Econometrica, Vol. 40, No. 4 (July, 1972)

TESTING FOR FOURTH ORDER AUTOCORRELATION IN QUARTERLY REGRESSION EQUATIONS[1]

BY KENNETH F. WALLIS

A test for fourth order autocorrelation in the error term of a regression equation estimated from quarterly data is described. The development draws on the finite sample results of Durbin and Watson and illustrates how their procedure for the first order case can be generalized. In the model $y = X\beta + u$ where X is a matrix of fixed regressors and $u_t = \rho u_{t-4} + \varepsilon_t$, an appropriate test statistic for $H_0: \rho = 0$ is the statistic $d_4 = \{\Sigma(z_t - z_{t-4})^2\}/\Sigma z_t^2$ computed from the least squares regression residuals $z = y - Xb$. Bounds to the significance points of d_4 are tabulated. Maximum likelihood estimation methods are described; these are equally appropriate when lagged values of the dependent variable appear among the regressors, and they provide asymptotic tests for general autoregressive error structures, as well as for the special case $u_t = \alpha_1 u_{t-1} + \alpha_4 u_{t-4} - \alpha_1\alpha_4 u_{t-5} + \varepsilon_t$. Examples from the empirical literature are presented.

1. INTRODUCTION

THE POSSIBILITY THAT the errors in a regression equation estimated from quarterly data possess fourth order autocorrelation was considered, among other things, in a recent paper [**28**], and a non-parametric test was proposed. Appropriate generalized least squares estimation methods were also described. In this paper we first present a more rigorous solution to the problem of testing for fourth order autocorrelation, which utilizes the approach introduced by Durbin and Watson [**6** and **7**]. We then describe non-linear estimation methods which simultaneously estimate the regression coefficients and the parameters of the simple fourth order or more general autoregressive error structures.

The usual interpretation of the error or disturbance term in econometric models is that it represents the effect of omitted or unobservable variables on the dependent variable. The error term might thus be expected to display certain features of observed economic variables, in particular, when quarterly data are employed, seasonal variation. Equally, when seasonally unadjusted data are being employed in order that one may attempt to explain seasonal variation in the dependent variable, along with other types of variation, by means of explanatory economic or seasonal dummy variables, then the presence of non-systematic seasonal variation, or an incomplete accounting for seasonality by the regressors, will produce seasonal effects in the error term, with the possible consequence of fourth order autocorrelation. Thus we require a test for correlation not between the errors

[1] Some of the results contained in this paper were reported in my paper "Estimation and Tests for Quarterly Regression Equations with Autocorrelated Errors" presented at the Second World Congress of the Econometric Society in Cambridge, September, 1970. I am grateful to David Hendry for comment and discussion and, in Section 3, for the use of his computer program, to Zvi Griliches and an anonymous referee for comments, to Andrew Tremayne for research assistance, and to M. I. Nadiri and Michael Parkin for supplying their data.

Added in proof: After this was written an unpublished paper by H. D. Vinod entitled "Generalization of the Durbin-Watson Statistic for Higher Order Autoregressive Processes" was brought to my attention; this considers statistics similar to d_4 for tests of higher order autocorrelation in the non-seasonal case.

in successive quarters, but between the errors in the corresponding quarters of successive years.

Durbin and Watson [**6** and **7**] consider the testing of the hypothesis that $\rho = 0$ when the regression disturbances $\{u_t\}$ are generated by the first order, stationary Markov scheme $u_t = \rho u_{t-1} + \varepsilon_t$. They propose a now well-known test statistic, based on the least squares regression residuals. Since the residuals depend on the particular regressors employed, exact significance points for their test statistic cannot be obtained once-and-for-all; instead, bounds to the significance points are tabulated, and the interval between the bounds comprises a range of values of the statistic for which the test is inconclusive. In this situation Durbin and Watson recommend that the true distribution of their test statistic, transformed to the range (0, 1), be approximated by a beta distribution with the same mean and variance. Subsequently, Theil and Nagar [**27**] and Henshaw [**16**] have developed further approximate procedures for use when the bounds test is inconclusive. Durbin [**3**] extends the basic results to the case where the estimated equation is one of a system of simultaneous regression equations, and Durbin [**5**] considers the asymptotic distribution of the test statistic in the case where some of the regressors are lagged values of the dependent variable. Asymptotic tests for the second order autoregression are described by Sargan [**26**] and by Hannan and Terrell [**13**], who use the second order partial autocorrelation coefficient as a test statistic. By use of frequency domain methods, Durbin [**4**] considers testing the null hypothesis of serial independence against an alternative hypothesis of a completely general autocorrelation structure.

In Section 2 we present a generalization of the Durbin-Watson bounds test to the case of fourth order autocorrelation, which we represent by means of a simple fourth order autoregression $u_t = \rho u_{t-4} + \varepsilon_t$. This provides a test appropriate to the situation discussed above, and also illustrates how the Durbin-Watson procedure can be generalized to higher order cases. The fourth order test statistic has a distribution which differs from that of the Durbin-Watson statistic, and tables of its critical values are presented. By way of comparison, we note an analogy with the problem of testing for first order autocorrelation when there are missing observations, for which there is as yet no general solution. Comments on the problem of lagged dependent variables are also presented. This simple fourth order autoregressive model should be distinguished from the general fourth order autoregression $u_t = \alpha_1 u_{t-1} + \ldots + \alpha_4 u_{t-4} + \varepsilon_t$, which presents somewhat greater difficulty for the finite sample test of Section 2, but for which we provide estimation methods and asymptotic tests in Section 3. In the presence of autocorrelated errors, least squares estimates are inefficient, and their standard error estimates are biased, and if lagged values of the dependent variable appear among the regressors, as is increasingly the case in applied econometric work, least squares estimates are inconsistent. In order to achieve consistent and asymptotically efficient estimates of equations containing lagged dependent variables and with errors following more general autoregressive schemes, we adopt maximum likelihood methods, since modern computational facilities make non-linear iterative methods easily available. These are described together with some examples in Section 3.

