MONASH

20/90

MONASH

UNIVERSITY

AUSTRALIA

TESTING THE LINEAR REGRESSION MODEL USING BURR CRITICAL VALUE AND p-VALUE APPROXIMATIONS

Merran A. Evans and Tim R.L. Fry

Working Paper No. 20/90

December 1990

DEPARTMENT OF ECONOMETRICS

ISSN 1032-3813

ISBN 0 86746 964 1

TESTING THE LINEAR REGRESSION MODEL USING BURR CRITICAL VALUE AND p-VALUE APPROXIMATIONS

Merran A. Evans and Tim R.L. Fry

Working Paper No. 20/90

December 1990

DEPARTMENT OF ECONOMETRICS, FACULTY OF ECONOMICS COMMERCE & MANAGEMENT

MONASH UNIVERSITY, CLAYTON, VICTORIA 3168, AUSTRALIA.

Testing the Linear Regression Model
Using Burr Critical Value and p-Value Approximations

Merran A. Evans and Tim R.L. Fry*

Department of Econometrics, Monash University, Clayton, 3168, Australia.

Abstract:

The accuracy of Burr approximations of critical values and *p*-values is evaluated for the usual tests of the general linear regression model. These include tests for coefficients and autocorrelation and heteroscedasticity, based both on standard distributions and those for which the distribution is unknown. The results suggest that the Burr approximations are reasonably accurate and should prove useful both in applied research and in teaching.

KEY WORDS: Burr distributions; approximations; *p*-values; critical values; hypothesis testing; regression.

* We are most grateful to Simon Peters for the Burr computer algorithm and for the support of the Monash Research Fund Grant 19.097.005.

1. INTRODUCTION

Frequently in econometrics one may wish to test a hypothesis or construct a confidence interval, but the distribution of the test statistic is either intractable or unknown, or the use of true critical values may not be feasible. This paper aims to explore the efficacy of using the Burr (1942) family of distributions to approximate the distribution of a variety of test statistics used in economic modelling based on the linear regression model.

Hypothesis tests play a crucial role in econometric modelling as diagnostic tests to help determine the validity of econometric models, economists being unable to perform experiments as verification. A statistical hypothesis test involves determining whether the value of a test statistic, based on the given data, is consistent with the probability distribution specified by the proposed hypothesis. A hypothesis is rejected if the test statistic lies in the extremes of this distribution, beyond a critical value, or if its associated p-value is very small. In econometric modelling it is important that hypothesis tests and associated confidence intervals be as accurate as possible, given the serious consequences of misspecification, which may include inefficient parameter estimates, misleading statistical inferences and poor forecasts.

Unfortunately, in applied economic analysis, the true probability distribution of a test statistic or estimator is often unknown . However, in practice the shape of any distribution can be described reasonably adequately by certain characteristics which are usually known or can be determined. These are the first four moments: the mean (location), variance (spread), skewness and kurtosis (thickness of tails). It is possible to use this information to approximate the unknown distribution with one selected from the Burr family. By matching these four moments, the Burr distribution which most closely approximates the distribution of a given statistic is identified.

Burr distributions cover a variety of different 'shapes' and have the ability to model distributions with a wide range of moment coverage (Fry(1988,1989)). Two particular distributions, Type XII and its 'reciprocal' Type III, have distribution functions and associated inverses in a simple closed mathematical form. Consequently, $p-$values for any calculated value of a test statistic can easily be generated from the distribution function, and critical values from its inverse. This explicit analytic form is not available for traditional moment matching methods using the Pearson and Johnson families of distributions, so gives the Burr approach a considerable computational advantage. All such approximation methods require computation of moments, but only with Burr distributions is numerical integration avoided in determining the distribution function and its inverse.

This study evaluates the usefulness in practice of the Burr approximations to distributions of a range of test statistics likely to be employed in analysis of the standard general linear regression model. In the next section the methodology is described. In section 3 the accuracy of Burr approximations is examined for the standard normal, χ^2, Students' t and F distributions, as well as for tests of autocorrelated and heteroscedastic disturbances for which the true distribution of the test statistic is unknown but true critical values can be determined. Section 4 presents practical examples in applied economics.

2. THE BURR XII AND III DISTRIBUTIONS

The cumulative distribution function of members of the Burr XII family of distributions is summarised by two parameters c and k, and is given by

$$F(x) = 1 - (1 + x^c)^{-k}, \qquad (x > 0, k > 0). \tag{1}$$

With the following existence conditions, $ck > r$ and $c > 0$, the moments about the origin are given by

$$\mu_r' = kB(1 + \frac{r}{c}, k - \frac{r}{c}) = \frac{\Gamma(1 + r/c)\Gamma(k - r/c)}{\Gamma(k)}. \tag{2}$$

The distributions are unimodal if $c > 1$ and L-shaped if $0 < c \leq 1$. The distribution function of members of the 'reciprocal' Burr III family is

$$F(x) = (1 + x^{-c})^{-k}, \qquad (x > 0, k > 0) \tag{3}$$

with its moments obtained from the expression for the Burr XII moments (2), by replacing c by $-c$, and hence the existence conditions are $c < -r$. These distributions are unimodal for $ck > 1$ and twisted L-shaped if $ck \leq 1$.

The existence regions of these two distributions are given in Rodriguez (1977, 1982) in terms of skewness and kurtosis. They cover an extensive range, which includes that of many standard distributions. Determining which of the family of Burr distributions most closely approximates the distribution of the test statistics of interest involves matching the first four moments. For a test statistic, t, with mean (μ_t), variance (σ_t^2), skewness $(\sqrt{\beta}_1 = \sqrt{\mu_3^2/\mu_2^3})$ and kurtosis $(\beta_2 = \mu_4/\mu_2^2)$ (where μ_r is the rth moment about the mean) coefficients should be determined. Its distribution can be approximated by finding values of k and c so that the skewness and kurtosis coefficients of the Burr distribution are equal to the corresponding quantities of the distribution of t.

