with three sets of "smoothly evolving" regressors and considered positive autocorrelation. They found the DW test size reasonably robust at the 5% significance level, but at lower levels it had higher rejection rates, and argued that this potential overcorrection for autocorrelation was not a problem. Knight (1985), by applying the Davis technique to non-normal distributions which can be characterised by an Edgeworth expansion, found the DW statistic very robust. However, using numerical methods for disturbances formed from mixtures of normal distributions, he found its sensitivity depended on the regressors used. Smith (1987) (extending Hillier and Smith (1983)), also using the Davis technique and moderate non-normality, found that disturbances with zero skewness and extreme kurtosis had the greatest effect on the mean, variance and kurtosis of the DW statistic, but that these were relatively robust, whereas with extreme skewness and moderate kurtosis, the skewness measure was not robust. For *some* X matrices, the test sizes were markedly affected, leading to fewer (more) rejections of the null hypothesis in favour of positive (negative) correlation than desirable. This result is in conflict with Bartels and Goodhew. In contrast to each, McCabe (1989), using generalisations of the Tukey-lambda distributions, found the von-Neumann ratio extraordinarily robust to both skewness and kurtosis. Bera and Jarque (1981) also found the DW and Lagrange Multiplier (LM) tests very robust.

For large samples, Box (1953) showed that the Type I error of Bartlett's test, a modification of the Likelihood Ratio (LR) test of equality of variances in several independent normal populations, depends heavily on kurtosis. Small sample examination of the robustness of some traditional tests for equality of variances, based on Bartlett's test and modifications of the F test, have been made by Gartside (1972), Brown and Forsythe (1974), Conover, Johnson and Johnson (1981) and Rivest (1986) in the context of the analysis of variance model. Examining tests of heteroscedasticity in the linear regression model, Barone-Adesi and Talwar (1983) found those of Bartlett, Glesjer, Kendall and Goldfeld and Quandt sensitive to the χ_4^2, t_4 and double exponential distributions, but Johnston's rank and Bickel's tests reasonably robust. Ali and Giacotto (1984) examined the robustness of non-parametric rank tests, Bickel's robust tests, and the parametric tests of Goldfeld and Quandt (GQF), Ramsey, Glesjer, White, and Breusch and Pagan, as well as the LR test. For disturbances from the skewed lognormal distribution, most of the non-parametric tests and the Glesjer, LR and Breusch and Pagan tests were either robust or moderately so, but the Ramsey and GQF tests were not. For long-tailed distributions, the non-parametric tests were reasonable with t_3, but not with the Cauchy distribution, and the Glesjer, White and Bickel tests were at best moderately robust with the Cauchy distribution. They considered the Bartlett and GQF tests sensitive to both skewness and kurtosis. In contrast, again, McCabe (1989) found the GQF test to be 'quite good' for the skewed case but not robust to light or heavy tails.

Several gaps appear in this literature. In particular, very little is known of the robustness of recent parametric tests of autocorrelation and heteroscedasticity, which are based on small sample properties. Further, some conclusions of past studies are in conflict. The following is an attempt to rectify these problems.

3. Outline of the Empirical Evaluation

In order to evaluate the accuracy of the normal critical values under alternative error distributions, empirical sizes were computed in a Monte Carlo simulation, and compared with the nominal size for a variety of tests and X matrices. These were computed in the context of the general linear regression model for testing against first-order autoregressive (AR(1)) and simple fourth-order autoregressive (AR(4)) disturbances, and additive heteroscedasticity of the form

$$var(u_t) = \sigma_t^2 = \sigma^2 f(1 + \lambda z_t), \tag{1}$$

where f is an unknown monotonically increasing non-negative function and z_t is a non-stochastic variable. The methodology reflects that of Evans and King (1985a), in which the accuracy of standard approximations to the critical values of a variety of tests for such disturbances was evaluated.

The actual and artificial X matrices considered are designed to reflect a range of behaviour characteristic of economic variables, as well as extreme behaviour to highlight any differences which may occur. All have been used previously in experiments concerning tests of autocorrelation and heteroscedasticity (see King (1985), Evans and King (1985a,1985b,1988), Griffiths and Surekha (1986)), and most in the empirical studies just discussed which examine the robustness to non-normality of these tests. The three real design matrices contain a constant term and two other regressors which are reasonably typical economic time series data, with various degrees of trend and seasonality. These regressors were: the annual spirit income and price data of Durbin and Watson (1951), as used by Smith (1987) and Knight (1985); the weakly seasonal quarterly Australian Consumer Price Index (CPI), and also lagged one quarter (as in Smith); and the quarterly

Australian capital movements, private and government, which are highly seasonal and subject to large fluctuations. Four artificial regressor matrices, each comprising a constant and a regressor determined from a time trend and the normal, lognormal and uniform distributions respectively, were used. These artificial data sets represent a range of alternative behaviour: a time trend is used often to characterise slowly evolving non-seasonal economic time series and was used by Bartels and Goodhew (1981) and Barone-Adesi and Talwar (1983); the lognormal distribution is employed in heteroscedasticity experiments to represent skewed data characteristic of cross-sectional data (see Griffiths and Surekha (1986) and McCabe (1989)); and the results of Cavanagh and Rothenberg(1985) suggest the inclusion of normal regressors, which are also used by Smith. Uniform data has been used by Ali and Giacotto (1984). Small ($n = 24$) and relatively large ($n = 64$) sample sizes were examined for each set of regressors.

Some invariance results simplify the analysis and extend the applicability of the conclusions reached. All tests considered depend on least squares residuals, which depend on the true (here known) disturbances and the X matrices, and not on the y or β vector. Hence there is no need to specify parameter values for the simulations. Further, as a constant term was included in each regressor matrix, the least squares residuals summed to zero, such that the tests were invariant to the mean of the error distribution. Invariance also applied to the disturbance variance, σ^2, as most of the test statistics considered can be expressed as a ratio of quadratic forms in disturbances. The others involved ratios of squared residuals. It is reassuring to know that these tests are invariant to transformations, considered by Durbin and Watson (1971), of the form $y \rightarrow \gamma_0 y + X\gamma$, (where γ_0 is a scalar, and γ a $k \times 1$ vector). As a consequence, they are invariant to a change in scale of the data and to multicollinearity (Evans (1985)).

