*On Testing for Serial Correlation in Regression when the Bounds Test is Inconclusive**

M. J. HARRISON

WHENEVER least squares regression is used to analyse economic time series, or cross-section data, the possibility of serially correlated disturbances presents a serious problem.[1] It is, therefore, of considerable importance to be able to test for the presence of serial correlation amongst regression residuals. Usually a one-tail test against positive serial correlation at the 5 per cent or 1 per cent probability level is employed. One such test is the well-known and much used "bounds test", proposed by Durbin and Watson [3]. Based on the modified von Neumann ratio $d = u'Du/u'u$,[2] where u is the vector of estimated residuals and D is the "first differencing matrix", the bounds test probably owes most of its popularity to its simplicity in application. The great drawback of the test, however, is that it is likely to prove inconclusive, particularly when the number of observations is small (less than 20, say) and/or the number of regressors is large, in which case the "region of ignorance", that is, the interval (d_L, d_U), tends to be large.[3]

Unfortunately, paucity of degrees of freedom is common in practical econo-

*The author is grateful to Mr. F. F. Murray, Senior Programmer, Trinity College, for programming advice, and to Dr. R. C. Geary, Economic and Social Research Institute, Dublin, for helpful discussions and several valuable suggestions on a previous draft of this paper. Errors and deficiencies are, of course, the author's sole responsibility.

1. The nature of the problem is explained well in [11], pp. 243–249.

2. The ratio is often written in non-matrix form as $d = \dfrac{\sum_{i=2}^{n}(u_i - u_{i-1})^2}{\sum_{i=1}^{n} u_i^2}$

3. For example, given 16 observations and 5 regressors, the inconclusive zone at the 5 per cent probability level is (0·74, 1·93).

metric work and the researcher using the bounds test is frequently confronted with a d-statistic from which no inference is possible at the conventional levels of significance. There is a danger in this insofar as the researcher feels he can interpret "bounds test inconclusive" as equivalent to "no need to reject the null-hypothesis of independent disturbances", for this will necessarily lead to too many cases of serial correlation being overlooked.[4] This danger is well-illustrated, albeit not mentioned, in a recent study by Pratschke [15], in which the performance of the bounds test and four other tests for serial correlation of errors were compared. The other tests were the "tau test", proposed by Geary [6], the similar "runs test", suggested by Wald and Wolfowitz [18] and Swed and Eisenhart [17], the "chi-squared test" suggested by Griliches *et alia* [7], and the "exact test" put forward by Fisher [4]. Because the chi-squared and exact tests gave such a "poor showing", however, Pratschke left their results aside. Each test was applied to 90 sets of regression residuals obtained when estimating various forms of Engel function using Irish cross-section data. Of the 90 bounds tests, no less than 17, that is, nearly 20 per cent, were inconclusive at the 5 per cent probability level;[5] and of those that were inconclusive, the corresponding tau test indicated significant positive serial correlation in 53 per cent of the cases, and the runs test in over 40 per cent of the cases.[6] To have accepted these bounds tests, in the absence of the other tests, as upholding the null-hypothesis of random disturbances would have been to err in about 50 per cent of the cases.

Thus a need clearly exists, either for an alternative test which does not have the disadvantage of an inconclusive zone, or, if the bounds test is to remain the basis of testing for serial correlation, for a reliable supplementary test for use when the bounds test is inconclusive. The tau and runs tests mentioned above warrant consideration in this respect. The tau test, for example, being based on a simple count of residual sign changes, is an outstandingly practical test, and it is always conclusive. However, it is unlikely to replace the bounds test altogether, for as Geary himself has stated, "d, as a test, is superior to τ"[7]. More recently, this statement has been borne out by Habibagahi and Pratschke [8], who, by means of a Monte Carlo experiment, have concluded that "the power of the familiar Durbin-Watson test . . . appears to exceed that of the Geary test".[8] Nevertheless, on the evidence available tau would seem a valuable supplementary test when the bounds test is inconclusive.

4. [12], p. 793.

5. Had Pratschke's residuals come from regressions using time series rather than cross-section data, it is reasonable to suppose that a larger number of inconclusive bounds tests would have resulted. It is a fact, as Theil [13, p. 201] points out, that a "substantial proportion" of the d-statistics published in the literature is inconclusive; this relates mainly to the time series analysis.

6. [15], Table 4, p. 503. Incidentally, there are minor discrepancies between Table 4 in [15] and Table A5 in [16] from which it is drawn. Indeed, the latter table itself is not entirely error-free. For example, according to Table A5 the d-value for equation 4.2 for clothing (1·013) is not significant; in fact, it is inconclusive.

7. [6], p. 125.

8. [8], p. 184.