Selecting the appropriate Burr distribution for any given $\sqrt{\beta}_1, \beta_2$, involves solving two non-linear simultaneous equations for the parameters c and k which characterise the Burr distribution (see Rodriguez (1982)).

$$\sqrt{\beta}_1 = \frac{\Gamma^2(k)\lambda_3 - 3\Gamma(k)\lambda_2\lambda_1 + 2\lambda_1^3}{[\Gamma(k)\lambda_2 - \lambda_1^2]^{3/2}},$$

$$\beta_2 = \frac{\Gamma^3(k)\lambda_4 - 4\Gamma^2(k)\lambda_3\lambda_1 + 6\Gamma(k)\lambda_2\lambda_1^2 - 3\lambda_1^4}{[\Gamma(k)\lambda_2 - \lambda_1^2]^2},$$

where $\lambda_r = \mu_r'\Gamma(k) = \Gamma(1 + r/c)\Gamma(k - r/c)$, $(r = 1, 2, 3, 4)$. Solutions c and k of these equations can be obtained approximately using the tables given in Burr (1973) by taking $\sqrt{\beta}_1, \beta_2$ nearest to those of the given test statistic. These published tables do not include negative skewness values, so for accuracy and completeness, we used an algorithm to obtain the solutions here. Note that Rodriguez (1980) has developed SHAPE, an interactive algorithm for use with the SAS package to fit Burr III (and hence also Burr XII) distributions.

These computed c and k identify which Burr distribution to use as an approximation to that of the given test statistic, t, and its mean (μ_B) and variance (σ_B^2) are then determined. Matching the first and second moments involves equating the standardised variates and solving:

$$\frac{x - \mu_B}{\sigma_B} = \frac{t - \mu_t}{\sigma_t}.$$

$$\Rightarrow \qquad t = \frac{\sigma_t}{\sigma_B}(x - \mu_B) + \mu_t = Sx + (\mu_t - S\mu_B) = Sx + L.$$

Hence by relocation ($L = \mu_t - S\mu_B$) and scaling ($S = \sigma_t/\sigma_B$), the two distributions with corresponding skewness and kurtosis are matched. Effectively this involves 'standardising' the distributions to have a zero mean and unit standard deviation, and matching the two shape parameters of skewness and kurtosis. This approach was used by Burr (1967) and Tadikamalla and Ramberg (1975) as an approximate method to generate normal and Gamma variates, respectively.

For tests with rejection regions in the left hand tail, the Burr $p-$value can be computed from the approximating Burr XII cumulative distribution function:

$$P(t \leq t_{crit}) = P(x \leq x_{crit}) = 1 - [1 + x_{crit}^c]^{-k} = 1 - [1 + (\frac{t_{crit} - L}{S})^c]^{-k}.$$

The inverse of this distribution function is easily obtained and the approximated Burr critical value for a rejection region of nominal significance level α in the lower tail of the distribution is given by:

$$t_{crit}^{appr} = L + Sx_{crit} = L + S[(1 - \alpha)^{-1/k} - 1]^{1/c}.$$

If $c < 0$, then using $x* = 1/x$ implies that the rejection region is in the right-hand tail of the distribution, with calculations changing appropriately. In this case the null hypothesis is rejected for large values of t or small values of $1/t$, which is the appropriate Burr III distribution if $c < 0$.

To guage the accuracy of this Burr approximating procedure, in the next section the true sizes of tests using these approximated Burr critical values, as well as the Burr $p-$values for tests involving the true critical values, were computed and compared with nominal sizes α = .01, .05 and .10. The criterion for accuracy was whether the computed values lie within 1%, 5% or 10% of the true value, i.e. within $\pm$ (.0001, .0005, .0010) of α = .01, within $\pm$ (.0005, .0025, .0050) of α = .05 and within $\pm$ (.0010, .0050, .0100) of α = .10. This was done for a variety of tests used in analysis of the general linear regression model,

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

where y is $n \times 1$, X is $n \times k$, β is $k \times 1$, and u is $n \times 1$. The matrix X is comprised of non-stochastic regressors, and the unknown disturbances u are assumed to be normally distributed. Note that this is **not** a simulation study as exact sizes are computed.

3. TEST STATISTICS USED IN REGRESSION ANALYSIS

3.1 TESTS WITH WITH KNOWN STANDARD DISTRIBUTIONS

Our first comparisons of the Burr approximations were with four distributions in common use in hypothesis testing situations, either through exact or asymptotic results, namely the standard normal, the Students' t, chi-squared (χ^2) and Fisher's F distribution.

Selected results are given in tables 1 to 3, but full results are available on request from the authors. Tables 1 and 2 give (a) the true test size when using the Burr approximation to the critical value and (b) the calculated Burr p-value of the true critical value for nominal sizes $\alpha = .01, .025, .05$ and $.10$. Table 3 presents corresponding results for the F distribution, but only for $\alpha = .05$ for reasons of space. These 'true' values were determined from double precision IMSL routines: DNORDF, DNORIN for the normal distribution; DTDF, DTIN for the t distribution; DCHIDF, DCHIIN for the χ^2 distribution; and DFDF, DFIN for the $F-$ distribution.