2. TESTING FOR FOURTH ORDER AUTOCORRELATION

2.1. *A Bounds Test*

The regression model is

$$y = X\beta + u$$

where y and u are $n \times 1$ vectors, β is a $k \times 1$ vector of coefficients, and X is an $n \times k$ matrix of rank k of observations on the regressors, treated as "fixed variables." The vector z of least squares regression residuals is given by

$$z = (I_n - X(X'X)^{-1}X')y = My,$$

say, where I_n is the $n \times n$ identity matrix, M is idempotent, and $z = Mu$. We consider the ratio of quadratic forms

$$r = \frac{z'Az}{z'z} = \frac{u'MAMu}{u'Mu}$$

where A is a real symmetric matrix. The basic lemma proved by Durbin and Watson [6] then states (i) there exists an orthogonal transformation $u = H\xi$ such that

$$r = \frac{\sum_{i=1}^{n-k} v_i\xi_i^2}{\sum_{i=1}^{n-k} \xi_i^2}$$

where $v_1, v_2, \ldots, v_{n-k}$ are the non-zero eigenvalues of MA; (ii) if s of the columns of X are linear combinations of s of the eigenvectors of A, and if the eigenvalues of A associated with the remaining $n - s$ eigenvectors are renumbered so that

$$\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_{n-s},$$

then

$$\lambda_i \leqslant v_i \leqslant \lambda_{i+k-s} \qquad (i = 1, 2, \ldots, n-k);$$

(iii) corollary: $r_L \leqslant r \leqslant r_U$ where

$$r_L = \frac{\sum_{i=1}^{n-k} \lambda_i\xi_i^2}{\sum_{i=1}^{n-k} \xi_i^2} \quad \text{and} \quad r_U = \frac{\sum_{i=1}^{n-k} \lambda_{i+k-s}\xi_i^2}{\sum_{i=1}^{n-k} \xi_i^2}.$$

The bound $r_L(r_U)$ is attained when the $k - s$ "other" columns of X are eigenvectors of A associated with the eigenvalues $\lambda_{n-k+1}, \ldots, \lambda_{n-s}$ $(\lambda_1, \ldots, \lambda_{k-s})$ omitted from the numerator expression. We also use a result of T. W. Anderson

[2], who shows that for cases in which the errors have density function

$$K \exp\left[-\frac{1}{2\sigma^2}\{(1+\rho^2)u'u - 2\rho u'\theta u\}\right] \tag{1}$$

and in which the regression vectors are eigenvectors of θ (or linear combinations of k of them), the uniformly most powerful (UMP) test of the null hypothesis $H_0 : \rho = 0$ against alternatives $\rho > 0$ is given by

$$\frac{z'\theta z}{z'z} > r_0$$

where r_0 is determined to give a critical region of appropriate size. Thus in particular, with $A = \theta$, the statistic r provides a test which is UMP against one-sided alternatives when the lower bound r_L is attained (or the upper bound).

Durbin and Watson [6 and 7] consider the testing of $H_0 : \rho = 0$ when the errors $\{u_t\}$ are generated by the first order autoregression

$$u_t = \rho u_{t-1} + \varepsilon_t, \qquad |\rho| < 1,$$

where the ε_t are independent $N(0, \sigma^2)$ random variables. These errors have covariance matrix $E(uu') = V_1$ with inverse

$$V_1^{-1} = \frac{1}{\sigma^2}\begin{bmatrix} 1 & -\rho & 0 & \dots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \dots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & \dots & 0 & 0 \\ \vdots & & & & & \\ & & & & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \dots & -\rho & 1 \end{bmatrix}$$

and density function

$$K \exp\left[-\frac{1}{2\sigma^2}\left\{(1+\rho^2)u'u - \rho^2(u_1^2 + u_n^2) - 2\rho \sum_{t=2}^{n} u_t u_{t-1}\right\}\right] \tag{2}$$

which unfortunately cannot be written in the form of (1) with a matrix θ independent of ρ. For convenience, Durbin and Watson choose the matrix

$$A_1 = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ 0 & -1 & 2 & \dots & 0 \\ \vdots & & & & \vdots \\ & & & 2 & -1 \\ 0 & & \dots & -1 & 1 \end{bmatrix}.$$

Taking $\theta = I_n - \frac{1}{2}A_1$ in (1) gives the density

$$(3)\qquad K \exp\left[-\frac{1}{2\sigma^2}\left\{(1+\rho^2)u'u - \rho(u_1^2 + u_n^2) - 2\rho\sum_{t=2}^{n} u_t u_{t-1}\right\}\right]$$

which is close to (2), becoming closer as ρ increases, and the eigenvectors of A_1 and θ are the same. Hence the test statistic

$$d_1 = \frac{z'A_1z}{z'z} = \frac{\sum_2^n (z_t - z_{t-1})^2}{\sum_1^n z_t^2}$$

provides a test which is approximately UMP against one-sided alternatives when the regression vectors are eigenvectors of A_1. That is, d_1 would be UMP if the joint density of the disturbances were given by (3) instead of being given (as assumed) by (2); Anderson [**2**] shows that in this latter case no UMP test exists. The eigenvalues of A_1 are given by von Neumann [**24**] as zero and

$$\lambda_j = 2\left(1 - \cos\frac{\pi j}{n}\right) \qquad (j = 1, \ldots, n-1).$$

The eigenvector corresponding to the zero root is a column of ones; hence, for regressions with an estimated constant term, we have from the Durbin-Watson lemma, with $s = 1$ and λ_j as given above,

$$d_{1,L} \leqslant d_1 \leqslant d_{1,U}$$

where

$$d_{1,L} = \frac{\sum_{i=1}^{n-k} \lambda_i \xi_i^2}{\sum_{i=1}^{n-k} \xi_i^2} \quad \text{and} \quad d_{1,U} = \frac{\sum_{i=1}^{n-k} \lambda_{i+k'} \xi_i^2}{\sum_{i=1}^{n-k} \xi_i^2}$$

and $k' = k - s = k - 1$ is the number of independent variables in the model in addition to the constant term. Significance points of $d_{1,L}$ and $d_{1,U}$ are tabulated by Durbin and Watson in [**7**].