Besides the tau test and runs test, however, there are a number of other alternatives which have been specifically designed for use when the bounds test is inconclusive. Of these, those tests based on the beta distribution are of special interest as they derive from the original Durbin-Watson proposal concerning a procedure for use in the event of inconclusiveness, and, more importantly, because, having so far been largely neglected, little is known about their use and about how they perform alongside the better known alternatives, such as those already mentioned. There are three such "beta tests".[9]

Originally, to circumvent inconclusiveness, Durbin and Watson suggested an approximation to the true significance points of d based on fitting a beta distribution with the correct mean and variance.[10] This does not seem to have been used very much in practice, possibly because it was implied in their original paper that the approximation could only be regarded as sufficiently accurate for more than 40 degrees of freedom. Of late, however, Durbin [1] has come to regard this implication as "unduly pessimistic". Indeed, when the basic bounds test is inconclusive and the approximation is not used, he seems to regard the d-test as being incomplete, and has criticised comparative studies involving the d-test on these grounds.[11] The value of Pratschke's comparisons, useful though they are, would seem to be diminished by the fact that no results are given for this approximate test for the 17 cases in which the bounds test is inconclusive.

A similar approximation was proposed by Theil and Nagar [12] for cases in which the regressors are "slowly changing" in the sense that their first and second differences are small in relation to the ranges of the variables themselves. This condition is probably satisfied by most economic time series; it would seem particularly suitable when the observations are arranged, not in chronological order, but according to increasing values of the dominant explanatory variable or principle component of the explanatory variables, as is usually done in cross-section studies. The significance points derived by Theil and Nagar correspond fairly closely with the original significance points of d_U given by Durbin and Watson. Indeed, Durbin has hinted that "the d_U values give a better approximation for the situation envisaged by Theil and Nagar than their own values".[12]

More recently, Henshaw [10] has suggested a method of fitting a beta distribution using the skewness and kurtosis coefficients of d whereby very accurate approximations to the true significance points can be discovered "even when the number of degrees of freedom is small or when the first and second differences of the explanatory variables are large compared with the range of the corresponding variable itself".[13] The test developed by Henshaw is very much a refinement of the one proposed by Theil and Nagar, which itself was based heavily on the procedure suggested by Durbin and Watson. There seems little doubt that Henshaw's version is superior to its predecessors; however, like them, it has not

9. The other alternatives (for example, the test proposed recently by Durbin in [1]) are not pursued here. Durbin and Watson's recent "$a + bd_U$ approximation" is also left aside.

10. [3], p. 163. 11. [1], p. 424. 12. [1], p. 423. 13. [10], p. 647.

been used much in practice. That this is because of the complexity of the calculation involved, as suggested by some writers,[14] is hard to comprehend considering the availability of modern, high-speed calculating equipment. Moreover, it has been suggested by some econometricians that it seems totally unnecessary with this type of test and modern computing equipment to use the Durbin-Watson bounds test in the first place.[15]

However, this paper is concerned with testing for serial correlation when the bounds test has been used but has proved inconclusive. Its purpose is to describe in more detail the three beta tests alluded to above and, in view of what has been said about his results, to extend Pratschke's case study by applying each test to his data for the 17 cases in which the bounds test is inconclusive and comparing the results with those of the corresponding tau and runs tests. Durbin's contention concerning the value of d_U as an approximation to the true significance points of d is also considered as it seems particularly appropriate to the type of data used by Pratschke. The aim is not to give a detailed theoretical account of the tests; this was done in the original papers. Rather, it is to bring together and outline the computational procedures involved when applying the tests; to indicate, in a relatively simple manner, how the tests are actually used. Consequently, in Section II only those results are given which have a role to play when the tests are used. First, material basic to all three tests is presented. This is followed by a brief description of the general method employed by the tests, and then a step by step account of the computational procedure of each individual test. Section III concerns the application of the tests to Pratschke's data for the 17 inconclusive cases; the results are presented and discussed, and a few tentative conclusions are drawn.

FITTING A BETA DISTRIBUTION TO d

In least squares regression, the distribution of d is dependent on the configuration of the $n \times k$ matrix, X, of n observations on the k regressors. The attainable limits of d are given by the smallest, r_L, and the largest, r_U, of the $n-k$ positive characteristic roots of the positive semi-definite matrix $A = D - DX(X'X)^{-1}X'$,[16] where D is the $n \times n$ first differencing matrix. The moments of d about its mean are similarly given in terms of the positive roots of A, being derived using the result that under the null-hypothesis of independent normal disturbances d is distributed independently of its own denominator, so that the moments of the ratio d are ratios of the corresponding moments of numerator and denominator.[17] In order to calculate the moments from data, however, use is made of certain