Both the $t(\nu)$ (with ν degrees of freedom) and the standard normal are symmetric distributions with first four moments $(\mu, \sigma, \sqrt{\beta}_1, \beta_2)$ given by $(0, \sqrt{\nu/(\nu-2)}, 0, 3(\nu-2)/(\nu-4))$ and (0, 1, 0, 3) respectively. In our approximations we exploited this symmetry, as suggested by Burr (1967), by averaging $F(x)$ and $1 - F(x)$. This improved the approximations marginally (results without averaging are available on request from the authors). Note that the moment existence condition ($\nu > r$ for the rth moment to exist) means that only $t(\nu)$ distributions with $\nu > 5$ can be approximated with methods that involve the first four moments. Further, distributions with $\nu \leq 8$ have skewness-kurtosis values outside the Burr existence regions. Hence our methodology can only be applied for $\nu=$ 9 to 120. The $t(9)$ distribution with $(\sqrt{\beta}_1, \beta_2) =$ (0,4.2) corresponds to the logistic distribution with scale parameter $\sigma = \sqrt{9/7}$, and also to the Burr distribution with $k = 1$

and $c \to \infty$ on the boundary of the existence region. Exact analytic results therefore can be obtained for $t(9)$ using the logistic distribution: exact p-values from the distribution function $F(x) = 1 + exp(-x/\theta)$, and p-values from its inverse $x_{crit} = \theta log[\alpha/(1-\alpha)]$, where $\theta = \sqrt{3}\sigma/\pi$. Table 1 shows that our approximating procedure, using averaging, gives results for the t and standard normal distributions which are very close to those for the 'true' case (typically within 5% and often within 1%). Before averaging, the true test size when using the Burr approximation to the true critical value was slightly below the nominal size in the lower tail, and above in the upper tail, whereas the converse applied for the Burr $p-$value approximations. This suggests a small shift in location between the true and approximating distributions. Any systematic difference in the shape of the tails is obliterated by averaging: the upper tail was generally closer to the true value, and hence the averaged values in both columns (a) and (b) are below the nominal size. Therefore any inference based upon the approximations is likely to be 'conservative'. The exception to this was for the $\alpha = .10$ significance level for the fatter tailed $t(\nu), \nu \leq 20$ distributions, where the opposite result applied.

The chi-squared distribution is not symmetric so averages are not used and the upper tail is generally of interest. For a $\chi^2(\nu)$ distribution, the skewness coefficient is $\sqrt{\beta}_1 = \sqrt{8/\nu}$ and the kurtosis coefficient is $\beta_2 = 3 + 12/\nu$. Approximations for $\chi^2(\nu)$ distributions with degrees of freedom $\nu = 1{,}120$ were examined. For all values of $\nu > 2$, the approximations were good, being typically within 5% of the 'true' values. The $\chi^2(2)$ distribution can be shown to correspond to that of an exponential variate with scale parameter 2 and skewness-kurtosis values of (2,9), as well the Burr distribution with $c = 1$ and $k \to \infty$, which is an extreme boundary point of the existence region. Hence exact p-values can be obtained from the exponential distribution function $F(x) = 1 - exp(-x/2)$, and critical values from its inverse, $x_{crit} = 2log[1/(1-\alpha)]$. For the more extreme $\chi^2(1)$, with $(\sqrt{\beta}_1, \beta_2) = (2.83, 15)$, the approximations were less accurate, usually only within 20% of the true value. Table 2 suggests that generally for the χ^2 distribution the approximating distribution has slightly thinner upper tails than the 'true' distribution as column (a) was typically above the nominal size and column (b) below it. The only exception was for the $\alpha =.10$ significance point with degrees of freedom $\nu \leq 10$, where the opposite consistently occurred, suggesting a cross-over of the true and approximating distributions between the $\alpha = .05$ and $.10$ significance points.

For Fishers' $F(\nu, \omega)$ distribution, the skewness coefficient is given by

$$\sqrt{\beta}_1 = \frac{(2\nu + \omega - 2)((8(\omega - 4))^{1/2}}{(\omega - 6)(\nu + \omega - 2)^{1/2}} \quad (\omega > 6).$$

The kurtosis coefficient (for $\omega > 8$) can be determined using the rth moments about the origin,

$$\mu'_r = \frac{(\omega/\nu)^r \Gamma(\nu/2 + r)\Gamma(\omega/2 - r)}{\Gamma(\nu/2)\Gamma(\omega/2)} \qquad \omega > 2r.$$

Existence conditions for the first four moments require $\nu \geq 1, \omega > 8$. Approximations to the upper critical points of this (non-symmetric) variance ratio distribution were determined for values of the degrees of freedom parameters $\nu = 1, 120$ and $\omega = 9{,}120$. For most degrees of freedom values, the approximations of critical values of the Fisher's F distribution were generally within 5% of the true value, and often within 1%, as can be seen in table 3. The true test size using the Burr critical value was generally slightly above the nominal size, and the Burr p-value slightly below it. Analytic results are readily obtained for $\nu = 2$, by equating the moments μ'_r of the Burr and F-distributions, respectively. These solutions are $c = 1, k = \omega/2$, and the Burr results correspond exactly to the true values for the $F(2, \omega)$ distribution. Analytic solutions for other distributions are being investigated.

3.2 TESTS OF DISTURBANCES WITH UNKNOWN DISTRIBUTIONS

The accuracy of normal, two- and four-moment beta approximations to critical values of a range of tests of disturbance behaviour has been explored by Henshaw (1966) and Evans and King (1985). Given the definite computational advantage of the Burr approximations, it is of interest to determine whether they are competitive with the highly accurate four-moment beta approximations. This is being explored in Evans (1990) for a wide range of such tests. Burr approximations are examined here, however, for tests which appear to be the most commonly used in practice against autocorrelation and heteroscedasticity. Against first-order autoregressive (AR(1)) disturbances the first-order Durbin and Watson

(1950) test was used, and against simple fourth-order autoregressive (AR(4)) disturbances its fourth-order analogue by Wallis (1972).