We now generalize these results to provide a bounds test for fourth order autocorrelation by considering the simple fourth order autoregressive model

$$(4)\qquad u_t = \rho u_{t-4} + \varepsilon_t$$

where the ε_t are independent normal random variables as before. The null hypothesis H_0 is the hypothesis that $\rho = 0$ in (4). The autocovariance $E(u_t u_{t-l})$ is zero unless l is an integer multiple of 4, in which case it is $(\sigma^2/(1-\rho^2))\rho^{l/4}$. If quarterly observations are available for m complete years, then the total number of observations is $n = 4m$. If $B = [b_{ij}]$ is a $p \times q$ matrix and C is a $g \times h$ matrix, the Kronecker product $B \otimes C$ is defined as the $pg \times qh$ matrix

$$B \otimes C = [b_{ij}C].$$

The covariance matrix of the errors generated by (4) is then given by

$$E(uu') = V_4 = V_1 \otimes I_4$$

where V_4 is $n \times n$ and V_1 is $m \times m$. We also have

$$V_4^{-1} = V_1^{-1} \otimes I_4,$$

and the errors have the density function

$$(5) \qquad K \exp\left[-\frac{1}{2\sigma^2}\left\{(1+\rho^2)u'u - \rho^2\left(\sum_{t=1}^{4} u_t^2 + \sum_{t=n-3}^{n} u_t^2\right) - 2\rho\sum_{t=5}^{n} u_t u_{t-4}\right\}\right]$$

which again cannot be written exactly in the form (1) with a matrix θ independent of ρ. However, by analogy with the Durbin-Watson statistic, we form the matrix

$$A_4 = A_1 \otimes I_4$$

where A_1 is now $m \times m$, and consider the statistic

$$(6) \qquad d_4 = \frac{z'A_4 z}{z'z} = \frac{\sum_5^n (z_t - z_{t-4})^2}{\sum_1^n z_t^2}.$$

This is again a special case of the statistic r and bounds $d_{4,L}$ and $d_{4,U}$ exist by the Durbin-Watson lemma. Taking $\theta = I_n - \frac{1}{2}A_4$ in (1) gives the density

$$K \exp\left[-\frac{1}{2\sigma^2}\left\{(1+\rho^2)u'u - \rho\left(\sum_{t=1}^{4} u_t^2 + \sum_{t=n-3}^{n} u_t^2\right) - 2\rho\sum_{t=5}^{n} u_t u_{t-4}\right\}\right]$$

which is close to (5), and the eigenvectors of θ and A_4 are the same. Thus, again, the statistic d_4 provides a test of H_0 which is approximately UMP against one-sided alternatives when the regression vectors are eigenvectors of A_4. In particular, the test given by d_4 when the lower bound $d_{4,L}$ is attained is approximately UMP (and $d_{4,U}$ similarly). The bounds are again given by

$$(7) \qquad d_{4,L} = \frac{\sum_{i=1}^{n-k} \lambda_i \xi_i^2}{\sum_{i=1}^{n-k} \xi_i^2} \quad \text{and} \quad d_{4,U} = \frac{\sum_{i=1}^{n-k} \lambda_{i+k-s} \xi_i^2}{\sum_{i=1}^{n-k} \xi_i^2}$$

where the λ's are now the roots of A_4, excluding the s roots associated with eigenvectors which appear as columns of X. Since the roots of a Kronecker product matrix are given by the products of the roots of the two constituent matrices in all possible pairs, the n roots of A_4 are 0 and $2(1 - \cos(\pi j/m))$, $j = 1, \ldots, m-1$, each appearing with a multiplicity of 4, which we may write as

$$\lambda_j = 2\left(1 - \cos\frac{\pi}{m}\left[\frac{j-1}{4}\right]\right) \qquad (j = 1, \ldots, n)$$

where $[f]$ denotes the integer part of the real quantity f.

In order to tabulate five per cent significance points, we compute the distribution (or a relevant section thereof) of the ratio of quadratic forms in normal variables (7) utilizing a computer program by Koerts and Abrahamse [**18**] which is based on methods described by Imhof [**17**]. In practice, we apply this method to compute the probability

$$P\{d_{4,L} < d\} = P\left\{\sum_{i=1}^{n-k} (\lambda_i - d)\xi_i^2 < 0\right\}$$

where the ξ_i^2 are independent chi-square variables with one degree of freedom, for a number of values of d, and by interpolation and successive approximation obtain that value, $d_{4,L}^*$ say, such that

$$P\{d_{4,L} < d_{4,L}^*\} = 0.05$$

to the required degree of accuracy (and $d_{4,U}^*$ similarly).