14. For example, see [1], p. 423; [11], p. 252; [13], p. 201.
15. For example, see [9], p. 486.
16. Concerning this matrix, see [10], pp. 648–649.
17. See Henshaw's equations (8) [10, p. 649]. Incidentally, there seems to be some confusion concerning the source of the result referred to which makes these equations possible. Henshaw cites Durbin and Watson [2, p. 419]. However, the general form of the result can be traced back to Geary [5].

properties of traces of matrices, and particularly of the fact of equality between sums of powers of the roots of A and the traces of the corresponding powers of A itself, to derive exact expressions for the moments of d in terms of traces of powers of A. Assuming independent normal disturbances, the first four moments are as follows:[18]

$$E(d) = \frac{\mathrm{tr}A}{n-k}; \tag{1}$$

$$\mathrm{var}(d) = 2\left[\frac{(n-k)\mathrm{tr}A^2-(\mathrm{tr}A)^2}{(n-k)^2(n-k+2)}\right]; \tag{2}$$

$$\mu_3 = 8\left[\frac{(n-k)^2\mathrm{tr}A^3-3(n-k)\mathrm{tr}A\,\mathrm{tr}A^2+2(\mathrm{tr}A)^3}{(n-k)^3(n-k+2)(n-k+4)}\right]; \tag{3}$$

$$\begin{aligned}\mu_4 = 12\{&(n-k)^3[4\mathrm{tr}A^4+(\mathrm{tr}A^2)^2]-2(n-k)^2[8\mathrm{tr}A^3\mathrm{tr}A+\mathrm{tr}A^2(\mathrm{tr}A)^2]\\ &+(n-k)[24\mathrm{tr}A^2(\mathrm{tr}A)^2+(\mathrm{tr}A)^4]-12(\mathrm{tr}A)^4\}\\ &\div[(n-k)^4(n-k+2)(n-k+4)(n-k+6)].\end{aligned} \tag{4}$$

The skewness and kurtosis coefficients of d are given by

$$\sqrt{\beta_1} = \frac{\mu_3}{[\mathrm{var}(d)]^{\frac{3}{2}}}, \text{ and} \tag{5}$$

(6) $\beta_2 = \dfrac{\mu_4}{[\mathrm{var}(d)]^2}$, respectively, which are easily evaluated once the moments of d have been calculated from data.

The corresponding moments, and the skewness and kurtosis coefficients, of a beta variable, x, with range (0, 1) may be expressed exactly in terms of the parameters, p and q, of the beta distribution.[19] The mean and variance of x are given by

18. The first three moments are given by Henshaw [10, p. 649]. The fourth moment, however, was derived by the present author as Henshaw seems to have made an incorrect transformation of this moment into trace form which manifests itself as an error in the third term of the numerator of his equation (13). Judging from recent journal and textbook references to his test, this error appears to have passed unnoticed, which is surprising considering the importance of equation (13) to the successful application of the test.

19. For a detailed explanation of the relation between the moments of the beta-function and its parameters, see [14], p. xxv.

(7) $$E(x) = \frac{p}{p+q}\text{; and}$$

(8) $$\text{var}(x) = \frac{pq}{(p+q)^2(p+q+1)}\text{, respectively;}$$

the skewness and kurtosis coefficients by

(9) $$\sqrt{\beta_1} = \frac{2(q-p)\sqrt{p+q+1}}{(p+q+2)\sqrt{pq}}\text{, and}$$

(10) $$\beta_2 - 3 = \frac{6[(p+q+1)(q-p)^2-(p+q+2)pq]}{pq(p+q+2)(p+q+3)}.$$

Since the range of d is (r_L, r_U), the link between these two distributions is the transformation of the test statistic into a beta variable, x, with range (0, 1) given by

(11) $$x = \frac{d-r_L}{r_U-r_L}\text{, or}$$

(12) $$d = r_L + (r_U - r_L)x.$$

Expression (11) also affords a convenient way of obtaining the limits of d without calculating the eigenvalues of A directly. For it follows from that equation that

(13) $$E(x) = \frac{E(d)-r_L}{r_U-r_L}\text{, and}$$

(14) $$\text{var}(x) = \frac{\text{var}(d)}{(r_U-r_L)^2}.$$

Simultaneous solution of (13) and (14) gives

(15) $$r_L = E(d) - E(x)\sqrt{\frac{\text{var}(d)}{\text{var}(x)}}\text{, and}$$

(16) $$r_U = \sqrt{\frac{\text{var}(d)}{\text{var}(x)}} + r_L,$$

which can be calculated from data by substituting $E(d)$, var(d), $E(x)$ and var(x) from (1), (2), (7) and (8), respectively.[20]