The tests against AR(1) disturbances considered were the first-order Durbin and Watson (1950) test (DW1), its Locally Best Invariant (LBI) alternative (DW1alt) of King (1981), the Berenblut and Webb (1973) test (BW1), and King's (1985) point optimal test (s1(.5)). For AR(4) disturbances the fourth order analogues of these first-order tests considered were the Wallis (1972) (DW4) test, the alternative LBI test (DW4alt), Webb's (1973) test (BW4) and King's (1984) s4(.5) test. The tests of heteroscedasticity included Szroeter's (1978) bounds test (SZ), and the Evans and King (1985b,1988) point optimal s(5.0) and approximate sa(2.5) tests, the one-sided LM test (LM1), which is LBI, and its approximate version (sa). These tests reject $H_0 : u \sim N(0, \sigma^2 I_n)$ for small values of the test statistic in the case of heteroscedasticity and for positive autocorrelation. Against negative autocorrelation, H_0 is rejected for large values, except for the point optimal tests, for which the analogous s1(-.5) and s4(-.5) tests are used.

These tests can be classified into two classes, and can be expressed as a ratio of quadratic forms in residuals and in disturbances. Tests which are based on OLS residuals include DW1, DW1alt, DW4, DW4alt, LM1, SZ and sa and can be written in the form

$$r = \hat{u}'A\hat{u}/\hat{u}'\hat{u} = u'MAMu/u'Mu = u'Bu/u'Mu, \tag{2}$$

where A is a real symmetric $n \times n$ matrix, $M = I - X(X'X)^{-1}X'$ and $B = MAM$. Tests also based on Generalised Least Squares (GLS) residuals include s1(.5), BW1, s4(.5), BW4, s(5.0) and sa(2.5), which can be written in the form

$$r = \tilde{u}'\Sigma^{-1}\tilde{u}/\hat{u}'\hat{u} = u'Bu/u'Mu, \tag{3}$$

where Σ is a positive definite $n \times n$ matrix, $\tilde{u}$ is the $n \times 1$ vector of GLS residuals assuming covariance matrix Σ, and here $B = \Sigma^{-1} - \Sigma^{-1}X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}$. From Durbin and

Watson's (1950) lemma, under normality, the α significance level critical value can be obtained by solving for $r*$ in

$$Pr(r < r^*) = Pr[u'(B - r^* I)u < 0] = \alpha.$$

The Breusch and Pagan (1979) LM test (BPtrue) was also used to test heteroscedasticity of the form (1). Here the true α level significance points are obtained by solving

$$Pr(BPtrue > r*) = Pr(\hat{u}'D\hat{u}/\hat{u}'\hat{u}) > r*^{1/2}) + Pr(\hat{u}'D\hat{u}/\hat{u}'\hat{u} < -r*^{1/2}) = \alpha$$

for $r*$, where $D = diag\{n(z_t - \bar{z})/[2\sum(z_t - \bar{z})^2]^{1/2}\}, t = 1, ..., n$. True critical values were computed for each of these tests at significance levels $\alpha = 0.01$, 0.05, 0.10, assuming normally distributed disturbances, using an approach analogous to that of Koerts and Abrahamse (1969) for the Durbin-Watson test, with maximum integration and truncation errors set to 10^{-6}.

The other five tests for heteroscedasticity studied used tabulated critical values, based on the assumption of normality, such that the nominal Type I error was α. These tests were: the Goldfeld and Quandt (GQF) (1965) test which used $F_{(n-c-2k)/2,(n-c-2k)/2}$, with $c = mod(n/5)$ omitted central observations; the asymptotic version (SZasym) of the Szroeter test using N(0,1) recommended by Griffiths and Surekha (1986); White's (1980) asymptotic test using $\chi^2_{k(k+1)/2-1}$; the Breusch and Pagan test (BPasym), and its modified 'studentised' version (BPmod), each using the asymptotic χ^2_1 critical value.

Breusch and Pagan, Godfrey (1978) and Griffiths and Surekha found the asymptotic Breusch-Pagan test rejected the null hypothesis less frequently than indicated by its nominal size. Bickel (1978), Koenker (1981), Koenker and Bassett (1982) suggested this test is sensitive to slight departures (particularly kurtosis) from normality and proposed a 'studentised' version, with the correct size for any disturbance law. The Breusch and Pagan test is given by the explained sum of squares of a regression of the vector of squared OLS residuals, $\hat{u}_t^2$, on an intercept and the postulated deflator z_t in (1), divided by $2\hat{\sigma}^4 = 2(\hat{u}'\hat{u}/n)^2$. The modified test (BPmod) substitutes an estimate of the variance of the squared disturbances in the denominator, based on the fact that $var(u_t^2) = 2\sigma^4$ under normality. The two tests are asymptotically equivalent and the modified version can be obtained as nR^2, where R^2 is the multiple correlation coefficient of the regression of $\hat{u}_t^2$ on an intercept and z_t. White's (1980) statistic is also of the form nR^2, where now R^2 corresponds to the regression of $\hat{u}_t^2$ on all products and cross products of the X regressors. Note that the modified BP and White's test require only second and fourth moments, not normality. Szroeter's asymptotic test (SZasym) is simply the approximately LBI test of King (sa), transformed by a function of the sample size (see Judge, Griffiths, Hill, Lutkepohl and Lee (1985)).

Knowledge of the deflator z_t is required for some of the heteroscedasticity tests, and for the purpose of this experiment, this variable was assumed to correspond to the first non-constant regressor. For all the heteroscedasticity tests, regressors were ordered according to increasing values of z_t. The accuracy of these "normal" critical values against non-normal distributions was investigated with 2000 replications, such that the sampling standard error was $(\alpha(1-\alpha)/2000)^{1/2}$. (Some experiments were conducted with 10000 replications also, but the results did not differ significantly from those using only 2000).

The experiment was undertaken in four stages, each examining a different range of alternative error distributions which are specified below. The first consisted of standard statistical distributions, reflecting a range of 'typical' and extreme behaviour, as well as those of other robustness studies. The later stages all considered distributions designed for Monte Carlo studies by Ramberg, Dudewicz, Tadikamalla and Mykytka (1979). These, denoted the RST family of distributions by McCabe, are a generalisation of Tukey's lambda distributions. In these experiments, each had a zero mean and unit variance and were characterised by the skewness and kurtosis coefficients. Departures from normality have generally been considered in terms of kurtosis and skewness, though for some distributions this may not be sufficient. Skewness is generally measured by $\sqrt{\beta}_1 = \mu_3/\mu_2^{3/2}$, where μ_i is the ith moment about the mean, and symmetric distributions have a value zero. Kurtosis is measured by $\beta_2 = \mu_4/\mu_2^2$, and the normal distribution has a value 3, with longer (shorter) tailed distributions having larger (smaller) values.