We tabulate significance points for two regression situations, according to whether or not the regression equation contains the seasonal dummy variables Q_{it}, $i = 1, \ldots, 4$, taking the value 1 in the ith quarter of each year and zero otherwise. Table I represents the situation in which such variables are not employed, but in which a constant term is estimated by including a column of ones among the variables. This regression vector corresponds to one of the zero roots of A_4;

TABLE I

FIVE PER CENT SIGNIFICANCE POINTS OF $d_{4,L}$ AND $d_{4,U}$ FOR REGRESSIONS WITHOUT QUARTERLY DUMMY VARIABLES ($k = k' + 1$)

	$k' = 1$		$k' = 2$		$k' = 3$		$k' = 4$		$k' = 5$	
n	$d_{4,L}$	$d_{4,U}$	$d_{4,L}$	$d_{4,U}$	$d_{4,L}$	$d_{4,U}$	$d_{4,L}$	$d_{4,U}$	$d_{4,L}$	$d_{4,U}$
16	0.774	0.982	0.662	1.109	0.549	1.275	0.435	1.381	0.350	1.532
20	0.924	1.102	0.827	1.203	0.728	1.327	0.626	1.428	0.544	1.556
24	1.036	1.189	0.953	1.273	0.867	1.371	0.779	1.459	0.702	1.565
28	1.123	1.257	1.050	1.328	0.975	1.410	0.898	1.487	0.828	1.576
32	1.192	1.311	1.127	1.373	1.061	1.443	0.993	1.511	0.929	1.587
36	1.248	1.355	1.191	1.410	1.131	1.471	1.070	1.532	1.013	1.598
40	1.295	1.392	1.243	1.442	1.190	1.496	1.135	1.550	1.082	1.609
44	1.335	1.423	1.288	1.469	1.239	1.518	1.189	1.567	1.141	1.620
48	1.369	1.451	1.326	1.493	1.281	1.537	1.236	1.582	1.191	1.630
52	1.399	1.475	1.359	1.513	1.318	1.554	1.276	1.595	1.235	1.639
56	1.426	1.496	1.389	1.532	1.351	1.569	1.312	1.608	1.273	1.648
60	1.449	1.515	1.415	1.548	1.379	1.583	1.343	1.619	1.307	1.656
64	1.470	1.532	1.438	1.563	1.405	1.596	1.371	1.629	1.337	1.664
68	1.489	1.548	1.459	1.577	1.427	1.608	1.396	1.639	1.364	1.671
72	1.507	1.562	1.478	1.589	1.448	1.618	1.418	1.648	1.388	1.678
76	1.522	1.574	1.495	1.601	1.467	1.628	1.439	1.656	1.411	1.685
80	1.537	1.586	1.511	1.611	1.484	1.637	1.457	1.663	1.431	1.691
84	1.550	1.597	1.525	1.621	1.500	1.646	1.475	1.671	1.449	1.696
88	1.562	1.607	1.539	1.630	1.515	1.654	1.490	1.677	1.466	1.702
92	1.574	1.617	1.551	1.639	1.528	1.661	1.505	1.684	1.482	1.707
96	1.584	1.626	1.563	1.647	1.541	1.668	1.519	1.690	1.496	1.712
100	1.594	1.634	1.573	1.654	1.552	1.674	1.531	1.695	1.510	1.717

hence we have $s = 1$ in (7), and three zeros are included in the remaining $n - 1$ roots. The number of independent variables additional to the constant term is $k' = k - 1$, and $d_{4,L}(d_{4,U})$ is a function of the $n - (k' + 1)$ smallest (largest) values of these $n - 1$ roots. Table II represents the situation in which the seasonal dummy variables are employed (equivalently, a constant term and three of these variables may be included). These variables are eigenvectors associated with the four-fold zero root of A_4; hence in this case we have $s = 4$ in (7), and the zero roots are excluded. Various values of $k'' = k - 4$, the number of "economic" variables among the regressors, are considered, $d_{4,L}(d_{4,U})$ being a function of the $n - (k'' + 4)$ smallest (largest) values of the $n - 4$ non-zero roots of A_4. We note that as far as $d_{4,U}$ is concerned, the two cases are identical for a given value of k, the total number of regressors, whenever $k > 4$—the upper bound is attained when the regression vectors coincide with eigenvectors corresponding to the smallest of the $n - s$ roots, and in regressions without quarterly dummy variables such vectors are nevertheless given by these variables. No similar relations exist for the lower bound.

The procedure for carrying out a test of size .05 of the null hypothesis $H_0: \rho = 0$ in (4) against the alternative $\rho > 0$ is as follows. The statistic d_4 is computed as in (6); if $d_4 < d^*_{4,L}$ the null hypothesis should be rejected; if $d_4 > d^*_{4U}$ the null hypothesis is accepted; and if $d^*_{4,L} \leqslant d_4 \leqslant d^*_{4,U}$ the test is inconclusive. Similarly,

TABLE II

FIVE PER CENT SIGNIFICANCE POINTS OF $d_{4,L}$ AND $d_{4,U}$ FOR REGRESSIONS INCLUDING A CONSTANT TERM AND QUARTERLY DUMMY VARIABLES ($k = k'' + 4$)

