MONASH

WP NO. 13/91

MONASH UNIVERSITY

AUSTRALIA

BURR CRITICAL VALUE APPROXIMATIONS FOR TESTS OF AUTOCORRELATION AND HETEROSCEDASTICITY

Merran A. Evans

Working Paper No. 13/91

November 1991

DEPARTMENT OF ECONOMETRICS

ISSN 1032-3813

ISBN 0 86746 974 9

BURR CRITICAL VALUE APPROXIMATIONS FOR TESTS OF AUTOCORRELATION AND HETEROSCEDASTICITY

Merran A. Evans

Working Paper No. 13/91

November 1991

DEPARTMENT OF ECONOMETRICS, FACULTY OF ECONOMICS COMMERCE & MANAGEMENT

MONASH UNIVERSITY, CLAYTON, VICTORIA 3168, AUSTRALIA.

BURR CRITICAL VALUE APPROXIMATIONS
FOR TESTS OF AUTOCORRELATION AND HETEROSCEDASTICITY

MERRAN A. EVANS

Department of Econometrics, Monash University

Summary

The accuracy of Burr approximations of critical values and p–values is evaluated for tests of autocorrelation and heteroscedasticity in the linear regression model.

Key words: Autocorrelation; Heteroscedasticity; Critical value approximations; Linear regression; Burr distributions.

1. Introduction

Many tests of autocorrelated and heteroscedastic disturbances in the general linear model have an unknown probability distribution so true critical values cannot be tabulated. Sometimes they can be determined but their use may not be feasible in practice, and a few have published bounds for these values, involving an inconclusive region which can be large for small samples. Accurate approximations therefore would be most useful for applied regression analysis, particularly given the serious consequences of misspecification of non-spherical disturbance behaviour.

Some approximations of these critical values, based on matching moments with the normal and beta distributions, have been evaluated by Evans and King (1985a). They found that normal and two-moment beta, and particularly the four-moment beta, approximations were reasonably accurate, with skewness being a determining factor. Subsequent investigation with more variable but 'characteristic' data sets has indicated that some tests of heteroscedasticity lie outside the skewness-kurtosis range of the beta distributions, so the four-moment beta approximation cannot be used. The original study is extended here to include more variable data sets and a popular test of heteroscedasticity, but primarily to consider approximations from two members of the Burr (1942) family of distributions, which have a considerably wider range of moment coverage and can model diverse shaped distributions.

Moments are matched with two related Burr distributions (Types III and XII) and the appropriate shape parameters identified as in the previous study. These Burr distributions are attractive in that, given these shape parameters, associated probabilities or 'p-'values for any specified value can easily be generated from the distribution function, and critical values for specified significance levels from its inverse, as each has a simple closed mathematical form. This is in contrast to the Pearson and Johnson families which traditionally have been used to approximate distributions. The beta (or Pearson Type I) and normal distributions, members of these families, also lack this feature. Matching-moment methods with all these families require the computation of moments, but only Burr approximations avoid numerical integration in determining the distribution function and its inverse. They therefore have a considerable computational advantage, as well as a wider moment coverage than the normal and beta approximations. It is of interest to determine if Burr approximations are competitive with the highly accurate four-moment beta approximations when both can be applied, and to explore their accuracy when they alone can be used. The focus is on the critical region in the tails of the distribution, so a high level of accuracy is demanded.

This study investigates the accuracy of Burr approximations, using a similar but expanded experimental design to Evans and King (1985a). Theoretical aspects are discussed in the next section and an empirical evaluation in section 3. Burr approximations are examined for tests of autocorrelated and heteroscedastic disturbances in the linear regression model for which the true distribution of the test statistic is unknown, but true critical values and sizes can be determined. The tests examined here are those most commonly used in practice, or which have been demonstrated to have good power properties.

2. Theoretical Discussion

2.1 Tests of Autocorrelation and Heteroscedasticity

Consider the linear regression model, with fixed regressors, and normally distributed disturbances,

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

where y is $n \times 1$, X is $n \times k$, β is $k \times 1$ and u is $n \times 1$. Consider, as in Evans and King (1985a), tests against first-order autoregressive ($AR(1)$) disturbances: the first-order

Durbin and Watson (1950) test (d_1), its locally best invariant (LBI) analogue (d_1') of King (1981), the Berenblut and Webb (1973) test (g_1), and King's (1985) point optimal test ($s_1(.5)$). Against simple fourth-order autoregressive ($AR(4)$) disturbances consider their fourth order analogues: the Wallis (1972) (d_4)test, King's (d_4') LBI version, Webb's (1973) g_4, and King's (1984) $s_4(.5)$ test. Also consider tests against heteroscedasticity of the additive form, $var(u_t) = \sigma_t^2 = \sigma^2 f(1 + \lambda z_t)$, where f is an unknown monotonically increasing non-negative function and z_t is a non-stochastic variable. These include Szroeter's (1978) bounds test (SZ), the LBI test discussed by King and Hillier (1985), the Evans and King (1985b,1988) point optimal $s(5.0)$ and approximate $s_a(2.5)$ tests and the one-sided approximate LBI s_a test, which is a transformation by the sample size of Szroeter's asymptotic test (see Judge *et al.* (1988, p452)). These tests reject the null hypothesis $H_0 : u \sim N(0, \sigma^2 I_n)$ for small values of the test statistic for heteroscedasticity and for positive autocorrelation. Note that all of these tests are invariant to multicollinearity in the regressors (see Evans (1985)).