The first stage of the experiment considered a range of empirically relevant as well as extreme disturbance distributions generated from IMSL subroutines:

. standard N(0,1) variates, a benchmark with $\beta_2 = 3$, generated by the inverse normal transformation from routine GGNML.

. Variates from a symmetric short-tailed uniform (0,1) distribution with $\beta_2 = 1.8$.

. Variates from a symmetric Student's t distribution (t(5)) with 5 degrees of freedom, generated as the ratio of N(0,1) and the square root of a $\chi_5^2/5$ variate. This distribution is near normal, but with heavier tails, with $\beta_2 = 7$.

. Variates from a skewed exponential distribution (Exptl) with scale parameter 2, such that $\sqrt{\beta}_1 = 2$, and $\beta_2 = 9$, generated from GGEXN.

. Skewed non-negative χ_1^2 variates (Chisq1), with $\sqrt{\beta}_1 = 2^{3/2}$, and $\beta_2 = 15$, generated from GGCHS.

. Symmetric heteroscedastic normal variates, as used by White and MacDonald (1980), generated by transforming N(0,1) variates to have a variance equal to that of a χ_1^2 variate, with $\beta_2 = 8.6$.

. Skewed lognormal variates (Lognor), generated as the exponent of N(0,1) variates, with $\sqrt{\beta}_1 = 6, \beta_2 = 82$.

. Stable errors, generated using GGSTA, with a characteristic exponent of 1.1, and skewed with parameter 1 (as in Bartels and Goodhew).

. Variates from a fat-tailed symmetric Cauchy distribution, with no finite moments, generated by GGCAY with zero location, unit scale, and unit shape parameter.

An attempt was then made to isolate systematically the effects of skewness and kurtosis by considering the RST distributions used by McCabe. These were: symmetric ($\sqrt{\beta}_1 = 0$) with light ($\beta_2 = 2$), medium ($\beta_2 = 4$) and heavy ($\beta_2 = 10$) tails; right ($\sqrt{\beta}_1 = 0.5$) and heavy right ($\sqrt{\beta}_1 = 0.8$) skewed, but with 'normal' ($\beta_2 = 3$) kurtosis; right skewed combined with medium and heavy tails: and heavy right skewed with medium and heavy tails. Note that the distribution with $\sqrt{\beta}_1 = 0, \beta_2 = 3$ is the approximation to the normal from the RST distributions. The effect of kurtosis alone was then examined by considering symmetric distributions, from this same RST family, with increasing degrees of kurtosis, namely $\beta_2 = 2$ (light tail), 2.6, 3 ('normal'), 3.4, 4 (medium tail), 5, 6, 7, 8, 9 and 10 (heavy tail), respectively. The distributions for the final experiment reflected these with systematically increasing kurtosis, except that each was also skewed with $\sqrt{\beta}_1 = .7$.

4. Results

For each of the seven data sets and nineteen tests and under each alternative error distribution, the tail probabilities for the nominal significance levels, $\alpha = .01$, .05 and .10, were determined. Robustness was evaluated in terms of the number of sampling standard errors difference between the empirical and nominal size. With 2000 replications, these computed standard errors were .222, .487, .671 for significance levels of 1%, 5% and 10%, respectively.

Ali and Giacotto declared a test robust at the 5% nominal level if its rejection rate did not exceed this value by 2 standard errors, and moderately robust if it was less than 7.5% plus this value. For Conover, Johnson and Johnson a robust test had a maximum Type I error rate less than 10% at the 5% level. Neither considered cases of underrejection of the null hypothesis, which may be more serious. The Pearson and Please (1975) criterion of "acceptable", requires that the size is in the range 3 - 7% at 5% nominal level and 0 - 2% at the 1% level, which corresponds to a deviation of just over ± 4 standard errors. In the following analysis, the criteria for "robustness" is 2 standard errors, and for "acceptable" is 4 standard errors, at the relevant significance level.

A sample of results of empirical sizes is given in the accompanying tables for three data sets, with 64 observations, for tests of positive autocorrelation and heteroscedasticity with a nominal size of 5%. To relieve congestion, rarely used tests such as LM1, and some distributions have been deleted. Complete results are available on request.

Table I presents selected results for alternative distributions of a known type. For tests of autocorrelated disturbances, the size was generally robust for all significance levels under mild non-normality, as represented by the symmetric, uniform and t(5) distributions. For the skewed and increasingly kurtotic exponential, heteroscedastic-normal and χ_1^2 distributions, results were often within 2 standard errors and, with the exception of DW4 (and occasionally BW4), the sizes were within the "acceptable" range of 4 standard errors. Sensitivity increased with increasing departures from normality, especially for lower significance levels. With the more extreme lognormal, Cauchy and Stable distributions, all but the fourth-order tests DW4 and BW4 were "moderately robust", by the Ali and Giacotto standard. For distributions with no finite moments, such as the Stable and Cauchy, the least-squares assumptions break down. Even then, test sizes were acceptable by the Conover, Conover and Johnson criterion, for the Cauchy distribution, and (with DW4 a borderline case) for the Stable distribution. Tests of negative correlation were similarly robust to moderate deviations from normality. However, their sizes were more susceptible to the extreme distributions, with variations of up to 10 standard errors for the extreme Stable and Cauchy distributions. As expected, sensitivity was heightened for smaller α values. There did not appear to be significant differences between the results for different data sets.

As Table I demonstrates, all the tests of heteroscedasticity are highly susceptible to non-normal disturbances, with the exception of the White and modified Breusch-Pagan (BPmod) tests, which are not based on the assumption of normality. The latter appears the more robust, as the White test was affected by extreme non-normal distributions for low α values for some data sets. The empirical size of all other tests was too small for the light tailed uniform distribution, and too large for the heavy tailed distributions, increasing dramatically and consistently with deviations from normality. None of the tests, derived under the assumption of normality, could be considered robust or "acceptable", by any of the two-sided criteria cited above, and the sizes of all but the light tailed uniform distribution, exceeded even the Conover, Conover and Johnson criterion of 10% at the 5% significance level. The Breusch-Pagan test, with either a true or asymptotic critical value, was the most afflicted by deviations from normality, with sizes exceeding 50% for the extreme Stable and Cauchy distributions, and less than 1% for the uniform distribution at the 5% nominal level.