Each of the three beta tests uses some or all of the above results, either as they stand, in exact form, or in some approximate form, to fit a beta distribution by the method of moments and thereby obtain estimates of the true significance points of d. To facilitate the actual fitting, two moments of d are required since the beta distribution depends on two parameters. The Durbin-Watson procedure uses the first two moments of d; the Theil-Nagar procedure uses approximations to the skewness and kurtosis coefficients of d; the Henshaw test uses the exact skewness and kurtosis coefficients of d. Unlike the mean and variance, the skewness and kurtosis coefficients, since they are independent of origin and scale, are independent of the range that is adjusted.[21] Use of these coefficients in the latter two tests, however, does not mean that other moments do not have to be calculated. It is clear from equations (5) and (6), for example, that a numerical value of the second moment of d, as well as of the third and fourth moments, is required before the skewness and kurtosis coefficients can be computed. Furthermore, a value for the first moment of d is necessary for the evaluation of the lower limit of d from equation (15); as will be seen, this is necessary for the Henshaw test.

The broad objective of each test is to solve for the beta parameters, p and q, in terms of the numerical values of the relevant two moments of d, calculated from data. This having been done, the estimates of p and q are replaced by the tabled values nearest them in the *Tables of the Incomplete Beta-Function* [14];[22] the values of the beta variables, x, corresponding to the required significance levels (for example, 5 per cent or 1 per cent) are then read from these tables. Finally, to estimate the significance points of d (that is, $d_{.05}$ or $d_{.01}$), the tabled values of x, and the numerical values of the limits of d, calculated from data, are substituted into equation (12). Consideration of the actual computational procedure of each test follows.

Durbin-Watson

For their "approximate procedure for use when the bounds test is inconclusive", Durbin and Watson suggest fitting a beta distribution with the correct mean and variance, that is, with the actual mean and variance of d calculated from data. Thus for this test, only the first two moments of d need be calculated. Furthermore, it is assumed that the limits of d are 0 and 4. These are the theoretical extremes; in practice they would rarely be attained. However, by making this assumption the computation involved in using the test is considerably simplified. Equation (11) becomes

20. It may not be immediately clear how $E(x)$ and var(x) are calculated from data using (7) and (8) as these equations involve the parameters p and q whose values must first be ascertained. However, this will be clarified below when it is dealt with in the section on Henshaw.

21 [12], p. 800.

22. For more accurate results interpolation can be used. See [14], pp. xii-xxii.

(17) $x = \frac{d}{4}$, and, using this, equation (7) becomes

(18) $$E(d) = \frac{4p}{p+q},$$ and equation (8) becomes

(19) $$\text{var}(d) = \frac{16pq}{(p+q)^2(p+q+1)}.$$

These two modified expressions, (18) and (19), can be solved such that

(20) $$p+q = \frac{E(d)[4-E(d)]}{\text{var}(d)} - 1,$$ and

(21) $$p = \tfrac{1}{4}(p+q)E(d).$$[23]

Application of the test proceeds as follows: values for $E(d)$ and $\text{var}(d)$ are calculated from data using (1) and (2); these values are substituted in (20) and (21) to yield estimates of p and q which are then used to ascertain, from the *Tables of the Incomplete Beta-Function*, the beta variables, x, for the desired significance levels. The corresponding critical values of d are finally given, using (17), by

(22) $$d = 4x.$$

Theil-Nagar

The Theil-Nagar procedure is the simplest of the three tests to apply in practice. Although it is based on fitting a beta distribution using the skewness and kurtosis coefficients of d, use of the complex formulae (3) and (4) is avoided by making approximations to them such that the parameters of the beta distribution can be estimated using only the number of observations, n, and the number of independent variables in the regression, k. Similarly, by making approximations to the expressions for the first and second moments of d, the limits of d can also be estimated in terms of n and k only. As was intimated above, the accuracy of these approximations improves the slower the rate of change of the variables in the regression. Moreover, it is possible to test whether this condition is met sufficiently well; but the necessary computation is considerable and if carried out would virtually eliminate the need to make the approximations in the first place, for it

23. [3], p. 165.

provides most of the information required for the more accurate Henshaw test which could then be used.[24]

The parameters of the beta distribution are approximated by

$$p \doteq \tfrac{1}{2}(n+k), \text{ and} \tag{23}$$

$$q = \tfrac{1}{2}(n-k+2);^{25} \text{ the limits of } d \text{ by} \tag{24}$$

$$r_L = \frac{4k^2-1}{n^2}, \text{ and} \tag{25}$$

$$r_U = 4-\frac{3}{n^2}.^{26} \tag{26}$$

When applying the test, these four expressions are evaluated first. Using the values of p and q from (23) and (24), respectively, the required beta variables, x, are then obtained from the Beta Tables. Finally, substitution into (12) of x, and r_L and r_U from (25) and (26), respectively, yields the required significance points of d.