These one-sided tests can be classified into two classes, each expressible as a ratio of quadratic forms in residuals and in disturbances. Tests based on Ordinary Least Squares (OLS) residuals include d_1, d_1', d_4, d_4', SZ and s_a and can be written in the form

$$t = \hat{u}'A\hat{u}/\hat{u}'\hat{u} = u'MAMu/u'Mu = u'Bu/u'Mu,$$

where A is some real symmetric $n \times n$ matrix, $\hat{u} = (I - X(X'X)^{-1}X')y = My = Mu$ is the OLS residual vector, $M = I - X(X'X)^{-1}X'$ and $B = MAM$. Tests based also on Generalised Least Squares (GLS) residuals include $s_1(.5), g_1, s_4(.5), g_4, LBI, s(5.0)$ and $s_a(2.5)$, which can be written in the form

$$t = \tilde{u}'\Sigma^{-1}\tilde{u}/\hat{u}'\hat{u} = u'Bu/u'Mu,$$

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 $B = \Sigma^{-1} - \Sigma^{-1}X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}$.

True critical value for a specified significance level α or, alternatively, the actual size for the approximated critical value t_{crit}, can be obtained from

$$Pr(t < t_{crit}) = Pr[u'(B - t_{crit}I)u < 0] = \alpha,$$

using an approach analogous to that of Koerts and Abrahamse (1969), with maximum integration and truncation errors set to 10^{-6}.

A popular test for heteroscedasticity is the two-sided Lagrange multiplier test of Breusch and Pagan (1979) using asymptotic χ^2 critical values which are known to be suspect. True critical values and sizes can be determined in a similar fashion to those above. For the Breusch-Pagan (1979,p1290) test, $BP = (\hat{u}'A\hat{u}/\hat{u}'\hat{u})^2$, where A is a diagonal matrix with ith element $\{n(z_i - \bar{z})/2[\sum(z_i - \bar{z})^2]^{1/2}\}$, for $i = 1, ..., n$, against heteroscedasticity of the form considered here,

$$Pr(BP > t_{crit}^2) = Pr(\hat{u}'A\hat{u}/\hat{u}'\hat{u} > t_{crit}) + Pr(\hat{u}'A\hat{u}/\hat{u}'\hat{u} < -t_{crit}).$$

This test was not included in the original study, so normal, two- and four-moment beta approximations are also calculated.

As each of these test statistics can be written as ratios of quadratic forms in normal variables, their moments can be obtained using the methods of Henshaw (1966) and Evans and King (1985a). These are based on the result (see Durbin and Watson (1950)) that under the null hypothesis, the test statistic t is distributed independently of its own denominator, so that the moments of t are ratios of the corresponding moments of the numerator $u'Bu$ and denominator $u'Mu$. The moments about the origin then involve expectations of traces of products of the matrices B, or equivalently products of powers of their eigenvalues. The centralised moments can thus be determined and matched with those of a distribution from the Burr family. The first four moments are used: the mean $(\mu_{(t)})$, variance $(\sigma^2_{(t)})$, skewness $(\sqrt{\beta}_1)$ and kurtosis (β_2). Bounds of these distributions can also be determined from the extreme eigenvalues of the matrix B.

2.2 Burr XII and III Distributions

Twelve solutions to a differential equation given in Burr (1942) can be classified by their functional forms, each defining a family of cumulative distribution functions within the Burr system. The distribution functions of two solutions, Type XII and its 'reciprocal' Type III, are defined by shape parameters c and k, and for $x > 0, k > 0$ are given by

$$F(x) = 1 - (1 + x^c)^{-k} = \alpha,$$

$$F(x) = (1 + x^{-c})^{-k} = \alpha,$$

respectively. The significance level, size or p—value can be computed for specified values of x from $F(x)$. The inverse of this distribution function is easily obtained and the approximated critical value x_{crit} for a rejection region of nominal significance level α thus determined. For example, with a Burr XII rejection region in the lower tail,

$$x_{crit} = [(1 - \alpha)^{-1/k} - 1]^{1/c}.$$

For $c < 0$, using $x* = 1/x$ implies that the rejection region is in the upper tail, with calculations changing appropriately. The null hypothesis is then rejected for large values of x or small values of $1/x$, which is the appropriate Burr III distribution if $c < 0$.