Under normality, the empirical size of the White test, using the asymptotic χ^2 critical value, was usually too low, and the χ^2 critical values of the modified BP test did not always give a reasonable size. Our results also support the previous findings that, under normality, the Breusch-Pagan test using the asymptotic critical value rejects the null hypothesis less frequently than indicated by the nominal size.

Tables 2 and 3 display selected results from the examination of the implications for the tests of different combinations of skewness and kurtosis. The RST distributions were symmetric, right ($\sqrt{\beta}_1 = .5$) or heavy right ($\sqrt{\beta}_1 = .8$) skewed, with each of these degrees of skewness combined with increasing kurtosis, including light ($\beta_2 = 2$), medium ($\beta_2 = 4$) and heavy ($\beta_2 = 10$) tails, respectively. Examples of results for distributions with both skewness and kurtosis are shown in Table 2, and for symmetric but increasingly kurtotic distributions in Table 3. The distribution denoted normal*, with $\sqrt{\beta}_1 = 0, \beta_2 = 3$, is the approximation to the normal from the RST family.

As demonstrated in Table 2, autocorrelation tests appeared robust to both skewness and kurtosis, the size generally being within the 2 standard errors range. The few exceptions, all within 3 standard errors and hence quite "acceptable", were generally for tests of negative autocorrelation, for the more marked departures from normality, i.e. with either a heavy skew, a heavy tail, or both, suggesting a slight underrejection of the null hypothesis. Similarly, for symmetric distributions (see Table 3), with kurtosis measures ranging from 2 to 10, the empirical sizes of tests for autocorrelation generally lay within 2 standard errors of the "normal" size for all data sets. Occasionally the sizes for some AR(1) tests lay within 3 standard errors, being slightly too high with light-tailed distributions against positive correlation, and slightly low for fat-tailed distributions and negative correlation.

Skewness did not appear to be a significant determining factor in the size of tests for heteroscedasticity. With the exception of the White and modified Breusch-Pagan tests, which were robust to the effect of both skewness and kurtosis, in each comparison, significant variation corresponded to a change in the kurtosis, rather than skewness. As can be seen in Table 2, the combination of skewness and kurtosis was similar to that of kurtosis alone, certainly with respect to the "acceptable" 4 standard error criterion, and usually within the robust measure of 2 standard errors. As further confirmation, when a skewness factor of $\sqrt{\beta}_1 = 0.7$ was combined with the RST distributions having systematically increasing kurtosis, the results did not significantly alter from those shown in Table 3. A minor exception was for a lighter tail distribution ($\beta_2 = 2.6$) for some data sets, suggesting a possible slightly counteracting effect, in that the empirical size was less of an understatement of the nominal value.

A systematic exploration of the effect of kurtosis (see Table 3) highlighted the sensitivity of tests for heteroscedasticity to this factor. Only the White and modified BP tests were not vulnerable to increasing kurtosis. For all the other tests, the empirical size was less than the nominal value for light-tailed distributions, and the reverse was true for heavy tails. The heavier the tails, the greater the divergence from the value expected under normality. Generally only distributions with small deviations ($\pm .4$) from normal kurtosis resulted in tests which could be considered "acceptable" and none were robust. For example, at the 5% level, with 64 observations, the empirical size was generally 1-2% for light tails, 8-9% for medium tails, and increased almost linearly between up to about 14% for $\beta_2 = 8$. This effect was most marked for the Breusch-Pagan test with true or asymptotic critical values, with empirical sizes approximating 20% for some data sets.

Generally sensitivity of tests decreased for higher levels of significance, which is consistent with other studies, and with intuition, as the tails of the distribution are likely to be most vulnerable. A similar pattern of behaviour was observed for samples sizes of 24 and 64. In contrast to some other studies, such as Knight and Smith, our findings were generally true of all X matrices considered, and no marked differences between the results from the various data sets were apparent.

Our results overall suggest that tests against autocorrelated disturbances are usually robust in the presence of non-normal disturbances, and always "acceptable", except with very extreme distributions. In contrast, tests for heteroscedasticity appear to be highly vulnerable to kurtosis, except for the White and modified Breusch-Pagan test, which specifically allow for this factor in their construction. The original Breusch and Pagan test, using either the asymptotic and true critical value, appears most susceptible of all tests to extreme departures from normality. The heteroscedasticity tests which are based on a small sample justification and which use true critical values, namely, LM1, sa, sa(2.5), s(5) and SZ, all appear to behave in a similar fashion with respect to both skewness and kurtosis. With normal disturbances, empirical sizes using asymptotic critical values were often poor for the White test, the Breusch-Pagan test (BPasym) and its modification (BPmod), and the asymptotic Szroeter (SZasym) tests, the first two usually being too low.

Overall we found that skewness was not a problem with tests for heteroscedasticity. In this we concur with McCabe, who found that the GQF test was insensitive to skewness, but not with the conclusions of Ali and Giacotto and Barone-Adesi and Talwar, who chose the lognormal and χ^2_4 distributions, respectively, to represent skewed distributions. Each of these distributions also has a high kurtosis measure, so their conclusions do not necessarily follow. Finally the lack of robustness to kurtosis, found in traditional tests of heteroscedasticity, is also shared generally by more recent tests based on a small sample justification.

5. Concluding Results

This study has attempted a comprehensive examination of the effects of non-normal error disturbances on the size of tests for non-spherical disturbances in the general linear regression model. It is reassuring that tests for autocorrelated disturbances appear relatively robust to departures of normality of the disturbances, except for the most extreme distributions. This appears to be the case for data with a range of characteristics.

Previous findings that high kurtosis of the error distribution tends to lead to higher rejection rates of the null hypothesis for tests of heteroscedasticity have been confirmed and also apply to more recent tests based on small sample properties. This sensitivity is not surprising: distinguishing between "normal" distributions with σ^2 generated randomly or with a few outliers and thick-tailed "non-normal" distributions may be difficult. In contrast to the conclusions of some other studies, skewness does not appear to be a critical factor for tests of heteroscedasticity. Of course, solving the problem of the size of the test does not guarantee high power, and this aspect will be the subject of future research.