	$k'' = 1$		$k'' = 2$		$k'' = 3$		$k'' = 4$		$k'' = 5$	
n	$d_{4,L}$	$d_{4,U}$	$d_{4,L}$	$d_{4,U}$	$d_{4,L}$	$d_{4,U}$	$d_{4,L}$	$d_{4,U}$	$d_{4,L}$	$d_{4,U}$
16	1.156	1.381	1.031	1.532	0.902	1.776	0.777	2.191	0.693	2.238
20	1.228	1.428	1.123	1.556	1.013	1.726	0.899	1.954	0.806	2.042
24	1.287	1.459	1.199	1.565	1.107	1.694	1.011	1.856	0.928	1.949
28	1.337	1.487	1.261	1.576	1.181	1.679	1.099	1.803	1.025	1.889
32	1.379	1.511	1.312	1.587	1.243	1.673	1.171	1.773	1.104	1.850
36	1.414	1.532	1.355	1.598	1.293	1.672	1.230	1.755	1.170	1.824
40	1.445	1.550	1.391	1.609	1.336	1.674	1.279	1.745	1.225	1.807
44	1.471	1.567	1.422	1.620	1.373	1.677	1.321	1.739	1.272	1.795
48	1.494	1.582	1.450	1.630	1.404	1.681	1.357	1.737	1.312	1.788
52	1.514	1.595	1.474	1.639	1.432	1.686	1.389	1.736	1.347	1.782
56	1.533	1.608	1.495	1.648	1.456	1.691	1.416	1.736	1.377	1.779
60	1.549	1.619	1.514	1.656	1.478	1.696	1.441	1.737	1.404	1.777
64	1.564	1.629	1.531	1.664	1.497	1.700	1.463	1.739	1.429	1.776
68	1.577	1.639	1.546	1.671	1.515	1.705	1.482	1.741	1.450	1.775
72	1.590	1.648	1.560	1.678	1.531	1.710	1.500	1.743	1.470	1.776
76	1.601	1.656	1.573	1.685	1.545	1.714	1.517	1.746	1.488	1.776
80	1.611	1.663	1.585	1.691	1.559	1.719	1.531	1.748	1.504	1.777
84	1.621	1.671	1.596	1.696	1.571	1.723	1.545	1.751	1.519	1.778
88	1.630	1.677	1.607	1.702	1.582	1.727	1.558	1.753	1.533	1.779
92	1.639	1.684	1.616	1.707	1.593	1.731	1.570	1.756	1.546	1.781
96	1.647	1.690	1.625	1.712	1.603	1.735	1.580	1.759	1.558	1.782
100	1.654	1.695	1.633	1.717	1.612	1.739	1.591	1.761	1.569	1.784

for testing against negative values of ρ (which lead to a value of d_4 between 2 and 4), H_0 should be rejected if $d_4 > 4 - d^*_{4,L}$ and accepted if $d_4 < 4 - d^*_{4,U}$, and the test is inconclusive whenever $4 - d^*_{4,U} \leqslant d_4 \leqslant 4 - d^*_{4,L}$. We note that as in the Durbin-Watson first order test, the size of the inconclusive region decreases as the sample size increases, and increases as the number of regressors increases.[2]

2.2 *The Durbin-Watson Statistic and Missing Observations*

An alternative approach to the calculation of the statistic d_4 would be to re-arrange the data so that the first quarter observations for all years are listed first, followed by the observations on all the second quarters, and so on, and then seek a measure of first order autocorrelation appropriate to this series of data. Since the least squares calculations are invariant with respect to data ordering, the effect is simply to rearrange the elements of the vector z of regression residuals to give

$$\tilde{z}' = (z_1, z_5, \ldots, z_{n-3}, z_2, z_6, \ldots, z_{n-2}, z_3, z_7, \ldots, z_{n-1}, z_4, z_8, \ldots, z_n).$$

The sum of squares of first differences of this series, eliminating differences $(z_2 - z_{n-3})$, $(z_3 - z_{n-2})$, and $(z_4 - z_{n-1})$ which relate to different quarters, is given by $\tilde{z}'(I_4 \otimes A_1)\tilde{z}$, which is identical with the quantity $z'(A_1 \otimes I_4)z = z'A_4 z$ considered earlier. That is, the statistic d_4 can be alternatively calculated as

$$(8) \qquad d_4 = \frac{\tilde{z}'(I_4 \otimes A_1)\tilde{z}}{\tilde{z}'\tilde{z}}.$$

(We note that this formulation can be readily adapted to situations in which the number of observations is not an integer multiple of four. For example, if $n = 4m + 2$, the time series beginning with year 1, quarter I and ending with year $m + 1$, quarter II, then the appropriate matrix for the quadratic form in the numerator of (8) is

$$\begin{bmatrix} A_1(m+1) & 0 & 0 & 0 \\ 0 & A_1(m+1) & 0 & 0 \\ 0 & 0 & A_1(m) & 0 \\ 0 & 0 & 0 & A_1(m) \end{bmatrix},$$

[2] Approximations to the distribution of d_4 under the null hypothesis for use when the bounds test is inconclusive can be obtained by again arguing by analogy with Durbin and Watson. For the first order case, Durbin and Watson [**8**] show that their original beta approximation and a linear approximation of the form $a + b\,d_U$ both perform well (whereas the Theil-Nagar [**27**] approximation cannot be recommended for practical use and the Henshaw [**16**] approximation, although very accurate, requires so much additional computation that its practical usefulness seems limited.) In our case, to apply the first approximation we assume that $\frac{1}{4}d_4$ has the beta distribution with mean and variance corresponding to those of d_4, which can be obtained for the particular X-matrix being used from expressions given by Durbin and Watson [**6**, pp. 419–420] for the exact mean and variance of $r = z'Az/z'z$. The second, linear approximation is obtained by assuming that d_4 has the distribution of $\tilde{d}_4 = a + bd_{4,U}$ where the distribution of $d_{4,U}$ depends on the particular values of n and k, and a and b are chosen so that d_4 and $\tilde{d}_4$ have the same mean and variance. We require the first two moments of $d_{4,U}$, and these again follow by specializing general expressions given by Durbin and Watson [**6**, p. 419]. These results, and further details of these approximations, are available upon request from the author.

where $A_1(m)$ denotes the mth order matrix A_1. The eigenvalues of this matrix comprise four zeros, the values $2(1 - \cos(\pi j/m))$, $j = 1, \ldots, m - 1$ twice, and the values $2(1 - \cos(\pi j/(m + 1)))$, $j = 1, \ldots, m$ twice, and from these the upper and lower bounding statistics can be constructed, and values of the significance points calculated in order to fill in Tables I and II for values of n other than multiples of 4. In practice, this has not been judged necessary.[3])