However, for most cases in practice even this simple calculation may be superfluous. By virtue of the approximations and the fact that the critical points of d ultimately depend only on n and k, Theil and Nagar were able to construct appropriate tables; these were referred to in Section I above. The tables contain the 5 per cent and 1 per cent significance points of d for certain values of n between 15 and 100 and values of k up to 6.[27] Only in cases where n is smaller than 15 and/or k is larger than 6, therefore, would it be necessary to resort to the actual Theil-Nagar calculation; these, incidentally, are just the cases when the bounds test is likely to prove inconclusive. It will be recalled that the limits for use in the bounds test also are not tabulated for less than 15 observations.

Henshaw

The Henshaw procedure also makes use of the skewness and kurtosis coefficients of d, but in this case they are calculated exactly from the original data. Substitution of the skewness coefficient from (5) and the kurtosis coefficient from (6) into the

24. Section 3 of [12] discusses the various approximations in detail. For further information on testing the validity of the approximations, see [10], pp. 655–656.

25. [12], p. 800.

26. [12], p. 801.

27. [12], Table 2, p. 802. When using this table, it is important to remember that the number of independent variables, k, is "adjusted"; that is to say, the dummy variable, unity, which takes account of the constant term in linear regression, is included in the number. Thus the smallest value of k is 2.

left hand sides of (9) and (10), respectively, allows solution for p and q in terms of the values of β_1 and β_2. The solution, a somewhat fearsome pair of expressions, is simplified by Henshaw by setting

(27) $$W = \frac{4\beta_2 - 3\beta_1}{3\beta_1 - 2\beta_2 + 6}, \text{ and}$$

(28) $$Z = \frac{\sqrt{\beta_1}(W^2 - 1)}{\sqrt{16W + \beta_1(W+1)^2}},$$ so that values of p and q are given by

(29) $$p = \tfrac{1}{2}(W - Z - 1), \text{ and}$$

(30) $$q = \tfrac{1}{2}(W + Z - 1).$$[28]

In applying the test, all four moments of d are first calculated from data using equations (1) through (4). Next, values for β_1 and β_2, derived from (5) and (6), are substituted, via (27) and (28), into (29) and (30) to give the estimates of p and q. These estimates serve two purposes: firstly, they allow $E(x)$ and var(x) to be estimated from (7) and (8), respectively, which, together with $E(d)$ and var(d) from (1) and (2), respectively, are used in equations (15) and (16) to determine the attainable limits of d, r_L and r_U, respectively. Secondly, as in the other two tests, they are used to obtain the required beta variables, x, from the Beta Tables. Finally, the values for x, r_L and r_U are substituted into (12) to obtain the significance points of d against which the original d-statistic can be judged. If the statistic which proved inconclusive on the bounds test is greater than the "Henshaw significance point" for the appropriate probability level, the null-hypothesis of random disturbances is accepted at that level of probability; if it is less, the alternative hypothesis of first-order positive autocorrelation is accepted. The significance points derived from the other two beta procedures are used in the same way.

RESULTS AND CONCLUSIONS

The three beta tests were applied to Pratschke's data for the 17 cases in which the bounds test was inconclusive.[29] For each test the values for p and q, and where relevant, r_L and r_U, were obtained by means of a double precision Fortran computer programme.[30] Using the computed values of p and q, the *Tables of the*

28. [10], pp. 650–651.

29. The data used in Pratschke's regression analyses [16] have not been published in their original form. On advice from Dr. Pratschke, however, they were obtained directly from the Central Statistics Office, Dublin, with help from Mr. D. Murphy.

30. The programme was written by the author; it is available for inspection on application to him.

Incomplete Beta-Function furnished values for x corresponding to the 5 per cent and 1 per cent probability levels. The corresponding significance points of d were given by equation (22) in the case of the Durbin-Watson beta procedure (D-W), and equation (12) in the cases of the Theil-Nagar procedure (T-N) and the Henshaw procedure (H). The significance points, along with the inconclusive d-statistics calculated by Pratschke, are presented in Table A3 of the Appendix. The final results of the tests, together with the corresponding results for the tau and runs tests reported by Pratschke, and the d_U approximation suggested by Durbin, are given in Table 1. To enable easy comparison of the performance of all six alternatives in these 17 cases, a simple concordance of results is given in Table 2.