References

Ali, M.M. and C. Giacotto (1984) "A Study of Several New and Existing Tests for Heteroscedasticity in General Linear Model", *Journal of Econometrics*, 26, 355-374.

Barone-Adesi, G. and P.P.Talwar (1983) "Market Models and Heteroscedasticity of Residual Security Returns", *Journal of Business and Economic Statistics*, 1, 2, 163-168.

Bartels, R. and J. Goodhew (1981) "The Robustness of the Durbin-Watson Test", *The Review of Economics and Statistics*, 63, 136-139.

Bera, A.K. and C.M. Jarque (1981) "Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals: Monte Carlo Evidence, " *Economics Letters*, 7, 313-318.

Berenblut, I. I. and G. I. Webb (1973) "A New Test for Autocorrelated Errors in the Linear Regression Model", *Journal of the Royal Statistical Society*, B, 35, 33-50.

Bickel, P.J. (1978) "Using Residuals Robustly I: Tests for Heteroscedasticity and Nonlinearity", *Annals of Statistics*, 6, 266-291.

Box, G.E.P. (1953) "Non-normality and Tests on Variances", *Biometrika*, 40, 318-355.

Box, G. E. P. and G. S. Watson (1962) "Robustness to Non-normality of Regression Tests", *Biometrika*, 49, 93-106.

Breusch, T.S. and A.R. Pagan (1979) "A Simple Test for Heteroscedasticity and Random Coefficient Variation", *Econometrica*, 47, 1287-1294.

Brown, M.B. and A.B. Forsythe (1974) " Robust Tests for the Equality of Variances", *Journal of the American Statistical Association*, 69, 364-367.

Cavanagh, C.L. and T.J. Rothenberg (1985) "Linear Regression with Non-normal Errors", mimeo.

Conover, W.J., M.E. Johnson and M.M.Johnson (1981) "A Comparative Study of Tests for Homogeneity of Variances, with Applications to the Outer Continental Shelf Bidding Data", *Technometrics*, 23, 4, 351-361.

Durbin, J. and G.S. Watson (1950) "Testing for Serial Correlation in Least Squares Regression, I", *Biometrika*, 37, 409-428.

Durbin, J. and G.S. Watson (1951) "Testing for Serial Correlation in Least Squares Regression, II", *Biometrika*, 38, 159-178.

Durbin, J. and G.S. Watson (1971) "Testing for Serial Correlation in Least Squares Regression, III", *Biometrika*, 58, 1-19.

Evans, M. A. (1985) "The Invariance of Tests of Disturbance Specification to Multicollinearity", *Bulletin of Economic Research*, 37, 3, 195-199.

Evans, M. A. and M. L. King (1985a) "Critical Value Approximations for Tests of Linear Regression Disturbances", *Australian Journal of Statistics*, 27, 68-83.

Evans, M.A. and M.L. King (1985b) "A Point Optimal Test for Heteroscedastic Disturbances", *Journal of Econometrics*, 27, 163-178.

Evans, M.A. and M.L.King (1988) "A Further Class of Tests for Heteroscedasticity", *Journal of Econometrics*, 37, 265-276.

Gartside, P.S. (1972) " A Study of Methods for Comparing Several Variances", *Journal of the American Statistical Association*, 67, 342-346.

Gastwirth, J. L. and M. R. Selwyn (1980) "The Robustness Properties of Two Tests for Serial Correlation, "*Journal of the American Statistical Association*, 75, 138-141.

Godfrey, L.G. (1978) "Testing for Multiplicative Heteroscedasticity", *Journal of Econometrics*, 8, 227-236.

Goldfeld, S.M. and R.E. Quandt (1965) "Some Tests for Homoscedasticity", *Journal of the American Statistical Association*, 60, 539-547.

Griffiths, W.E. and K. Surekha (1986) "A Monte Carlo Evaluation of the Power of Some Tests for Heteroscedasticity", *Journal of Econometrics*, 31, 219-231.

Hillier, G. and M. Smith (1983) "On the Size of the Durbin-Watson Test Under Moderate Non-Normality". Mimeo. Monash University.

Judge, G.G., W.E. Griffiths, R.C. Hill, H. Lutkepohl and T-C. Lee (1985) *The Theory and Practice of Econometrics*, Wiley.

Kariya, T. (1977) "A Robustness Property of Tests for Serial Correlation", *The Annals of Statistics*, 5, 1212-1220.

Kariya, T. and M.L. Eaton (1977) "Robust Tests for Spherical Symmetry", *The Annals of Statistics*, 5, 206-215.

King, M.L. (1979) "Some Aspects of Statistical Inference in the Linear Regression Model", Unpublished Ph.D. thesis, University of Canterbury, Christchurch.

King, M.L. (1980) "Robust Tests for Spherical Symmetry and their Application to Least Squares Regression", *The Annals of Statistics*, 8, 1265-1271.

King, M. L. (1981) "The Alternative Durbin-Watson Test:An Assessment of Durbin and Watson's Choice of Test Statistic", *Journal of Econometrics*, 17, 51-66.

King, M.L. (1984) " A New Test for Fourth-order Autoregressive Disturbances", *Journal of Econometrics*, 24, 269-277.

King, M.L. (1985) " A Point Optimal Test for Autoregressive Disturbances", *Journal of Econometrics*, 27, 21-37.

Knight, J. L. (1985) "The Sensitivity of the Durbin- Watson Statistic to the Assumption of Normal Disturbances", Mimeo. University of New South Wales.

Koenker, R. (1981) " A Note on Studentizing a Test for Heteroscedasticity", *Journal of Econometrics*, 17, 107-112.

Koenker, R. and G. Bassett (1982) " Robust Tests for Heteroscedasticity based on Regression Quantiles", *Econometrica*, 50, 43-61.

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

McCabe, B. (1989) "Misspecification Tests in Econometrics Based on Ranks", *Journal of Econometrics*, 40, 261-278.

Pearson, E. S. and N.W. Please, (1975) "Relation between the Shape of Population Distribution and the Robustness of Four Simple Test Statistics", *Biometrika*, 62, 223-241.