For each of these tests, critical values cannot be tabulated as they depend on the actual regressors used, so bounds are required, and these involve an inconclusive region, which can be large for small samples. However, true critical values can be determined, even though the true distribution is unknown. This is a consequence of being able to express each of these one-sided tests as a ratio of quadratic forms in normal variables:

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

where A is the first- and fourth-order differencing matrix for the first- and fourth-order tests, respectively, $\hat{u} = (I - X(X'X)^{-1}X')y = My = Mu$ is the OLS residual vector and $M = I - X(X'X)^{-1}X'$. From Durbin and Watson's (1950) lemma, under normality, the true critical value for a specified significance level α or, alternatively, the actual size for any specified critical value t_{crit}, can be obtained from

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

The null hypothesis of independence is rejected for small values of the test statistic for positive, and large values for negative, autocorrelation.

Two standard popular tests for heteroscedasticity are those of Goldfeld and Quandt and of Breusch and Pagan (1979). The former has an F-distribution, and the latter is asymptotically χ^2, both distributions which were covered in the preceding section. Although the true distribution of the Breusch and Pagan test statistic is unknown, true critical values and sizes can be determined in a similar fashion to those above, and compared with the asymptotic values (which are known to be suspect), as well as the Burr approximations. The two-sided Breusch and Pagan (1979,p1290) Lagrangian multiplier test statistic, $t^2 = (\hat{u}'A\hat{u}/\hat{u}'\hat{u})^2 = (u'MAMu/u'Mu)^2$, has the matrix A in diagonal form 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 additive form, $var(u_i) = \sigma_i^2 = \sigma^2 f(1 + \lambda z_i)$, (where f is an unknown monotonically increasing non-negative function and z_i is a non-stochastic variable). The true sizes and critical values can be determined from

$$Pr((\hat{u}'A\hat{u}/\hat{u}'\hat{u})^2 > t^2_{crit}) = Pr(\hat{u}'A\hat{u}/\hat{u}'\hat{u} > t_{crit}) + Pr(\hat{u}'A\hat{u}/\hat{u}'\hat{u} < -t_{crit}) = \alpha.$$

Both true critical values and true sizes of the Burr approximations to the critical values were computed for each of these tests, assuming normally distributed disturbances, using the Imhof procedure in 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 first four moment characteristics for each test were obtained using the methods of Evans and King (1985) which involve traces of products of the matrices MAM.

Several actual and artificial X matrices were chosen to reflect a range of behaviour characteristic of economic variables, and all have been used previously in experiments concerning tests of autocorrelation and heteroscedasticity. All include a constant intercept term. The real regressors include: the annual spirit income and price data of Durbin and Watson (1951); the mildly 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. Artificial regressors represent alternative characteristic behaviour, such as a time trend for slowly evolving non-seasonal economic time series, observations with a lognormal distribution for skewed cross-sectional data, and a set of 0-1 dummies for quarterly additive seasonal behaviour. Uniformly distributed data is common is such experiments. Small, moderate and relatively large sample sizes (n) were examined for each set of regressors, with n = 20,40,60 for the autocorrelation tests, and n = 24,40,64 for the Breusch and Pagan heteroscedasticity test.

Selected results on the approximations for the one-sided $AR(1)$ and $AR(4)$ tests are given in table 4. For several data sets and for nominal sizes α = .01,.05, .10, these show: (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. Corresponding results for the two-sided Breusch and Pagan test are presented in table 5, as well as the true size of the test when the standard χ_1^2 asymptotic value is used. (The deflator z_t was assumed in this study to correspond to the first non-constant regressor in each data set).

For the autocorrelation tests, the true size of the Burr approximation always 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 density function curves near the tails between $\alpha = .01$ and .05. The converse 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. All these tests were characterised by near symmetry with a range of $\sqrt{\beta}_1 \in$ (-.03, .04) and short tails with $\beta_2 \in$(2.86,2.89). With the exception of the CPI data for $AR(4)$ with small samples and the extreme tail ($n = 20, \alpha = .01$), all sizes and p-values were within at least 5% of the nominal sizes.

For the two-sided Breusch and Pagan test shown in table 5, the skewness and kurtosis of each data set is also shown, as more variation was found than for the autocorrelation tests. The two tails of the distribution are used, such that approximations for $\alpha = .01$ involve summing values for $\alpha = 0.005$ and .995, reaching further into the tails. The few cases where the Burr approximation was poor occurred this extreme situation, and for the liquidity data with $n = 64$ the p-value could not be obtained within our limits of computation. With $\alpha = .01$ and small samples (n=24), the true sizes of tests with the Burr approximated critical value and the p-value approximations were sometimes within 10% of the nominal size (and within 5% for the trend data). With larger samples they were usually within 10% and often within 5%. However, for larger values of α, the approximations improved: for $\alpha = .05$, they always were within 10%, usually within 5%, and often within 1% of the nominal size; and for $\alpha = .10$, they were usually within 1% and always within 5% (with the exception of the highly skewed lognormal data for $n = 24$).

The usefulness in applied work of the Burr approximation is highlighted by a comparison of the Breusch and Pagan results with those obtained using the conventional χ^2 asymptotic critical value also shown in table 5, for all sample sizes. The test using this asymptotic critical value is well known to underestimate the true size, which is confirmed by our results. None of the values were within even 10% of the nominal size for $n = 24$. Only for test statistics with moderate skewness and kurtosis (for the trend and uniform data) were they within 10% of the nominal value, at $\alpha = .10$ for $n = 40$ and $\alpha \geq .05$ for $n = 64$. Given this asymptotic value is the standard recommended procedure in textbooks and econometric software, the Burr approximation technique has obvious practical value.