TABLE 1: *Results of Six Alternative Tests for the Cases in which the Bounds Test is Inconclusive*

	Equation[a]	τ[b]	u[b]	*Beta Tests*			d
				D-W	*T-N*	*H*	*approx.*
		1	*2*	*3*	*4*	*5*	*6*
Food	5	*	ns	*	*	*	*
	10	ns	ns	*	*	*	*
	13	**	**	**	**	**	**
Clothing	1	*	*	*	*	*	*
	2	*	*	*	*	*	*
	3	ns	ns	*	*	*	*
	10	**	**	**	**	**	**
	11	*	ns	**	**	**	**
	13	ns	ns	ns	*	*	*
	17	ns	ns	ns	*†	ns	*†
	18	ns	ns	ns	*†	ns	*†
Fuel and Light	9	*	**	*	*	*	*
	13	**	**	*	*	*	*
Housing	4	ns	ns	*	*	*	*
	12	ns	ns	*	*	*	*
	17	ns	ns	ns	*†	ns	*†
Sundries	8	*	*	*	*	*	*

Notes: *a* Equation numbers correspond with those used in [15, Table A]. The functional forms of the equations are described in [16, Table 4].
b The τ and u results are those reported in [15, Table 4].
* Indicates significance at the 5 per cent level.
** Indicates significance at the 1 per cent level.
ns Indicates not significant at the 5 per cent level.
† Theil-Nagar approximations inappropriate; test inconclusive.

TABLE 2: *Comparison of the Performance of Six Alternative Tests for the Cases in which the Bounds Test is Inconclusive*

Significance Level	Tests					
	τ 1	u 2	D-W 3	T-N 4	H 5	d_U 6
1 per cent	3	4	3	3	3	3
5 per cent	6	3	10	11	11	11
Not Significant at 5 per cent level	8	10	4	—	3	—
Inconclusive	—	—	—	3	—	3
Total	17	17	17	17	17	17

The incidence of significant positive autocorrelation indicated by τ and u at the conventional significance levels has already been noted in percentage terms in Section I above. The first and second columns of Table 2 give the actual number of rejections and acceptances of the null-hypothesis suggested by these two tests at the two levels, as reported by Pratschke. The similarity between the two sets of results is reasonably close, but this was to be expected considering the similarity of the two tests. Both tests are always conclusive.

Of the results for the beta tests, those for the Durbin-Watson procedure and the Henshaw procedure are strikingly similar. Both tests are always conclusive, and both indicate significant autocorrelation at the 5 per cent level in about 80 per cent of the 17 cases. The number of rejections and acceptances of the null-hypothesis suggested by the two tests at the 5 per cent and 1 per cent significance levels is given in the third and fifth columns of Table 2. If the results of the Henshaw test are accepted as being accurate to the order claimed by Henshaw, then such close agreement of these results would seem to corroborate Durbin's belief that the Durbin-Watson beta approximation is not as bad, even for small numbers of observations, as is usually thought. Had it been used by Pratschke, there would have been no cases of inconclusiveness in his case study. Moreover, at the 5 per cent significance level, the total number of significant d-tests would have been raised considerably, from 20 to 33; the total number of non-significant d-tests would have been raised slightly, too, from 53 to 57.

The results of the Theil-Nagar test are not too dissimilar from those of the other two beta tests, but certain of the results should be viewed with suspicion. Similarity between the Theil-Nagar results and those of the Henshaw test was not entirely unexpected since the ordered cross-section data which were used seemed to fulfil the necessary condition of "smoothness". Moreover, a test for this condition forms part of the calculation embodied in the computer programme that was used. It is similar to, but more stringent than the test proposed by Theil and Nagar and used by Henshaw,[31] involving comparisons of the trace of the matrix D with that of A, and the trace of D^2 with that of A^2. In each case, the difference between the traces, which is small when the condition is satisfied, is expressed as a percentage of the trace of the appropriate power of A. For 14 of the 17 sets of data the test percentages were small, less than about 5 per cent, signifying that the Theil-Nagar approximation is valid for those cases. However, in 3 of the 17 cases the test percentages were greater than 5 per cent. It is felt that it is not coincidental that these are the cases of equations 17 and 18 for clothing, and equation 17 for housing in which the Theil-Nagar procedure indicated significant serial correlation at the 5 per cent level while the Henshaw procedure indicated no significant serial correlation at that level. Since the Theil-Nagar approximation seems inappropriate for these 3 cases, the corresponding results are best left aside. The most that can be said about the Theil-Nagar results, therefore, is that at the 5 per cent level, as can be seen from Table 2, they indicate significant autocorrelation in about 80 per cent of the 17 cases, but for the remaining 20 per cent of the cases the results are inconclusive.