Ramberg, J.S., E.J. Dudewicz, P.R. Tadikamalla and E.F. Mykytka (1979) "A Probability Distribution and Its Uses for Fitting Data", *Technometrics*, 21, 201-214.

Rivest, L-P. (1986) "Bartlett's, Cochran's and Hartley's Tests on Variances are Liberal When the Underlying Distribution is Long-Tailed", *Journal of the American Statistical Association*, 81, 124-128.

Simkowitz M.A. and W.L. Beedles (1980) "Asymmetric Stable Distributed Security Returns", *Journal of the American Statistical Association*, 75, 306-312.

Smith, M. (1987) Unpublished Ph.D. thesis, Monash University.

Szroeter, J. (1978) "A Class of Parametric Tests for Heteroscedasticity in Linear Econometric Models", *Econometrica*, 46, 1311-1327.

Wallis, K.F. (1972) " Testing for Fourth Order Autocorrelation in Quarterly Regression Equations", *Econometrica*, 40, 617-636.

Webb, G.I. (1973) "Autocorrelations and the General Linear Model", Unpublished Ph.D. thesis, The City University, London.

White, H. (1980) "A Heteroskedasticity-consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity", *Econometrica*, 48, 817-838.

White, H. and G.M.MacDonald (1980) " Some Large-Sample Tests for Nonnormality in the Linear Regression Model", *Journal of the American Statistical Association*, 75, 16-27.

Table 1
Sizes of tests under alternative standard distributions
64 observations, nominal size 5%, standard error .5

$\sqrt{\beta_1}, \beta_2=$	Uniform 0,1.8	Normal 0,3	t(5) 0,7	Exptl 2,9	Chisq1 2.8,15	Lognor 6,82	Stable	Cauchy
				Lognormal data				
DW1	5.5	5.2	5.4	5.7	5.8	6.7	5.8	4.5
DW1alt	5.6	5.2	5.0	5.1	4.9	6.2	4.1	2.4
s1(.5)	5.5	5.0	5.2	5.3	5.4	6.1	5.1	3.1
BW1	5.5	5.0	5.1	5.6	5.8	6.7	5.8	4.5
DW4	5.1	5.3	6.3	7.0	7.2	8.6	10.5	9.9
DW4alt	5.8	5.7	5.5	6.2	4.9	5.4	3.5	1.9
s4(.5)	5.7	5.6	5.7	6.2	5.8	6.3	6.1	5.1
BW4	5.7	5.9	6.3	6.8	6.5	7.8	7.7	6.2
sa	1.3	5.1	10.8	16.8	21.3	26.2	33.7	35.7
sa(2.5)	1.2	5.0	11.1	18.0	23.4	30.1	38.9	40.6
s(5.0)	1.7	5.0	10.0	17.8	23.5	29.6	40.0	40.9
SZ	1.2	5.0	11.0	16.9	21.8	27.4	34.9	37.1
SZasym	1.4	5.3	11.3	17.3	21.7	26.9	33.9	35.9
GQF	1.4	4.7	11.7	17.2	23.7	30.0	37.5	40.5
BPtrue	.7	5.0	13.3	21.0	29.5	39.4	51.6	54.4
BPasym	.3	3.8	11.1	17.1	26.3	35.6	49.9	52.5
BPmod	5.8	3.6	3.7	4.6	5.1	5.3	5.9	5.6
White	4.0	3.9	4.9	5.2	6.2	6.0	5.5	4.4
				Trend data				
DW1	5.7	5.3	5.2	5.4	5.9	7.1	6.1	4.4
DW1alt	5.7	5.4	5.0	4.9	5.3	6.0	4.2	2.4
s1(.5)	5.7	5.2	5.0	5.1	5.7	6.5	5.2	3.3
BW1	5.8	5.3	5.2	5.4	6.0	7.1	6.1	4.4
DW4	5.1	5.4	6.0	6.5	7.7	8.3	10.2	9.5
DW4alt	5.7	5.5	5.5	6.0	5.3	5.2	3.8	1.6
s4(.5)	5.5	5.2	5.7	6.3	6.2	6.4	6.3	5.6
BW4	5.5	5.2	5.9	7.1	6.6	7.4	7.5	6.0
sa	1.4	4.8	10.9	17.3	21.5	26.6	34.0	36.0
sa(2.5)	1.2	4.9	11.2	17.9	23.6	30.4	39.3	41.1
s(5.0)	1.1	4.9	11.2	18.2	25.0	31.0	41.3	43.3
SZ	1.4	4.7	10.9	17.2	22.1	27.5	35.1	37.2
SZasym	1.2	4.7	10.6	16.9	21.1	26.4	33.9	35.8
GQF	1.4	4.8	11.0	16.9	23.5	29.4	37.4	40.2
BPtrue	.5	4.7	15.4	25.0	34.2	46.1	62.8	66.1
BPasym	.5	4.0	14.8	24.0	33.4	45.8	61.7	65.7
BPmod	5.6	4.6	3.8	5.1	5.0	4.7	3.3	2.4
White	5.3	4.4	4.5	4.8	5.3	6.0	6.5	6.0
				Uniform data				
DW1	5.4	5.4	5.1	4.8	5.9	7.3	6.0	4.7
DW1alt	5.3	5.1	5.2	4.5	5.0	5.9	4.2	2.4
s1(.5)	5.5	5.0	5.2	4.8	5.5	6.2	5.2	3.2
BW1	5.6	5.3	5.1	5.0	6.1	7.0	6.0	4.7
DW4	5.3	4.9	6.3	6.6	7.8	8.3	10.2	9.6
DW4alt	5.5	5.7	5.6	5.8	5.0	5.0	3.8	1.9
s4(.5)	5.7	5.5	6.2	6.0	6.3	6.0	5.7	4.6
BW4	5.8	5.4	6.5	6.4	6.7	7.2	7.3	6.1
sa	1.4	4.7	11.1	17.1	21.6	26.6	34.1	36.1
sa(2.5)	1.2	4.7	11.1	18.0	23.5	30.4	39.2	41.3
s(5.0)	1.3	4.6	11.1	18.0	23.9	30.6	39.8	41.9
SZ	1.4	5.0	10.9	17.1	22.2	27.4	35.2	37.1
SZasym	1.2	4.4	10.6	16.6	20.7	26.2	33.6	35.7
GQF	1.4	4.8	10.9	16.9	23.7	29.5	37.4	40.3
BPtrue	.6	4.9	14.1	25.0	33.8	45.4	62.6	65.5
BPasym	.5	4.4	13.6	24.3	32.7	44.8	61.8	64.9
BPmod	5.5	4.4	3.6	5.1	5.6	5.2	5.3	4.1
White	5.6	4.4	4.4	5.1	5.3	6.1	7.4	7.4