The main point of this section is that although d_4 as calculated above can be regarded as a Durbin-Watson statistic, d_1, corrected for the "gaps" in the data (between the observations for the ith quarter of year m and the $(i + 1)$th quarter of year 1, $i = 1, 2, 3$), the results of the previous section indicate that the distributions of d_4 and d_1 are quite different. A feature of at least one regression program in widespread use (see [**12**]) is the calculation of a Durbin-Watson statistic "adjusted for j gaps," that is, correcting for nonadjacent observations by eliminating the squared difference between the last residual before a set of missing observations and the first residual after such a set from the numerator of d_1. (The term "missing observation" as used here covers cases where the investigator has good economic grounds for fitting a relation to selected sub-periods of the total sample period.) No advice is given on the test procedure to be employed with such a statistic, and a comparison of our tables with those of Durbin and Watson [**7**] indicates that no simple rules of thumb apply at least in the case we consider. The value of d_4 calculated as in (8) from the vector of residuals $\tilde{z}$ would be reported as a "Durbin-Watson statistic adjusted for three gaps." One might casually compare this with the Durbin-Watson significance points for the appropriate value of k' and an effective sample size of $n - 3$, but a comparison of these significance points with those presented in Tables I and II indicates that this procedure is subject to considerable inaccuracy, and that Durbin-Watson statistics "adjusted for ... gaps" should be treated with considerable caution. What similarities there are between our tables and those of Durbin and Watson occur for the lower bound, $d_{4,L}$, in Table II, for regressions with quarterly dummy variables, comparing their k' with our k''. These similarities relate, broadly speaking, to the treatment of the zero root. However, one hesitates to suggest any rule of thumb for the testing of Durbin-Watson statistics adjusted for gaps on the basis of such slender evidence.

2.3 *Lagged Dependent Variables*

In the preceding discussion, it is assumed that the explanatory variables can be regarded as "fixed variables," whereas a feature of much empirical work in econometrics is the presence of lagged values of the dependent variable among the explanatory variables. For the first order case, Nerlove and Wallis [**23**] calculate

[3] In this situation, where the roots except for the zeros are equal in pairs, we could utilize the exact distribution derived by R. L. Anderson [**1**] for the circular serial correlation coefficient. For a few such cases, we have applied Anderson's results to calculate exact probabilities for values of $d_{4,U}$, as a check on Imhof's numerical integration methods. However, Anderson's distribution function becomes difficult to compute as the number of roots increases.

the asymptotic bias in the Durbin-Watson statistic in the model

$$(9)\qquad \begin{aligned} y_t &= \alpha y_{t-1} + u_t, \\ u_t &= \rho u_{t-1} + \varepsilon_t, \end{aligned}$$

and show that d_1 is asymptotically biased towards 2 (the value which is expected if no autocorrelation is in fact present). In this section we assess the effect of the lagged dependent variable on the test for fourth order autocorrelation presented above by considering the model

$$(10)\qquad \begin{aligned} y_t &= \alpha y_{t-1} + u_t, \\ u_t &= \rho u_{t-4} + \varepsilon_t. \end{aligned}$$

We concentrate on the effect of the variable y_{t-1}; the addition of exogenous variables to the right-hand side of the regression equation again modifies the biases which we observe. It might be argued that a correct formulation of dynamic behavior in a quarterly model requires the presence of higher order lagged values of the dependent variable. While this is not a commonly observed feature in empirical econometrics, it is the case that replacing the term y_{t-1} in (10) by y_{t-4} produces exactly the same asymptotic bias in d_4 as that calculated for d_1 in the model (9) by Nerlove and Wallis. The results presented below indicate that this bias is rather less pronounced in the model (10).

Relegating the details to the Appendix, the least squares estimate $\hat{\alpha} = \Sigma\, y_t y_{t-1}/\Sigma\, y_{t-1}^2$ has probability limit

$$\operatorname{plim} \hat{\alpha} = \frac{\alpha + \alpha^3\rho}{1 + \alpha^4\rho} = \alpha + \frac{\alpha^3\rho(1 - \alpha^2)}{1 + \alpha^4\rho}.$$

Thus for $\rho > 0$, $\hat{\alpha}$ is again asymptotically biased upwards. Considering the statistic d_4 calculated from the regression residuals $z_t = y_t - \hat{\alpha} y_{t-1}$, we have

$$\operatorname{plim} d_4 = \operatorname{plim} 2\left(1 - \frac{\Sigma\, z_t z_{t-4}}{\Sigma\, z_t^2}\right) = 2(1 - \operatorname{plim} \hat{\rho}),$$

say. The asymptotic bias

$$b = \operatorname{plim} d_4 - 2(1 - \rho) = -2 \operatorname{plim}(\hat{\rho} - \rho)$$

is shown in the Appendix to be

$$b = \frac{2\rho\alpha^6(1 - \alpha^2)(1 - \rho^2)}{(1 + \alpha^4\rho)(1 - \alpha^6\rho^2)}.$$

This has the sign of ρ, and so d_4 is again biased towards 2; b is an even function of α, and is zero to three decimal places for $\alpha = \pm 0.3$. For greater values of α, the magnitude of the bias is illustrated in Table III, where $\operatorname{plim} d_4$ is compared with the "true" value of $2(1 - \rho)$ for various values of ρ. A comparison with Table I of Nerlove and Wallis [**23**] shows that the biases in d_4 are much smaller

TABLE III

VALUES OF PLIM d_4 FOR SELECTED VALUES OF ρ AND α

	$-.9$	$-.7$	$-.5$	$-.3$	$-.1^{\rho}$	.1	.3	.5	.7	.9
$2(1-\rho)$ / α	3.8	3.4	3.0	2.6	2.2	1.8	1.4	1.0	.6	.2
$\pm.9$	3.652	3.220	2.870	2.528	2.178	1.819	1.448	1.066	.667	.238
$\pm.7$	3.771	3.345	2.947	2.564	2.188	1.812	1.431	1.041	.639	.219
$\pm.5$	3.796	3.391	2.991	2.593	2.198	1.802	1.406	1.009	.608	.204

than those in d_1, and hence indicates that a first order lag in the dependent variable presents a relatively less serious problem when testing for fourth order autocorrelation.