4. PRACTICAL EXAMPLES

To illustrate the potential application of our methodology, consider the examples in chapter 9 of the econometrics text book by Judge, Griffiths, Hill, Lutkepohl and Lee (1988). These concern autocorrelation and heteroscedasticity in the general linear model and use a common data matrix with 20 observations on a constant and two regressors x_2 and x_3. The data can be found in tables 9.1 and 9.2 in the textbook and results using SHAZAM in White, Haun and Gow (1988).

In the heteroscedasticity example (Table 9.1), an OLS estimation is carried out with the following results (with calculated t-statistics given in brackets below the estimated coefficients):

$$\underset{}{\hat{y}_t =} \underset{(-0.110)}{-0.991} + \underset{(3.404)}{1.651}\ x_{2t} + \underset{(2.431)}{0.997}\ x_{3t}, \qquad F_{calc} = 18.44.$$

The second example, focussing on autocorrelation (Table 9.2), has OLS results:

$$\hat{y}_t = \underset{(.861)}{3.842} + \underset{(7.556)}{1.811}\ x_{2t} + \underset{(3.132)}{0.634}\ x_{3t}, \qquad F_{calc} = 63.66.$$

Note that critical values and sizes are independent of the data for the t and F tests, and depend only on the X matrix for the disturbance behaviour tests, so are common to each example. (True values are also calculated as a benchmark).

The two-sided 5% critical value for the t distribution with 17 degrees of freedom is 2.110 and the corresponding Burr approximation is 2.119. The F-test on the regression has (2,17) degrees of freedom and both the tabulated and Burr 5% critical values are 3.59. Thus identical conclusions will be drawn from each approach, but our methodology can also provide p-values for the calculated t and F statistics. These are: 0.5480, 0.0019, 0.0131 and 0.0000 respectively, for the first example, and 0.1991, 0.0000, 0.0033 and 0.0000 for the second.

In the heteroscedasticity example, it is postulated that the disturbance variance depends upon x_2, and the Breusch and Pagan test yields a value of 4.03. With the standard procedure of comparing this with the asymptotic χ^2 critical value 3.84 at the 5% significance level, the null hypothesis of homoscedasticity would be rejected. As this asymptotic value may not be appropriate with 20 observations, we determine the relevant Burr critical and p-values. For this data the test statistic has skewness and kurtosis values $(\sqrt{\beta}_1, \beta_2)$ of 0.095 and 3.102, respectively. The selected approximating Burr distribution gives a 5% critical value of 3.47 (compared to the true 3.50) and a p-value for the calculated statistic of 0.0354 (compared to the true size of 0.0359). Hence we would now reject the null hypothesis at the 5% level with more confidence, but accept it at the 1% level (which has a true critical value of 6.14 and an asymptotic value of 6.64).

Initial interest usually is in testing for an AR(1) process. The calculated value for the Durbin-Watson test statistic is 0.91 in the second autocorrelation example. Tabulated bounds for critical values of this test statistic are $\{d_L = 1.100, d_U = 1.537\}$ at the 5% significance level, and $\{d_L = 0.863, d_U = 1.271\}$ at the 1% level. Hence at the 5% level, we would reject the null hypothesis of no autocorrelation, but at the 1% level the test is inconclusive. For this data, the skewness and kurtosis of the Durbin-Watson statistic are -.057 and 2.756, respectively. The approximating Burr distribution gives the 5% critical value as 1.425 and the 1% value as 1.167, (which compare well with the exact critical values of 1.431 and 1.166). The Burr p-value for the calculated statistic corresponds to the true value of .001. Thus we would now reject the null hypothesis of no first order autocorrelation at both levels.

If we were to use this example to test for fourth order autocorrelation (as would be done in practice with quarterly data), the calculated value for the Wallis variant of the Durbin and Watson test statistic is 2.62, indicating a test for negative correlation is appropriate, which involves the upper tail of the distribution. Tabulated bounds (see Johnson (1984)) for critical values at the 5% significance level are $\{d_{4L} = .827, d_{4U} = 1.203\}$. Unfortunately, the distribution of the test statistic is asymmetric for regression models with an intercept such as these, so we cannot compare 4 - 2.62 = 1.38 with (interpolated) values of these bounds, as would be done in the first-order case. However, bounds can be obtained from King and Giles (1977) and Giles and King (1978), and these are $\{d_{4L} = 1.977, d_{4U} = 2.673\}$ at the 5% level. The test statistic of 2.62 just falls in the inconclusive region. Our methodology readily provides further information. The skewness and kurtosis values for this test statistic are -0.057 and 2.756, respectively. The corresponding Burr upper 5% and 1% critical values are 2.427 and 2.706 (compared to the true values of 2.433 and 2.705), respectively. The Burr $p-$value of the calculated statistic 2.62 is 0.0173 (true size = .0175). Hence we would reject the null hypothesis of no fourth order autocorrelation at the 5% level, but accept it at the 1% level.

In summary, for the t, F and χ^2 tests conventional practice is to compare calculated values of the test statistic with tabulated critical values at standard significance levels. For the usual Durbin-Watson type tests of autocorrelation, the calculated values are compared with tabulated bounds for these true critical values, but the test is often inconclusive in small samples. In all these cases, our methodology approximates the true critical values well, and has the added attraction of producing quite accurate p-values for any calculated statistic. To our knowledge no p-value is produced routinely in a standard econometrics package for tests with critical regions dependent on the regressors, with the exception of the first-order Durbin-Watson test in SHAZAM. However, even this package does not automatically produces the true critical values.