The use of the d_U-value in place of an inconclusive bounds test for cases in which the regressors are "slowly changing", as suggested by Durbin, is clearly equivalent to "playing safe". Whenever a d-statistic is inconclusive at some level of significance, it is assumed to be indicative of significant serial correlation at that level. If the d-statistic is inconclusive at more than one significance level (the 1 per cent as well as the 5 per cent level, for example), then use of the lower (1 per cent) critical value of d_U ensures that the statistic is judged significant at the lower level. Thus in this study, all 17 d-statistics, being inconclusive at the 5 per cent level, are assessed by the d_U approximation as significant at the 5 per cent level at least. Those relating to equation 13 for food, and equations 10 and 11 for clothing are judged as being significant at the 1 per cent level also.

The d_U results, as can be seen from the fourth and sixth columns of Tables 1 and 2, are identical to those given by the Theil-Nagar test. Hence, they would seem to bear out Durbin's claim that the d_U approximation is as good as, if not better than, the Theil-Nagar beta procedure for cases in which the variables are changing slowly.[32] Like the Theil-Nagar test, however, the accuracy of the d_U results depends on the condition of smoothness being met. As has already been pointed out, there are 3 cases for which this condition is not met sufficiently well. For the d_U approximation as well, therefore, it can be safely concluded only that 14 cases of significant autocorrelation are indicated at the 5 per cent level. Nothing can be said in the other 3 cases; they are inconclusive.

Leaving aside the 3 cases for which the Theil-Nagar procedure and the d_U approximation are not valid, the most interesting aspect of the results is the marked difference in the performance of τ and u on the one hand, and of D-W, T-N, H and d_U on the other, in detecting serial correlation in the cases in which

31. [10], pp. 655–656.

32. The similarity between the d_U values and the Theil-Nagar significance points can be clearly seen by comparing Table A1 with Table A2 in the Appendix.

the bounds test is inconclusive. τ and u indicate significance at the 5 per cent level in roughly 50 per cent of the 17 cases; the other 4 tests in about 80 per cent of the cases. There is very close similarity amongst all 6 tests at the 1 per cent level; the differences occur at the 5 per cent level. The indications are that τ and u are less sensitive to the presence of autocorrelation than the other 4 tests. Two things add credence to this. First, the available evidence indicates that the power of the Geary tau test declines as the number of observations falls below about 30,[33] and the regressions which formed the basis of Pratschke's case study, and hence form the basis of the present study, made use of only 16 observations. Second, the evidence on the Henshaw test, sparse as it is, suggests that it is a powerful test even for small numbers of observations.[34] It should be stressed, however, that no firm, general conclusions can be drawn from the results presented here since the "truth", as in Pratschke's original case study, is unknown, and in any case because the sample of cases is small.

Nevertheless, the 17 cases have provided useful material with which to illustrate the tests described in this paper, and the results are not uninteresting nor entirely valueless. Indeed, it is felt that the results, and those of the Durbin-Watson beta procedure in particular, have made a positive contribution in filling the gap in the original case study caused by Pratschke's failure to use the procedure suggested by Durbin and Watson for the cases in which the bounds test is inconclusive.

Finally, the findings suggest that it would be worthwhile to subject Henshaw's test, the best of the beta tests, to more rigorous testing alongside other tests. For if it proves as powerful as it is thought to be, there is no reason why it should not be incorporated as part of any standard regression package. Research to compare the power of Geary's test and Henshaw's test amongst others is being pursued by the present author using a Monte Carlo approach along the lines of that used by Habibagahi and Pratschke (*op. cit.*), and will be reported in a forthcoming paper.

Trinity College, Dublin.

33. See [8].

34. To test the power of his test for an extreme case of 16 observations on 6 regressors, Henshaw applied it to the calculation of significance points for a related distribution for which accurate significance points are available. Even for this small number of degrees of freedom, the significance points of d obtained by his beta procedure agreed with the corresponding exact significance points "to an order of accuracy that is adequate in applied work with economic time series" [10, p. 652].

APPENDIX

TABLE A1: *Critical Values of Durbin-Watson d (one-tail)*

Number of Observations	*Significance Level Per Cent*	*Number of Independent Variables*	d_L	d_U
16	5	1	1·10	1·37
		2	0·98	1·54
		3	0·86	1·73
	1	1	0·84	1·09
		2	0·74	1·25
		3	0·63	1·44

Source: Durbin and Watson, *op. cit.*

TABLE A2: *Critical Vales of d from the Theil-Nagar Beta Approximation*

Number of Observations	*Significance Level Per Cent*	*Number of Independent Variables*	*Critical d*
16	5	1	1·37
		2	1·53
		3	1·71
	1	1	1·08
		2	1·24
		3	1·42