2.4. *An Illustration*

As an illustration of the proposition that in general not only seasonal variables but also procedures to account for fourth order autocorrelation may be required in regressions estimated from quarterly data, Thomas and Wallis [**28**] present a regression of United Kingdom consumers' expenditure on "other alcohol," 1955–67, on a trend and three dummy variables, which although acceptable on the conventional criteria (R^2, d_1, t-ratios) exhibits significant fourth order autocorrelation in the residuals according to their sign reversal test. However, a log-linear form is preferred to a linear form, according to a likelihood ratio test described by Sargan [**26**, p. 42], and the log-linear regression yields a value of 2.77 for their $\chi^2(1)$ test statistic for fourth order autocorrelation, which is not significant. Nevertheless, the value of $d_4 = 1.140$ is significant, when compared with the critical value of $d_{4,L} = 1.514$ given in Table II ($k'' = 1, n = 52$), and this may provide an indication of the greater power of the Durbin-Watson-type procedure. The results are as follows, with standard errors in parentheses:

$$\log C_t = \underset{(0.024)}{4.634} + \underset{(0.001)}{0.017t} - \underset{(0.025)}{0.818Q_{1t}} - \underset{(0.025)}{0.515Q_{2t}} - \underset{(0.025)}{0.457Q_{3t}},$$

$$n = 52, \qquad R^2 = 0.978, \qquad d_1 = 1.945, \qquad d_4 = 1.140.$$

On simultaneously estimating the regression coefficients and the fourth order autocorrelation parameter by methods described below, we obtain

$$\log C_t = \underset{(0.045)}{4.598} + \underset{(0.001)}{0.018t} - \underset{(0.040)}{0.821Q_{1t}} - \underset{(0.040)}{0.513Q_{2t}} - \underset{(0.040)}{0.454Q_{3t}}, \quad \underset{(0.136)}{\hat{\rho} = 0.402}.$$

The standard errors of the estimated coefficients have increased, indicating that they were initially underestimated in the presence of autocorrelated errors. Note however that the explanatory variables used in this illustration are examples of regression variables for which ordinary least squares estimates are asymptotically efficient even when the errors are autocorrelated.

The test for fourth order autocorrelation presented in this section draws on the results of Durbin and Watson, and it serves to illustrate how their finite sample results for the first order case can be generalized to provide a test for autocorrelation of higher orders. A test for sth order autocorrelation in the context of the simple model $u_t = \rho u_{t-s} + \varepsilon_t$ can be obtained by replacing 4 by s at appropriate points in the preceding discussion.

3. EFFICIENT ESTIMATION

3.1. *General Considerations*

In this section efficient estimation in the presence of higher order autocorrelation is considered, and the simple fourth order autoregressive model is considerably generalized. However, finite sample tests as described in the previous section are not available for the more general situations, and the tests carried out in the various examples below are valid only asymptotically. In this situation we can relax the requirement of fixed regressors, and consider the regression model

$$y_t = \beta_1 x_{1t} + \ldots + \beta_k x_{kt} + \beta_{k+1} y_{t-1} + \ldots + \beta_{k+l} y_{t-l} + u_t \qquad (t = 1, \ldots, n) \tag{11}$$

where the error term satisfies the general rth order autoregression

$$u_t = \alpha_1 u_{t-1} + \ldots + \alpha_r u_{t-r} + \varepsilon_t. \tag{12}$$

The case considered in the previous section, $u_t = \rho u_{t-4} + \varepsilon_t$, gives a special case of this general structure. We also consider the case given by a multiplication of the first order and fourth order quasi-difference operators, namely,

$$(1 - \alpha_1 L)(1 - \alpha_4 L^4) u_t = \varepsilon_t \tag{13}$$

where L is the lag operator. It is often convenient to set up the autoregressive generating function in terms of its roots rather than its coefficients, and we have observed patterns of this form when estimating the autoregressive structure of single time series. The general form (12) is also considered for values of $r = 1, \ldots, 5$. All this is undertaken in an exploratory spirit—lag structures, seasonal effects, adjustment patterns, data construction, or various misspecifications may be such that little can be said about the form of (12) a priori.