Determining which of the family of Burr XII and III distributions most closely approximates that of a given statistic, t, involves matching the first four moments. For the selected Burr distribution to have the skewness and kurtosis corresponding to those calculated for t, $(\sqrt{\beta}_1, \beta_2)$, two non-linear simultaneous equations are solved for the shape parameters c and k:

$$\sqrt{\beta}_1 = \mu_3 / \sqrt{\mu_2^3} = \frac{\mu_3' - 3\mu_2'\mu_1' + 2\mu_1^3}{[\mu_2' - \mu_1^2]^{3/2}},$$

$$\beta_2 = \mu_4 / \mu_2^2 = \frac{\mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1^2 - 3\mu_1^4}{[\mu_2' - \mu_1^2]^2},$$

where μ_r are the rth moments about the mean, and μ_r' about the origin. For Burr XII $(c > 0)$ and Burr III $(c < 0)$,

$$\mu_r' = kB(1 + \frac{r}{c}, k - \frac{r}{c}) = \frac{\Gamma(1 + r/c)\Gamma(k - r/c)}{\Gamma(k)}.$$

Moments exist for $ck > r$ for Burr XII and $c < -r$ for Burr III. The distributions are unimodal if $c > 1$ for Burr XII and $|ck| > 1$ for Burr III.

For many distributions approximate or starting solutions can be obtained by interpolating Burr's (1973) table, but they do not cater for negative skewness, and coverage is sparse near the boundaries. An algorithm was developed to obtain the solutions. SHAPE, an interactive algorithm for use with the SAS package to fit Burr III (and hence also Burr XII) distributions has been developed by Rodriguez (1980).

The Burr distribution which matches the third and fourth moments is identified by the computed c and k, and its mean and variance are then determined as $\mu_B(c, k)$ and $\sigma_B^2(c, k)$. Matching the first two moments involves scaling and relocating each distribution to have a zero mean and unit standard deviation then equating these 'standardised' variates $(t - \mu_{(t)})/\sigma_{(t)}$ and $(x - \mu_B)/\sigma_B$, such that $t = \frac{\sigma_{(t)}}{\sigma_B}(x - \mu_B) + \mu_{(t)}$.

No general pattern is obvious, but some particular solutions c, k correspond to standard distributions. Relationships with other distributions and the extensive moment-coverage of the Burr XII and III distributions are shown in Rodriguez (1977, 1982) and

Tadikamalla (1980). Beta distributions correspond to the bounded Type I distributions of the Pearson family, which cover the smaller kurtosis areas of the moment-ratio diagrams. The potential area of moment coverage is bounded 'below' by the limit of all frequency distributions, $\beta_2 - \beta_1 - 1 = 0$. Near this limit are the U-shaped bimodal distributions, then with higher kurtosis for given skewness, the J-shaped, and next the unimodal Type I distributions which are bounded by the Pearson Type III or Gamma distributions line, $2\beta_2 - 3\beta_1 - 6 = 0$. The Burr coverage encompasses a much larger range of more skewed and kurtotic distributions. It overlaps the Type I area coverage for most beta distributions, but not the extreme U-shapes. It also includes the unbounded Pearson Types IV (in both directions), and VI (one direction) or F, distributions.

Consideration of the moment-coverage of each of the approximating distributions, suggests that the normal and two-moment beta approximations, which only consider the mean and variance of the matching distributions, will not be successful at approximating distributions with any significant skewness or kurtosis. The four-moment beta approximation will be applicable for all distributions for which $\beta_2 \leq 1.5(\beta_1 + 2)$. In general, methods of matching moments using approximating distributions, such as the Pearson, Johnson or Burr, do not explicitly take into account end points of the statistic, so that 'approximations to extreme percentage points will deteriorate sooner or later' (Bowman and Shenton (1982)). An exception to this is the beta approximation, based on a bounded distribution. As interest is focussed on the end points for critical regions this is likely to be an important factor and this success could be anticipated. However, Burr approximations have the appeal of simple analytic form for both sizes and critical values, so it is hoped that they might be competitive with the four-moment beta approximations. They also cover wider range of moments.

3. Empirical Evaluations

3.1 Outline of Experiment

To evaluate accuracy, true sizes of Burr critical values were computed and compared with nominal sizes $\alpha = .01$, .05 and .10, as was done in Evans and King (1985a). In addition, Burr p- values or size approximations, when the test involved the true critical values, were compared. Again the criterion for accuracy was whether the computed values lay within 1%, 5% or 10% of the true value, i.e. within $\pm$ (.0001, .0005, .0010) of $\alpha = .01$, within $\pm$ (.0005, .0025, .0050) of $\alpha = .05$ and within $\pm$ (.0010, .0050, .0100) of $\alpha = .10$.