A recent research example where these Burr approximations might be useful is in a recent study of specification analysis in dynamic models by Fiebig and Maasoumi (1990). For example, in their regressions (for regime 1) of interest rates on seasonal dummies and trend and also on these regressors plus the lagged dependent variable, they report Durbin-Watson statistics of 1.59 and 1.55, respectively, which lie in the 5% inconclusive region. For dynamic models, such as the latter, Inder (1986) has found that the Durbin-Watson test with the distribution approximating that of the model without the lag is superior to the $h-$ test. Adopting this approach, the set of fixed regressors and hence the approximating Burr distribution used for each of the two models is the same, the Burr 5% critical value is 1.43 and the p-values are .12 and .10, respectively. As seasonal variation is of interest, the Burr p-values for the fourth-order Wallis analogue of the Durbin-Watson test also can be computed analogously. For these two models the p-values are .06 and .116, respectively, indicating significant fourth order residual correlation in the model with trend and seasonal dummies as regressors. In contrast, this test for their model with seasonal dummies, a lagged dependent variable but no trend has a p-value of .58.

In an applied research study involving financial data by Faff, Lee and Fry (1990), Burr p-values approximations were used for a locally best invariant test for $AR(1)$ behaviour of the model coefficients. The test statistic has an unknown distribution, but is expressible as a ratio of quadratic forms in normal variables so our methodology is applicable.

5. CONCLUSIONS AND FUTURE RESEARCH

Our empirical study of approximating critical values and determining p-values of standard tests in applied regression analysis by matching the first four moments with those of distributions from the Burr family is most encouraging. It provides an incentive to extend this methodology to test statistics and estimators in a range of econometric applications. In summary, it appears that tests using Burr approximations are reasonably accurate: they generally lie within 5%, and at least within 10%, of the true size, when the skewness and kurtosis parameters fall within the Burr existence regions. A useful feature of this approach is that an initial inspection of these permissible regions in terms of skewness and kurtosis of a test statistic of interest will indicate whether the Burr approximation is likely to be successful. These regions are given in Rodriguez (1977,1982) and cover most distributions likely to be encountered in practice.

This methodology has marked computational advantages over other approximating methods and should prove useful in applied research and teaching. Critical values can be determined for tests which are often inconclusive, as well as for some where only the asymptotic critical value is known but is inappropriate in small samples characteristic of economic analysis. Further, intuitively appealing and readily comprehensible p-values can be calculated easily for all the standard test statistics.

REFERENCES

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.

– – – – – (1967), "A Useful Approximation to the Normal Distribution, with Application to Simulation," *Technometrics*, 647-651.

– – – – – (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.

– – – – – (1951), "Testing for Serial Correlation in Least Squares Regression, II," *Biometrika*, 38, 159-178.

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

Evans, M. A. (1990) "Burr Critical Value Approximations for Tests of Autocorrelation and Heteroscedasticity", Mimeo, Monash University.

Faff, R. W., J. H. H. Lee and T. R. L. Fry (1990), "Time Stationarity of Systematic Risk: Some Australian Evidence,"*Journal of Business, Finance and Accounting*, forthcoming.

Feibig, D.G. and E. Maasoumi (1990), "Specification Analysis in Dynamic Models", Journal of Business and Economic Statistics," *Journal of Business and Economic Statistics*, 8, 4, 443-451.

Fry, T. R. L. (1988), "On the Use of Some Burr Family Distributions in Econometrics," unpublished Ph. D. thesis, University of Manchester.

– – – – – (1989), "Univariate and Multivariate Burr Distributions: A Survey. " Working Paper 9/89, Econometrics, Monash University.

Giles, D.E.A. and M.L. King (1978), "Fourth-order Autocorrelation: Further Significance Points for the Wallis Test", *Journal of Econometrics*, 8, 255-259.

Henshaw, R.C. (1966), "Testing Single-Equation Least Squares Regression Models for Autocorrelated Disturbances", *Econometrica*,3, 646-660.

Inder, B. (1986), "An Approximation to the Null Distribution of the Durbin-Watson Statistic in Models Containing Lagged Dependent Variables", *Econometric Theory*, 2, 413-428.

Johnson, J. (1984), "Econometric Methods," 3rd.edition, McGraw-Hill.

Judge, G.G., W.E.Griffiths, R.C.Hill, H.Lutkepohl and T-C.Lee (1988), "Introduction to the Theory and Practice of Econometrics", Wiley.

King, M.L. and D.E.A. Giles (1977), "A Note on Wallis' Bounds Test and Negative Autocorrelation", *Econometrica*, 45, 1023-1026.

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.

– – – – – (1980), "SHAPE: An Interactive Graphics SAS Procedure for Fitting Histograms," presented at the Fifth Annual SAS Users Conference International Conference. Texas.

– – – – – (1982), "Burr Distributions" in S. Kotz & N. L. Johnson, eds. *Encyclopedia of Statistical Sciences,* Vol. 1, Wiley.

Tadikamalla, P. R. and J. S. Ramberg (1975), "An Approximate Method for Generating Gamma and Other Variates" *Journal of Statistical Computational and Simulation*, 3, 275-282.

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

White, K. J., S. A. Haun and D. J. Gow (1988), *Introduction to the Theory and Practice of Econometrics*, Second Edition: A Computer Handbook Using SHAZAM and SAS", Wiley.