Source: Theil and Nagar, *op. cit.*

TABLE A3: *Calculated Critical Values of d for the Durbin-Watson, Theil-Nagar, and Henshaw Beta Tests*

	Equation[a]	Number of Independent Variables	Inconclusive d-statistic[b]	Beta Tests D-W 1%	D-W 5%	T-N 1%	T-N 5%	H 1%	H 5%
Food	5	2	1·280	1·180	1·464	1·254	1·531	1·195	1·481
	10	2	1·457	1·180	1·464	1·254	1·531	1·253	1·538
	13	2	1·144	1·180	1·464	1·254	1·531	1·258	1·542
Clothing	1	2	1·357	1·180	1·464	1·254	1·531	1·258	1·542
	2	2	1·259	1·180	1·464	1·254	1·531	1·244	1·533
	3	1	1·143	1·064	1·348	1·084	1·367	1·065	1·360
	10	2	1·044	1·180	1·464	1·254	1·531	1·253	1·538
	11	2	1·013	1·180	1·464	1·254	1·531	1·253	1·536
	13	2	1·498	1·180	1·464	1·254	1·531	1·258	1·542
	17	3	1·571	1·232	1·516	1·417	1·709	1·176	1·474
	18	3	1·677	1·232	1·516	1·417	1·709	1·173	1·471
Fuel and Light	9	1	1·157	1·028	1·360	1·084	1·367	1·004	1·293
	13	2	1·425	1·180	1·464	1·254	1·531	1·258	1·542
Housing	4	2	1·352	1·180	1·464	1·254	1·531	1·267	1·543
	12	1	1·206	1·052	1·320	1·084	1·367	1·013	1·286
	17	3	1·691	1·232	1·516	1·417	1·709	1·176	1·474
Sundries	8	2	1·273	1·180	1·464	1·254	1·531	1·195	1·481

Notes: a Equation numbers correspond with those used in [15, Table A]. The functional forms of the equations are described in [16, Table 4].

b *Source:* [16], Table A5, p. 25. All of the *d*-statistics are inconclusive at the 5 per cent level. Those for equation 13 for food, and equations 10 and 11 for clothing, are inconclusive at the 1 per cent level also.

REFERENCES

[1] Durbin, J., "An Alternative to the Bounds Test for Testing for Serial Correlation in Least Squares Regression", *Econometrica,* 38 (1970), 422–429.

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

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

[4] Fisher, R. A., *Statistical Methods for Research Workers* (10th Ed.), Oliver and Boyd, 1946.

[5] Geary, R. C., "A General Expression for the Moments of Certain Symmetrical Functions of Normal Samples", *Biometrika,* 25 (1933), 184–186.

[6] Geary, R. C., "Relative Efficiency of Count of Sign Changes for Assessing Residual Autoregression in Least Squares Regression", *Biometrika*, 57 (1970), 123–127.
[7] Griliches, Z., Maddala, G. S., Lucas, R., and Wallace, N., "Notes on Estimated Aggregate Quarterly Consumption Functions", *Econometrica*, 30 (1962), 491–500.
[8] Habibagahi, H. and Pratschke, J. L., "A Comparison of the Power of the Von Neumann Ratio, Durbin-Watson and Geary Tests", *Review of Economics and Statistics*, 54 (1972), 179–185.
[9] Hannan, E. J., "A Note on an Exact Test for Trend and Serial Correlation", *Econometrica*, 37 (1969), 485–489.
[10] Henshaw, R. C., "Testing Single-Equation Least Squares Regression Models for Autocorrelated Disturbances", *Econometrica*, 34 (1966), 646–660.
[11] Johnston, J., *Econometric Methods* (2nd Ed.), McGraw-Hill, 1972.
[12] Theil, H. and Nagar, A. L., "Testing the Independence of Regression Disturbances", *Journal of the American Statistical Association*, 56 (1961), 793–806.
[13] Theil, H., *Principles of Econometrics*, North-Holland, 1971.
[14] Pearson, K., *Tables of the Incomplete Beta-Function* (2nd Ed.), Cambridge U.P., 1968.
[15] Pratschke, J. L., "Testing for Uniformity of Fit in Regression: An Econometric Case Study", *Economic and Social Review*, 2 (1971), 497–505.
[16] Pratschke, J. L., *Income-Expenditure Relations in Ireland 1965–1966*, Economic and Social Research Institute, Dublin, Paper No. 50.
[17] Swed, F. S. and Eisenhart, C., "Tables for Testing Randomness of Grouping in a Sequence of Alternatives", *Annals of Mathematical Statistics*, 14 (1943).
[18] Wald, A. and Wolfowitz, J., "On a Test Whether Two Samples are From The Same Population", *Annals of Mathematical Statistics*, 11 (1940).