The real design matrices contain a constant term and one or more other regressors which are reasonably typical economic time series data, with various degrees of trend and seasonality. These regressors include: the annual spirits income and price data of Durbin and Watson (1951); the weakly seasonal quarterly Australian Consumer Price Index (CPI), and also lagged one quarter; and quarterly Australian liquidity or capital movements, private and government, which are highly seasonal and subject to large fluctuations (not used in the previous study). Cross-sectional household population census data for the Australian population in 1961 and 1976 was also used. The artificial design matrices each comprised a constant and a regressor determined from a time trend (representing slowly evolving non-seasonal economic time series behaviour), the lognormal (representing skewed cross-sectional data) and uniform (standard in such experiments) distributions, respectively. Stable additive quarterly seasonal behaviour was represented by a set of 0-1 seasonal quarterly dummies, and Watson's X matrix is customary in experiments concerning autocorrelation. In summary, five data sets each were used for the four $AR(1)$ and four $AR(4)$ tests (each with sample sizes $n = 20,40,60$), as well as for the six heteroscedasticity tests (with $n = 24,40,64$), totalling 210 cases. Knowledge of the deflator z_t is required for some of the heteroscedasticity tests and this was assumed here to correspond to the first non-constant regressor, and regressors were ordered according to its increasing values.

Burr approximations were compared with those of the normal, two-moment beta $\beta(2)$ and four-moment beta $\beta(4)$, as described in Evans and King (1985). The normal approximation is based on assuming that $(t-\mu)/\sigma \sim N(0,1)$. For the $\beta(2)$ approximation the test statistic is first transformed by its extreme bounding values to lie in the range (0,1), then the mean and variance of this scaled test statistic are equated with those of a beta distribution, $B(p,q)$, to solve for its parameters p and q. The $\beta(4)$ approximation fits a beta distribution to this scaled test statistic by matching the skewness and kurtosis parameters to obtain the parameters p and q, then solving for the extreme eigenvalues using the first two moments.

Most of the autocorrelation and OLS-based heteroscedasticity tests studied were characterised by near symmetry and slightly short tails, so have moments which lie in the area covered by the beta distribution. Hence for these, approximations based on the normal and beta distributions, particularly the four-moment beta approximation, are likely to be reasonably accurate. The Breusch-Pagan and the GLS-based heteroscedasticity tests generally were more variable for the more extreme data sets, and some had skewness- kurtosis values outside the beta distribution range such that four-moment beta approximations are not applicable.

3.2 Results

Selected results on the approximations are given in the accompanying tables. For tests of autocorrelation ($AR(1)$ and $AR(4)$) and heteroscedasticity, respectively, Tables 1 to 3 show for each data set with 40 observations and for nominal sizes $\alpha = .01, .05, .10$: (a) the true test size when using the Burr approximation to the critical value; and (b) the calculated Burr $p-$value when the test involves the true critical value. Table 4 shows the true test size when using Burr and four-moment beta (when obtainable) approximations for the $s(5.0)$ and Breusch-Pagan tests of heteroscedasticity for sample sizes of 24 and 64. Full results are available on request from the author.

For the one-sided tests, some consistent patterns emerge from these results. The true size of the Burr approximation generally slightly exceeded the nominal size at $\alpha = 0.01$, and was slightly smaller for $\alpha = 0.05$ and 0.10, suggesting a cross-over of the true and approximating probability distribution function curves near the tails between $\alpha = .01$ and .05. The reverse results for the Burr p-values of tests using true critical values confirm this: for $\alpha = 0.01$, the Burr $p-$ values were slightly lower than the nominal size, but for higher values $\alpha = .05$ and .10, they were slightly higher. The exceptions to this pattern were for the highly skewed GLS-based tests, where the cross-over occurred between $\alpha =$ .05 and .10 for positive skewness, and a systematic shift with no cross-over for negative skewness.

Consider first the true sizes of tests involving Burr approximations to the critical values. For the $AR(1)$ tests, all these sizes were within 5% of the nominal value, and for $\alpha = .01$ usually within 1%. For the $AR(4)$ tests, approximations generally were within 5% of the nominal value except a few within 10% for the extreme tail ($\alpha = .01$) and small samples ($n = 20$). For $n = 40$ some, and for $n = 64$ most, were within 1% for $\alpha = 0.01$. The OLS-based heteroscedasticity tests, SZ and s_a, were all within 5% of the nominal sizes, and for $\alpha = 0.01$, s_a was within 1% for $n \geq 40$ as was SZ for $n = 64$ and usually for $n = 40$. Approximations for the more skewed GLS-based tests, $s(5.), s_a(2.5)$ and LBI, were not as accurate: for the extreme tail true sizes were usually within 10% of the nominal value, except for a few with small or moderate ($n = 40$) samples. However for higher nominal sizes, these tests were often within 5% and always within 10% of the nominal size.