Table 1

Burr Accuracy for Students' $t(\nu)$ and $N(0,1)$ distributions

Nominal size	1%		2.5%		5%		10%	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
ν								
9	1.000	1.000	2.500	2.500	5.000	5.000	10.000	10.000
10	.940	.982	2.413	2.431	4.948	4.883	10.092	9.868
11	.952	.978	2.427	2.407	4.951	4.839	10.062	9.819
12	961	.978	2.437	2.395	4.953	4.818	10.043	9.794
13	.967	.977	2.443	2.389	4.954	4.806	10.031	9.779
14	.971	.978	2.448	2.386	4.956	4.798	10.022	9.770
15	.975	.979	2.452	2.384	4.957	4.794	10.016	9.764
16	.977	.980	2.455	2.384	4.958	4.791	10.011	9.760
17	.979	.982	2.457	2.383	4.959	4.789	10.008	9.757
18	.981	.983	2.459	2.383	4.960	4.788	10.005	9.755
19	.983	.984	2.461	2.384	4.960	4.787	10.003	9.754
20	.984	.985	2.462	2.384	4.961	4.787	10.001	9.752
21	.985	.986	2.464	2.385	4.961	4.787	9.999	9.752
22	.986	.987	2.465	2.385	4.962	4.787	9.998	9.751
23	.986	.987	2.466	2.386	4.962	4.787	9.997	9.751
24	.987	.988	2.466	2.386	4.963	4.787	9.996	9.751
25	.988	.989	2.467	2.387	4.963	4.787	9.995	9.750
26	.988	.989	2.468	2.387	4.964	4.787	9.995	9.750
27	.989	.990	2.468	2.388	4.964	4.788	9.994	9.750
28	.989	.991	2.469	2.388	4.964	4.788	9.994	9.750
29	.990	.991	2.469	2.389	4.964	4.788	9.993	9.750
30	.990	.992	2.470	2.389	4.965	4.789	9.993	9.750
40	.992	.995	2.473	2.393	4.966	4.792	9.990	9.752
60	.994	.999	2.475	2.397	4.968	4.795	9.989	9.754
120	.996	1.002	2.477	2.401	4.969	4.800	9.988	9.757
N(0,1)	.997	1.006	2.479	2.405	4.971	4.805	9.987	9.760

(a) True size (%) of Burr critical value approximations

(b) Burr p-value (%) approximations using true critical values

Table 2

Burr Accuracy for the $\chi^2(\nu)$ disturbances

Nominal size	1%		2.5%		5%		10%	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
ν								
1	1.105	.877	2.235	2.870	4.075	6.147	8.687	11.169
2	1.000	1.000	2.500	2.500	5.000	5.000	10.000	10.000
3	1.011	.989	2.512	2.448	5.002	5.002	9.958	10.043
4	1.019	.981	2.525	2.475	5.008	4.992	9.948	10.053
5	1.025	.976	2.536	2.464	5.021	4.979	9.951	10.050
6	1.028	.972	2.545	2.455	5.033	4.967	9.958	10.042
7	1.031	.970	2.553	2.448	5.044	4.956	9.968	10.033
8	1.033	.969	2.558	2.442	5.053	4.946	9.977	10.023
9	1.034	.967	2.563	2.437	5.062	4.938	9.987	10.013
10	1.035	.967	2.567	2.433	5.069	4.931	9.996	10.004
11	1.035	.966	2.571	2.430	5.075	4.924	10.004	9.996
12	1.036	.966	2.574	2.428	5.081	4.918	10.012	9.988
13	1.036	.966	2.576	2.425	5.086	4.913	10.020	9.980
14	1.036	.966	2.578	2.423	5.091	4.909	10.026	9.973
15	1.036	.966	2.580	2.422	5.095	4.905	10.033	9.967
16	1.036	.966	2.581	2.420	5.099	4.901	10.039	9.961
17	1.036	.966	2.583	2.419	5.102	4.898	10.044	9.955
18	1.035	.966	2.584	2.418	5.105	4.895	10.049	9.950
19	1.035	.966	2.585	2.417	5.108	4.892	10.054	9.945
20	1.035	.967	2.586	2.416	5.111	4.889	10.058	9.941
30	1.032	.969	2.591	2.412	5.128	4.872	10.091	9.907
40	1.029	.973	2.593	2.410	5.138	4.863	10.112	9.887
50	1.027	.975	2.593	2.409	5.144	4.857	10.126	9.873
60	1.025	.976	2.593	2.409	5.148	4.853	10.136	9.863
70	1.026	.978	2.594	2.410	5.149	4.851	10.136	9.855
80	1.024	.979	2.594	2.410	5.151	4.848	10.143	9.848
90	1.023	.980	2.594	2.410	5.153	4.847	10.149	9.843
100	1.022	.978	2.593	2.410	5.155	4.847	10.154	9.845
120	1.019	.982	2.592	2.411	5.157	4.845	10.162	9.837

(a) True Size (%) of Burr critical value approximations

(b) Burr p-value (%) approximations with true critical values

Table 3

Burr Accuracy for $F(\nu,\omega)$ distribution-approximations to the 5% nominal size