Table 2
Sizes under distributions with right skewness and tail kurtosis
64 observations, nominal size 5%, standard error .5

	normal tail			medium tail		heavy tail	
$\sqrt{\beta_1},\beta_2=$	normal* 0,3	skew .5,3	heavyskew .8,3	skew .5,4	heavyskew .5,10	skew .8,4	heavyskew .8,10
Lognormal data							
DW1	5.6	6.1	5.8	5.6	5.8	5.3	5.4
DW1alt	5.5	6.0	5.7	5.6	5.7	5.2	5.2
s1(.5)	5.4	5.5	5.6	5.6	5.5	5.1	5.2
BW1	5.5	5.7	5.7	5.5	5.7	5.3	5.4
DW4	4.9	4.9	4.9	5.0	5.0	5.1	5.5
DW4alt	4.6	4.7	4.5	4.5	4.6	4.6	4.7
s4(.5)	4.4	4.4	4.6	4.4	5.0	4.5	4.5
BW4	4.3	4.2	4.5	4.5	4.7	4.6	4.5
sa	6.1	6.0	6.0	9.3	9.1	14.5	14.0
sa(2.5)	6.1	6.0	6.3	9.5	9.1	14.7	14.6
s(5.0)	5.7	6.0	6.5	7.9	8.2	12.6	12.1
SZ	6.3	5.7	6.0	9.3	9.1	15.2	14.6
SZasym	6.4	6.4	6.2	9.7	9.6	14.9	14.4
GQF	6.1	6.1	5.9	9.3	8.9	14.3	14.5
BPtrue	6.0	6.6	6.3	9.7	10.2	17.8	17.4
BPasym	4.6	4.9	4.8	7.9	8.0	15.6	15.3
BPmod	4.2	5.4	6.0	4.9	5.0	4.5	4.9
White	4.4	5.1	5.6	4.5	5.2	5.0	5.2
Trend data							
DW1	5.8	5.8	5.8	5.8	5.7	5.5	5.5
DW1alt	5.9	5.8	5.9	5.7	5.6	5.0	5.3
s1(.5)	5.7	5.9	6.1	5.7	5.5	5.2	5.4
BW1	5.8	5.8	5.8	5.7	5.6	5.5	5.5
DW4	4.8	4.7	4.8	4.8	5.1	5.4	5.6
DW4alt	4.9	4.7	4.9	4.8	5.0	4.6	4.5
s4(.5)	4.7	4.6	4.9	4.9	4.5	4.6	4.8
BW4	4.5	4.6	4.8	4.7	4.8	4.4	4.4
sa	6.1	6.1	6.1	9.7	9.4	14.3	14.0
sa(2.5)	6.1	5.9	6.7	9.4	9.1	14.5	14.3
s(5.0)	6.0	6.1	6.3	8.7	8.8	14.2	13.9
SZ	6.2	5.9	6.2	9.1	8.9	15.1	14.5
SZasym	5.9	5.9	5.9	9.4	9.3	14.0	13.7
GQF	6.3	6.0	5.9	9.1	8.7	14.7	14.5
BPtrue	5.8	5.9	6.9	10.6	11.0	21.3	21.1
BPasym	5.6	5.3	6.3	10.2	10.3	20.6	20.5
BPmod	5.4	6.0	6.8	5.8	5.8	4.7	4.9
White	4.9	5.7	6.5	5.1	5.6	4.9	5.4
Uniform data							
DW1	6.0	6.0	5.8	6.2	5.9	6.0	6.0
DW1alt	6.0	6.0	6.2	6.0	5.8	5.9	6.0
s1(.5)	6.0	5.8	6.0	6.0	6.0	5.7	5.7
BW1	5.9	5.8	5.8	5.8	5.8	5.7	5.8
DW4	4.8	5.0	4.8	4.9	5.0	5.2	5.3
DW4alt	4.6	4.9	4.7	4.7	4.6	4.8	4.8
s4(.5)	4.4	4.9	5.0	4.7	5.0	4.6	4.6
BW4	4.6	4.8	5.0	4.9	4.9	4.6	4.9
sa	6.0	6.2	6.2	9.7	9.5	14.3	13.9
sa(2.5)	6.1	5.9	6.8	9.2	9.4	14.3	14.5
s(5.0)	6.2	5.9	6.6	9.1	9.2	14.5	14.6
SZ	6.0	5.8	6.4	9.1	9.0	15.1	14.4
SZasym	5.6	5.7	5.7	9.0	9.0	13.6	13.3
GQF	6.3	6.2	6.2	9.1	8.7	14.7	14.6
BPtrue	5.4	5.7	6.6	11.0	11.7	21.3	20.9
BPasym	4.9	5.3	6.2	10.1	11.1	20.4	19.9
BPmod	5.3	5.7	6.8	5.6	5.9	5.4	5.1
White	4.7	5.7	5.8	4.8	5.6	4.8	4.7

* This is the normal approximation from the RST distributions.

Table 3
Sizes under symmetric distributions with increasing tail kurtosis
64 observations, nominal size 5%, standard error .5