Next consider the reverse situation: the Burr $p-$ value approximation using the true critical value. For the $AR(1)$ tests, for $\alpha = .01$, more than half the values lay within 1%, and all within 5% of the nominal size, and for higher α all were within 5%. For the $AR(4)$ tests, a few approximations for the extreme tail and small samples were only within 10%, but generally results were within 5%, and often within 1% for $\alpha = .01$. For the OLS based heteroscedasticity tests, s_a and SZ, approximations were always within 5% of the nominal size, and within 1% for $n = 64$, and s_a was within 1% for $n = 40$. Again approximations with the GLS-based tests were less precise: generally within 10% for $\alpha \geq .05$ and sometimes for $\alpha = .01$.

For the two-sided Breusch-Pagan test probabilities in the two tails are added, such that approximations reach further into the tails. For the extreme tail the true size of the Burr approximated critical value was within 10% of the nominal size for two data sets and within 5% for the trend data, but with larger samples it was generally within 10% and often within 5%. For larger values of α, accuracy improved: for $\alpha = .05$, approximations were within 10%, usually within 5%, and often within 1% of the nominal size; and for $\alpha = .10$, usually within 1% and always within 5%. For Burr p-value approximations for $\alpha \geq .05$, values were within 5% and often within 1% of the nominal value, whereas for the extreme tail values were generally within 10% and often within 5%. The few cases where Burr approximations for the Breusch-Pagan test were poor occurred for highly skewed data with $\alpha = 0.01$, involving examining the tails at significance levels .005 and .995. Even these, though, were considerably more accurate than those obtained using the conventional χ^2 asymptotic critical value, which illustrates the potential for this methodology in applied work (see Evans and Fry (1991)).

Overall, the few cases which gave less accurate approximations did so in both ways: (a) and (b) results were consistent. Of the 210 cases studied, ten lay on the Weibull boundary, which is the limiting distribution of Burr XII with $c > 0, k \to \infty$. These included some autocorrelation tests and one of heteroscedasticity, all nearly symmetric ($|\sqrt{\beta_1}| \simeq .02$) and with short tails ($\beta_2 \simeq 2.7$). The $s(5.0)$ heteroscedasticity test with the lognormal data had values of $(\sqrt{\beta_1}, \beta_2) = (1.3, 5.61), (1.6, 6.17), (1.53, 7.19)$, respectively, for n=24,40,64, on the same Weibull boundary. Problems arise as computation of $\Gamma(k)$ is not feasible for very large k and, as each of these ten had $k > 350$, the Weibull distribution was used:

$$F(x) = 1 - exp(-x^c), \qquad x = log[\alpha/(1-\alpha)]^{1/c}.$$

These Weibull approximations were successful: within 5% and often within 1%, of the nominal value, for all but one case with the extreme tail and a small sample, which was within 10%.

Compare now the relative performance of the Burr approximations with those previously examined in Evans and King (1985a). Consider first the normal approximations. For the autocorrelation tests approximations for $\alpha = .01$ were often not even within 10% of the nominal value. For $\alpha \geq .05$ they were generally within 5%, always 10% and sometimes 1% for $n \geq 40$. For the OLS-based heteroscedasticity tests approximations were within 5% and sometimes within 1% of the nominal size for $\alpha \geq .05$ and $n \geq 40$. However, in the extreme tail they were less successful: the s_a test was within 10% for all sample sizes, but the SZ test only for larger n. The GLS-based heteroscedasticity test approximations were poor, being within 10% of the nominal value only for a few cases with $n \geq 40, \alpha = .10$. Normal approximations to the Breusch-Pagan test generally were not even within 10% of the nominal value when there was any significant skewness and kurtosis. Results were reasonable only with the trend data, and some were within 10% of the true value for $\alpha = .10$.

Beta approximations based on the first two moments were reasonable when there was no significant skewness and kurtosis, such as the autocorrelation tests, where results were generally within 5% of the nominal size. However, for the $AR(1)$ and especially the $AR(4)$ tests, with small samples and the extreme tail, some results were only within 10% and some not even that. Generally for the autocorrelation tests, Burr approximations were more precise for $\alpha = .01$, but not for $\alpha \geq .05$, except for $s_4(.5)$. For the OLS-based heteroscedasticity tests, the $\beta(2)$ approximations were generally within 1% for SZ, and with $\alpha \geq .05$, for s_a. For $n = 24$, the Burr approximation was superior for s_a. For the highly skewed and variable GLS-based heteroscedasticity tests, these approximations were poor, and generally inferior to Burr: most results were not even within 10% of the true value. As with the normal approximations, results were poor when the test statistic had significant skewness or kurtosis.