ω \ ν	2	3	4	5	6	7	8	9	10	15	20
20(a)	5.000	5.029	5.030	5.026	5.020	5.015	5.010	5.006	4.952	4.990	4.984
(b)	5.000	4.972	4.971	4.975	4.981	4.986	4.990	4.994	5.048	4.973	4.966
22(a)	5.000	5.021	5.017	5.008	5.000	4.992	4.960	4.979	4.974	4.957	4.948
(b)	5.000	4.980	4.984	4.992	5.000	5.008	5.041	5.020	5.025	5.041	5.049
24(a)	5.000	5.014	5.005	4.963	4.964	4.971	4.963	4.955	4.949	4.928	4.917
(b)	5.000	4.987	4.995	5.038	5.038	5.038	5.036	5.043	5.049	5.069	5.079
26(a)	5.000	5.008	4.995	4.980	4.966	4.954	4.943	4.935	4.927	4.902	4.993
(b)	5.000	4.992	5.005	5.019	5.033	5.044	5.054	5.063	5.070	5.094	5.007
28(a)	5.000	5.003	4.396	4.968	4.952	4.938	4.926	4.916	4.908	4.996	5.002
(b)	5.000	4.998	5.031	5.031	5.047	5.060	5.071	5.080	5.089	5.005	4.999
30(a)	5.000	4.998	4.479	4.958	4.978	4.981	4.911	4.900	4.991	4.906	5.009
(b)	5.000	5.002	5.021	5.040	5.024	5.018	5.085	5.097	5.009	5.119	4.991
35(a)	5.000	4.989	4.963	4.937	4.986	4.991	4.995	4.999	5.003	5.010	5.069
(b)	5.000	5.010	5.036	5.061	5.015	5.009	5.005	5.000	4.997	4.987	4.916
40(a)	5.000	4.983	4.951	4.987	4.992	4.998	5.003	5.007	4.969	5.027	5.035
(b)	5.000	5.017	5.048	5.013	5.008	5.003	4.997	4.992	5.040	4.973	4.964
45(a)	5.000	4.977	4.942	4.991	4.997	5.003	4.938	5.014	5.018	5.035	5.044
(b)	5.000	5.022	5.057	5.009	5.003	4.997	5.078	4.986	4.982	4.966	4.955
50(a)	5.000	4.973	4.988	4.994	5.001	4.920	5.013	5.019	5.024	5.199	5.051
(b)	5.000	5.027	5.012	5.006	5.000	5.102	4.986	4.980	4.977	4.765	4.948
55(a)	5.000	4.990	4.990	4.996	5.004	5.011	5.014	5.064	5.105	5.046	5.057
(b)	5.000	5.010	5.010	5.004	5.003	4.990	4.984	4.921	4.871	4.953	4.943
60(a)	5.000	4.966	4.991	4.998	5.006	4.980	5.052	5.013	5.032	5.051	5.062
(b)	5.000	5.033	5.009	5.002	4.994	5.020	4.949	4.987	4.968	4.950	4.937

(a) True size (%) of Burr critical value approximations

(b) Burr p-value (%) approximations with the true critical value

Table 4

Burr Accuracy for Autocorrelation Tests

Nominal Size / Data	1% (a)	1% (b)	5% (a)	5% (b)	10% (a)	10% (b)
AR(1) Durbin Watson test						
n = 20						
spirit	1.011	.989	4.834	5.170	9.811	10.187
trend	1.018	.981	4.823	5.182	9.791	10.280
CPI	1.013	.986	4.808	5.197	9.782	10.216
liquidity	1.021	.971	4.788	5.218	9.750	10.247
n = 40						
spirit	1.010	.989	4.803	5.202	9.774	10.225
trend	1.012	.987	4.800	5.205	9.767	10.232
CPI	1.010	.989	4.800	5.205	9.769	10.229
liquidity	1.012	.989	4.801	5.204	9.767	10.231
n = 60						
spirit	1.008	.991	4.795	5.211	9.763	10.235
trend	1.010	.990	4.794	5.212	9.760	10.238
CPI	1.008	.991	4.794	5.211	9.763	10.235
liquidity	1.009	.990	4.792	5.213	9.759	10.239
AR(4) Wallis test						
n = 20						
CPI	1.056	.939	4.826	5.191	9.729	10.271
seasonals	1.039	.958	4.824	5.181	9.763	10.236
liquidity	1.041	.956	4.824	5.182	9.763	10.237
n = 40						
CPI	1.020	.978	4.790	5.210	9.750	10.248
seasonals	1.015	.984	4.800	5.207	9.761	10.237
liquidity	1.020	.978	4.797	5.209	9.752	10.248
n = 60						
CPI	1.014	.985	4.789	5.216	9.748	10.250
seasonals	1.011	.988	4.792	5.214	9.756	10.242
liquidity	1.013	.986	4.791	5.215	9.753	10.246

(a) True size (%) of Burr critical value approximations.

(b) Burr p-value (%) approximations with true critical values.

Table 5

Burr accuracy for the Breusch-Pagan heterscedasticity test

Nominal size		1%			5%			10%		
		(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
Data	$\left(\sqrt{\beta_1},\beta_2\right)$									
n=24										
trend	(0, 2.89)	1.017	.824	.746	4.974	5.035	2.706	9.958	10.051	6.650
lognormal	(-.67,3.65)	.738	1.258	.554	4.939	5.292	2.462	10.464	9.625	5.019
liquidity	(-.42,3.46)	.932	1.068	.221	4.938	4.962	1.817	10.022	9.831	5.040
uniform	(.13,2.98)	1.060	.922	.641	4.984	5.013	3.873	9.922	10.153	8.369
n=40										
trend	(0, 2.94)	1.015	.980	.779	4.967	5.043	4.464	9.952	10.057	9.301
lognormal	(-.59,3.71)	.951	1.046	.678	4.938	5.073	2.993	9.991	10.004	6.780
liquidity	(-.18,3.31)	.967	1.024	.641	4.911	5.017	3.464	9.948	9.968	7.495
uniform	(.14,2.99)	1.062	.918	.782	4.991	4.999	4.366	9.929	10.083	9.140
n=64										
trend	(0, 2.96)	1.014	.981	.863	4.964	5.046	4.672	9.950	10.066	9.574
lognormal	(-.48,3.57)	.926	1.081	.824	4.920	5.107	3.667	10.006	9.991	7.811
liquidity	(.79,4.20)	1.463	-	.604	5.256	4.670	2.355	9.895	10.251	8.643
uniform	(.12,2.99)	1.049	.933	.863	4.984	5.011	4.611	9.934	10.077	9.475

(a) True size (%) with Burr critical value approximation
(b) Burr p-value (%) approximations with true critical values
(c) True size of test using asymptotic $\chi^2(1)$ critical value.