$\sqrt{\beta_1},\beta_2=$	light 0,2	normal* 0,3	medium 0,4	--- 0,5	increasing 0,6	kurtosis 0,7	--- 0,8	heavy 0,10
			Lognormal data					
DW1	5.6	5.6	5.5	5.5	5.4	5.4	5.3	5.3
DW1alt	5.5	5.5	5.5	5.3	5.2	5.3	5.3	5.2
s1(.5)	5.4	5.4	5.4	5.2	5.3	5.3	5.3	5.3
BW1	5.4	5.5	5.5	5.4	5.4	5.3	5.3	5.3
DW4	4.6	4.9	5.0	5.0	5.2	5.2	5.2	5.2
DW4alt	4.8	4.6	4.5	4.7	4.6	4.6	4.5	4.3
s4(.5)	4.7	4.4	4.3	4.3	4.3	4.2	4.2	4.3
BW4	4.3	4.3	4.4	4.4	4.4	4.3	4.3	4.4
sa	1.9	6.1	9.4	11.3	12.3	13.1	13.6	14.3
sa(2.5)	2.1	6.1	9.2	11.3	12.9	13.6	14.0	14.8
s(5.0)	2.6	5.7	8.1	9.6	10.6	11.4	11.9	12.6
SZ	1.7	6.3	9.3	11.3	12.9	13.7	14.4	15.2
SZasym	2.0	6.4	10.0	11.6	12.6	13.3	14.3	14.8
GQF	2.2	6.1	8.9	11.0	12.4	13.3	14.3	14.6
BPtrue	1.4	6.0	10.1	13.3	15.3	16.8	17.5	18.2
BPasym	.9	4.6	8.0	10.5	12.6	14.2	14.9	15.5
BPmod	5.5	4.2	4.4	4.2	4.3	4.3	4.2	4.3
White	4.6	4.4	4.6	4.7	4.7	4.7	4.7	4.8
			Trend data					
DW1	5.7	5.8	5.6	5.5	5.4	5.6	5.7	5.7
DW1alt	5.6	5.9	5.5	5.4	5.3	5.4	5.4	5.2
s1(.5)	5.6	5.7	5.5	5.3	5.2	5.4	5.5	5.4
BW1	5.6	5.8	5.6	5.5	5.4	5.6	5.7	5.7
DW4	4.4	4.8	4.9	5.2	5.1	5.2	5.5	5.5
DW4alt	4.7	4.9	5.0	4.7	4.7	4.7	4.7	4.7
s4(.5)	4.5	4.7	4.6	4.5	4.4	4.4	4.5	4.5
BW4	4.4	4.5	4.5	4.5	4.5	4.5	4.5	4.5
sa	1.7	6.1	9.2	11.0	12.3	13.4	13.9	14.2
sa(2.5)	2.0	6.1	9.1	11.0	12.6	13.1	13.7	14.3
s(5.0)	2.0	6.0	8.8	10.9	12.2	13.2	13.9	14.6
SZ	1.7	4.6	9.2	11.2	12.6	13.4	14.3	15.3
SZasym	1.5	5.9	8.9	10.8	12.1	13.0	13.6	14.1
GQF	2.2	6.3	9.0	10.9	12.3	13.4	14.4	14.8
BPtrue	.9	5.8	10.4	14.8	16.9	18.3	20.1	21.2
BPasym	.8	5.6	10.0	14.3	16.5	18.1	19.3	20.6
BPmod	5.3	5.4	5.2	4.9	4.9	4.8	4.6	4.6
White	5.2	4.9	4.5	4.7	4.4	4.4	4.4	4.3
			Uniform data					
DW1	5.7	6.0	5.9	5.7	5.7	5.8	6.0	6.0
DW1alt	5.9	6.0	5.9	5.9	5.7	5.7	5.7	5.7
s1(.5)	5.8	6.0	5.8	5.7	5.8	5.7	5.6	5.6
BW1	5.9	5.9	5.6	5.7	5.6	5.6	5.5	5.5
DW4	4.5	4.8	4.8	4.8	4.9	4.9	4.9	5.0
DW4alt	4.7	4.6	4.6	4.7	4.7	4.6	4.6	4.6
s4(.5)	4.8	4.4	4.3	4.3	4.4	4.2	4.1	4.1
BW4	4.5	4.6	4.6	4.5	4.5	4.5	4.7	4.6
sa	1.6	6.0	9.1	11.0	12.1	13.2	13.9	14.4
sa(2.5)	2.1	6.1	9.1	11.1	12.7	13.2	13.8	14.3
s(5.0)	2.1	6.2	8.9	11.3	12.4	13.2	13.9	14.3
SZ	1.6	6.0	9.2	11.2	12.8	13.5	14.1	15.4
SZasym	1.5	5.6	8.7	10.4	11.8	12.5	13.5	13.9
GQF	2.3	6.3	9.2	10.9	12.4	13.4	14.3	14.7
BPtrue	1.1	5.4	10.7	15.0	17.2	19.1	20.5	21.3
BPasym	.9	4.9	9.8	14.1	16.6	18.3	19.5	20.5
BPmod	5.3	5.3	4.8	4.9	4.9	4.9	4.8	4.8
White	5.1	4.7	4.9	4.6	4.7	4.6	4.7	4.8

* This is the normal approximation from the RST distribution.

<u>MONASH UNIVERSITY</u>

<u>DEPARTMENT OF ECONOMETRICS</u>

<u>WORKING PAPERS</u>

<u>1988</u>

6/88 Grant H. Hillier, "On the Interpretation of Exact Results for Structural Equation Estimators".

7/88 Kuldeep Kumar, "Some Recent Developments in Non-Linear Time Series Modelling.

8/88 Maxwell L. King, "Testing for Fourth-Order Autocorrelation in Regression Disturbances When First-Order Autocorrelation is Present".

9/88 Ralph D. Snyder, "Kalman Filtering with Partially Diffuse Initial Conditions".

10/88 Asraul Hoque, "Indirect Rational Expectations and Estimation in a Single Equation Macro Model".

11/88 Asraul Hoque, "Efficiency of OLS Relative to C-O for Trended x and Positive Autocorrelation Coefficient".

12/88 Russel J. Cooper & Keith R. McLaren, "Regular Alternatives to the Almost Ideal Demand System".

13/88 Maxwell L. King, "The Power of Student's t Test: Can a Non-Similar Test Do Better?".

14/88 John Hamilton, "On-Line Management of Time Series Databases: Database Retrieval Program DBR".

<u>1989</u>

1/89 R. D. Snyder, C. Chah & C. Lehmer, "Multi-series Heuristics for Exponential Smoothing.

2/89 Maxwell L. King and Phillip M. Edwards, "Transformations for an Exact Goodness-of-Fit Test of Structural Change in the Linear Regression Model".

3/89 Muhammad I. Bhatti and Maxwell L. King, "A Beta-Optimal Test of the Equicorrelation Coefficient".

4/89 Muhammad I. Bhatti, "Null Distribution of the Small Sample Mean Correlation Coefficient: An Application to Medical Research."

5/89 Grant H. Hillier, "On the Normalisation of Structural Equations: Properties of Direction Estimators".

6/89 Brett A. Inder, "A New Test for Autocorrelation in the Disturbances of the Dynamic Linear Regression Model".