Finally the four-moment beta approximations were very accurate, when obtainable. For the autocorrelation and OLS-based heteroscedasticity tests, results were usually within 1%, with only a few exceptions for small samples and the extreme tail. The GLS-based approximations, when obtainable, were within 1% generally for $n \geq 40$, and for $n = 24$ they were within 5% and usually within 1%. However, for most of the more variable regressors, approximations could not be determined for $s(5.0)$, the LBI and the Breusch-Pagan tests as $\beta_2 \geq 1.5(\beta_1 + 2)$, so that they did not lie in the beta range of moment coverage.

In summary, for tests with significant skewness and kurtosis, Burr approximations are generally superior to those from the normal and usually the $\beta(2)$ distributions. This is not particularly surprising, given that this skewness and kurtosis is specifically taken into account by the test statistic. However, although the Burr approximations are quite accurate, they are not as good as the $\beta(4)$ approximations, where the test statistic is first scaled and relocated by the estimated endpoints, and then two non-linear equations used to solve for the skewness and kurtosis coefficients, which is algebraically equivalent to scaling by the mean and standard deviation. All the test statistics in this study are bounded, such that bounded distributions are matched for the beta approximations, whereas bounded distributions are matched with unbounded with the Burr approach. As interest is in the tails of distributions, such extra knowledge of endpoints is obviously an advantage. However, in cases of marked skewness and kurtosis, which seems to occur often with some tests of heteroscedasticity, the four-moment beta approximation cannot be obtained, but Burr approximations can be determined and are reasonably accurate.

4. Concluding Remarks

This empirical evaluation of approximating critical values and p-values of tests of autocorrelation and heteroscedasticity by matching moments with the Burr distribution family suggests that tests using Burr approximations are reasonably accurate: values are generally within 5%, and at least within 10%, of the nominal size.

Burr approximate critical values, obtained by matching the first four moments, are superior to the normal and two-moment beta approximation for test statistics with any significant skewness and kurtosis. They are not as accurate as the four-moment beta approximation when used to match bounded distributions with skewness-kurtosis in the range covered by beta distributions. However these beta approximations cannot be determined for distributions with marked skewness and kurtosis, such as some heteroscedasticity tests, but quite reasonable Burr approximations can be obtained. Burr approximations have a wider moment coverage and marked computational advantages over other methods. An initial inspection of the existence regions of the Burr and beta distributions in terms of skewness and kurtosis of the test statistic of interest will indicate whether each these approximations is likely to be successful.

This methodolology is applicable to the many test statistics and estimators which can be written as a ratio of quadratic forms in normal variables. It should prove most useful in applied work.

References

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.

BOWMAN, K. O. and L. R. SHENTON (1982) Approximations to distributions, in S. Kotz and N. L. Johnson, eds. *Encyclopedia of Statistical Sciences,* Vol. 1, Wiley.

BREUSCH, T. S. and A. R. PAGAN (1979). A simple test for heteroscedasticity and random coefficient variation, *Econometrica,* 47, 1287-1294.

BURR, I. W. (1942). Cumulative frequency functions, *Annals of Mathematical Statistics*, 13, 215-232.

BURR, I. W. (1973). On a general system of distributions to match a grid of α_3 and α_4. *Communications in Statistics*, 2, 1-21.

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.

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 T. R. L. FRY (1991) Testing the linear regression model using Burr critical value approximations, provisionally accepted by the *Journal of Quantitative Economics.*

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.

EVANS, M. A. (1991). Robustness and size of tests of autocorrelation and heteroscedasticity, *Journal of Econometrics*, forthcoming.

HENSHAW, R. C. (1966). Testing single-equation least squares regression models for autocorrelated disturbances", *Econometrica*, 3, 646-660.

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.

KING, M. L. and G. H. HILLIER (1985). Locally best invariant tests of the error covariance matrix of the linear regression model. *Journal of the Royal Statistical Society, B*, 47, 98-102.

KOERTS, J. and A. P. J. ABRAHAMSE (1969). *On the theory and application of the general linear model*. Rotterdam University Press, Rotterdam.

RODRIGUEZ, R. N. (1977). A guide to the BURR Type XII distributions. *Biometrika*, 64, 129-134.

RODRIGUEZ, R. N. (1980). SHAPE: An interactive graphics SAS procedure for fitting histograms, presented at the Fifth Annual SAS Users Conference International Conference. Texas.

RODRIGUEZ, R. N. (1982). Burr Distributions, in S. Kotz and N. L. Johnson, eds. *Encyclopedia of Statistical Sciences,* Vol. 1, Wiley.

SROETZER, J. (1978). A class of parametric tests for heteroscedasticity in linear econometric models. *Econometrica*, 46, 1311-1327.

TADIKAMALLA, P. R. (1980) A look at the Burr and related distributions. *International Statistical Review*, 48, 337-334.

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.

TABLE 1

True size of Burr critical value approximations (a), and Burr p-value approximations of true critical values (b), for tests of AR(1) disturbances, n=40.

Nominal size		1%		5%		10%	
Test/data	$\left(\sqrt{\beta_1},\beta_2\right)$	(a)	(b)	(a)	(b)	(a)	(b)
d_1							
spirits	(-.03,2.86)	1.010	.989	4.803	5.202	9.774	10.225
trend	(-.01,2.86)	1.012	.987	4.800	5.205	9.767	10.232
CPI	(-.02,2.87)	1.011	.989	4.800	5.205	9.769	10.229
liquidity	(-.01,2.86)	1.012	.987	4.801	5.204	9.767	10.231
Watson's X	(0,2.89)	1.014	.985	4.797	5.208	9.758	10.239
d_1'							
spirits	(-.04,2.87)	1.007	.920	4.804	5.201	9.778	10.221
trend	(-.03,2.87)	1.010	.990	4.802	5.203	9.772	10.226
CPI	(-.04,2.87)	1.008	.991	4.802	5.203	9.774	10.226
liquidity	(-.01,2.90)	1.011	.988	4.798	5.207	9.763	10.236
Watson's X	(-.03,2.87)	1.009	.990	4.804	5.201	9.775	10.223
$s_1(.5)$							
spirits	(-.03,2.87)	1.009	.990	4.803	5.202	9.773	10.224
trend	(-.02,2.87)	1.011	.988	4.801	5.204	9.769	10.226
CPI	(-.01,2.86)	1.012	.987	4.799	5.206	9.765	10.233
liquidity	(-.01,2.86)	1.014	.985	4.801	5.204	9.766	10.232
Watson's X	(.04,2.86)	1.022	.975	4.793	5.213	9.746	10.253
g_1							
spirits	(-.02,2.86)	1.101	.987	4.803	5.202	9.767	10.229
trend	(-.01,2.86)	1.010	.987	4.800	5.205	9.767	10.232
CPI	(-.01,2.86)	1.014	.985	4.799	5.206	9.763	10.235
liquidity	(.01,2.86)	1.017	.982	4.799	5.207	9.758	10.241
Watson's X	(.03,2.86)	1.205	.978	4.795	5.211	9.749	10.250

TABLE 2

True size of Burr critical value approximations (a), and Burr p-value approximations of true critical values (b), for tests of AR(4) disturbances, n=40.

Nominal size		1%		5%		10%	
Test/data	$\left(\sqrt{\beta_1},\beta_2\right)$	(a)	(b)	(a)	(b)	(a)	(b)
d_4							
CPI	(.03,2.86)	1.010	.978	4.796	5.210	9.750	10.248
seasonals	(0,2.86)	1.015	.984	4.799	5.207	9.761	10.237
CPI/seas	(-.02,2.86)	1.012	.987	4.800	5.205	9.766	10.232
liquidity	(.03,2.86)	1.020	.978	4.797	5.209	9.752	10.248
Watsons's X	(.04,2.88)	1.021	.978	4.794	5.212	9.744	10.254
d'_4							
CPI	(-.05,2.87)	1.008	.992	4.802	5.202	9.774	10.225
seasonals	(-.03,2.88)	1.007	.993	4.806	5.199	9.780	10.218
CPI/seas	(-.06,2.88)	1.004	.996	4.808	5.196	9.786	10.212
liquidity	(-.02,2.87)	1.010	.989	4.802	5.203	9.771	10.228
Watsons's X	(-.01,2.90)	1.011	.988	4.797	5.208	9.762	10.236
$s_4(.5)$							
CPI	(-.01,2.87)	1.018	.981	4.794	5.212	9.751	10.247
seasonals	(.02,2.87)	1.013	.986	4.800	5.205	9.765	10.233
CPI/seas	(-.02,2.88)	1.012	.987	4.800	5.205	9.767	10.232
liquidity	(.02,2.89)	1.020	.979	4.796	5.210	9.751	10.247
Watson's X	(.02,2.89)	1.028	.969	4.789	5.219	9.731	10.267
g_4							
CPI	(.01,2.86)	1.017	.982	4.795	5.211	9.754	10.244
seasonals	(0,2.86)	1.015	.984	4.799	5.207	9.761	10.237
CPI/seas	(-.01,2.86)	1.014	.985	4.799	5.206	9.764	10.234
liquidity	(.02,2.87)	1.019	.980	4.797	5.208	9.754	10.244
Watson's X	(.05,2.87)	1.024	.974	4.792	5.214	9.741	10.258

TABLE 3

True size of Burr critical value approximations (a), and Burr p-value approximations of true critical values (b), for tests of heteroscedasticity, n=40.

Nominal size		1%		5%		10%	
Test/data	$\left(\sqrt{\beta_1},\beta_2\right)$	(a)	(b)	(a)	(b)	(a)	(b)
SZ							
trend	(0,2.87)	1.015	.984	4.798	5.207	9.761	10.823
lognormal	(.01,2.87)	1.016	.983	4.798	5.208	9.758	10.241
households	(-.01,2.87)	1.012	.987	4.800	5.205	9.765	10.233
liquidity	(.02,2.87)	1.018	.981	4.780	5.208	9.756	10.242
uniform	(0,2.86)	1.014	.985	4.799	5.206	9.762	10.235
s_a							
trend	(0,2.94)	1.009	.990	4.788	5.217	9.752	10.247
lognormal	(.01,2.94)	1.010	.989	4.792	5.216	9.755	10.244
households	(-.01,2.94)	1.008	.992	4.789	5.218	9.756	10.242
liquidity	(.01,2.93)	1.011	.989	4.791	5.214	9.755	10.245
uniform	(0,2.94)	1.009	.990	4.790	5.216	9.755	10.242
$s_a(2.5)$							
trend	(.39,3.21)	1.129	.857	4.757	5.260	9.574	10.432
lognormal	(.39,3.21)	1.130	.855	4.758	5.258	9.572	10.437
households	(.38,3.20)	1.125	.860	4.755	5.201	9.576	10.430
liquidity	(.38,3.20)	1.126	.859	4.760	5.254	9.579	10.424
uniform	(.39,3.20)	1.129	.972	4.760	5.183	9.577	10.219
LBI							
households	(-.80,3.93)	.950	1.058	4.600	5.415	9.632	10.353
trend	(0,2.94)	1.009	.991	4.788	5.217	9.751	10.245
lognormal	(.59,3.71)	1.250	.716	4.717	5.315	9.359	10.664
liquidity	(.18,3.31)	1.009	.990	4.696	5.317	9.628	10.371
uniform	(-.14,2.99)	.987	1.014	4.815	5.183	9.819	10.219
s(5.0)							
trend	(.58,3.54)	1.238	.735	4.786	5.235	9.498	10.519
lognormal	(1.36,6.17)	2.533	0	5.531	4.299	9.488	10.601
households	(-.28,3.00)	.981	1.020	4.854	5.147	9.877	10.122
liquidity	(1.10,5.46)	1.884	.151	5.130	4.833	9.323	10.777
uniform	(.45,3.34)	1.147	.990	4.763	5.216	9.558	10.242
B&P							
trend	(0,2.94)	1.015	.980	4.967	5.043	9.952	10.057
households	(-.80,3.93)	1.517	.952	5.336	4.355	10.005	10.008
lognormal	(-.59,3.71)	.951	1.046	4.938	5.073	9.991	10.004
liquidity	(-.18,3.31)	.967	1.024	4.911	5.017	9.948	9.968
uniform	(.14,2.99)	1.062	.918	4.991	4.999	9.929	10.083

TABLE 4

True size of four-moment beta $\beta(4)$ and Burr critical value approximations for s(5.0) and Breusch-Pagan heteroscedasticity tests

Nominal size		1%		5%		10%	
Approximation		$\beta(4)$	Burr	$\beta(4)$	Burr	$\beta(4)$	Burr
Data	$\left(\sqrt{\beta_1}, \beta_2\right)$						
		s(5.0) point optimal test					
n=24							
trend	(.72, 3.78)	.958	1.399	4.974	4.849	9.859	9.439
lognormal	(1.31, 5.61)	-	2.521	-	5.466	-	9.479
households	(-.17, 2.78)	.997	1.046	4.983	4.904	9.987	9.885
liquidity	(1.12, 5.29)	-	2.146	-	5.310	-	9.360
uniform	(.65, 3.63)	-	1.361	-	4.813	-	9.394
n=64							
trend	(.47, 3.36)	-	1.156	-	4.758	-	9.542
lognormal	(1.53, 7.19)	-	3.016	-	5.891		9.605
households	(-.23, 2.99)	1.001	.980	5.003	4.833	10.005	9.850
liquidity	(-.07, 3.30)	-	.979	-	4.863	-	9.883
uniform	(.35, 3.21)	-	1.093	-	4.763	-	9.614
		Breusch-Pagan (B&P) test					
n=24							
trend	(0, 2.89)	1.002	1.017	5.002	4.974	10.000	9.958
lognormal	(-.67, 3.65)	1.023	.738	5.091	4.939	10.059	10.464
households	(.73, 3.66)	1.123	1.508	5.003	5.369	9.868	10.059
liquidity	(-.42, 3.46)	-	.932	-	4.938	-	10.022
uniform	(.13, 2.98)	1.004	1.060	4.984	4.984	9.983	9.922
n=64							
trend	(0, 2.96)	1.000	1.014	5.002	4.964	9.998	9.950
lognormal	(-.48, 3.57)	-	.926	-	4.920	-	10.006
households	(.64, 3.56)	1.039	1.345	4.973	5.200	9.872	9.946
liquidity	(.79, 4.20)	-	1.463	-	5.256	-	9.895
uniform	(.12, 2.99)	1.000	1.049	5.003	4.984	10.003	9.934

- indicates that the approximation was not obtainable.