Applying the operator $(1 - \alpha_1 L - \ldots - \alpha_r L^r)$ to equation (11) gives a regression equation with a serially uncorrelated error term ε_t, but whose $k(r + 1) + (l + r)$ coefficients are subject to certain restrictions, as the model contains only $k + l + r$ parameters. These are estimated simultaneously by minimizing the residual sum of squares

$$\phi = \sum_{t=r+1}^{n} e_t^2$$

with respect to the $k + l + r$ parameters. (We assume that we have n observations on the variables of equation (11), so that the y-series begins with the observation y_{-l+1}.) We use a program by Hendry [**14**], which utilizes a minimization algorithm by Powell [**25**] designed specifically for minimizing sums of squares of non-linear functions. On assuming that the ε_t are normally distributed random variables, the resulting estimates become very close to, and asymptotically equivalent to, maximum likelihood estimates.[4]

As a simple illustration, we take $k = l = r = 1$, whereupon the model is

$$y_t = \beta_1 x_t + \beta_2 y_{t-1} + u_t,$$

$$u_t = \alpha_1 u_{t-1} + \varepsilon_t,$$

and solving out gives the equation

$$y_t = \beta_1 x_t - \alpha_1 \beta_1 x_{t-1} + (\alpha_1 + \beta_2) y_{t-1} - \alpha_1 \beta_2 y_{t-2} + \varepsilon_t.$$

This equation has a serially uncorrelated error term, but least squares coefficient estimates will not in general satisfy the non-linear restriction implied by the fact that these four coefficients are functions of three parameters. We impose this restriction by applying an iterative method to minimize

$$\phi = \sum_{t=2}^{n} e_t^2 = \sum_{t=2}^{n} \{y_t - b_1 x_t + a_1 b_1 x_{t-1} - (a_1 + b_2) y_{t-1} + a_1 b_2 y_{t-2}\}^2$$

with respect to a_1, b_1, and b_2 (cf. Fuller and Martin [**10**]).

The close connection between a given regression relation with an autoregressive error and a relation with a more complicated lag structure but an independent error is noted by Griliches [**11**], who suggests that one may distinguish between the model

$$y_t = \beta_1 x_t + u_t, \qquad u_t = \alpha_1 u_{t-1} + \varepsilon_t,$$

and the partial adjustment model

$$y_t = \gamma_1 x_t + \gamma_2 y_{t-1} + \eta_t, \qquad \eta_t \text{ independent},$$

by adding the term $\gamma_3 x_{t-1}$ and seeing whether $\hat{\gamma}_3 \simeq -\hat{\gamma}_1 \hat{\gamma}_2$; if so, the former specification is preferred. Effectively, this amounts to a test of the non-linear restriction imposed in our estimation procedure, and when $r = 1$ we adopt the likelihood ratio test described for the general relation (11) by Sargan [**26**, pp. 27–28], which is easily generalized to higher order autoregressions (see also [**20**, Ch. 9.5]). However, in the unrestricted transformed equation the number of coefficients to be estimated is $k(r + 1) + (l + r) - m$, where m is the number of variables which are redundant in Sargan's terminology, that is, variables which when lagged are identical with, or linear combinations of, other variables appearing in the equation, examples being the constant term, trend and seasonal dummy variables, and x-variables which have their own lagged values appearing among the k explanatory exogenous variables in (11). Thus when the error term obeys the

[4] The discrepancy between our estimates and maximum likelihood estimates arises from the treatment of initial observations ($e_1, \ldots, e_r$ in ϕ above).

general structure (12), (least squares) estimation of this equation may provide some difficulty in practice, for the number of coefficients may no longer be small relative to the effective number of observations $n - r$. Rejection of the autoregressive error hypothesis in favor of the unrestricted transformed equation would suggest that one might profitably revise the lag structure originally specified in (11).

The maximum likelihood approach permits various likelihood ratio tests on the autoregressive parameters, jointly or singly, to be carried out, analogous to those described by Fisher [**9**]. Standard errors of the estimates of the α's and β's are based on analytic second derivatives of ϕ, except in the case of (13) which is non-linear in α's, when numerical differentiation is applied. Thus "t-tests" may be carried out on individual α's which are asymptotically equivalent to the likelihood ratio test of a single coefficient, and the standard errors of the regression coefficients are not conditional on the estimates of the autoregressive parameters, unlike those presented by Fisher [**9**]. (For further details see Hendry [**14** and **15**].)

3.2. *The Demand for Money by United States Manufacturing Firms*

The first example is an equation estimated by Nadiri [**22**] in a study of the demand for money by United States manufacturing firms. The least squares estimates are

$$\begin{aligned} m_t = \underset{(0.210)}{0.289} - \underset{(0.019)}{0.054}v_t + \underset{(0.028)}{0.066}(c/w)_{t-2} + \underset{(0.111)}{0.346}X_t - \underset{(0.095)}{0.240}X_{t-1} \\ + \underset{(0.073)}{0.741}m_{t-1} - \underset{(0.012)}{0.045}Q_{1t} + \underset{(0.010)}{0.009}Q_{2t} + \underset{(0.011)}{0.018}Q_{3t}, \\ n = 52, \qquad R^2 = 0.901, \qquad h = 1.004. \end{aligned}$$

The data are quarterly, seasonally unadjusted, logarithms of the following variables: m is real cash holdings, v is the opportunity cost of real cash balances, c/w is the factor price ratio, and X is sales; h is the test statistic for first order autocorrelation introduced by Durbin [**5**] for cases in which the regression equation contains lagged values of the dependent variable; it should be tested as a standard normal deviate. According to the conventional criteria, the estimates are acceptable, and a joint test of the seasonal dummy variables gives an F-ratio of 13.54 with (3,43) degrees of freedom, which is highly significant. On calculating d_4 we obtain a value of 1.140; hence, the null hypothesis of no fourth order autocorrelation is rejected (the bias caused by m_{t-1} having been shown to be small in Section 2.3). Estimation subject to the simple fourth order autoregressive error $u_t = \rho u_{t-4} + \varepsilon_t$ yields:

$$\begin{aligned} m_t = \underset{(0.363)}{1.153} - \underset{(0.016)}{0.026}v_t + \underset{(0.029)}{0.063}(c/w)_{t-2} + \underset{(0.088)}{0.233}X_t - \underset{(0.075)}{0.280}X_{t-1} \\ + \underset{(0.069)}{0.707}m_{t-1} - \underset{(0.020)}{0.076}Q_{1t} - \underset{(0.018)}{0.008}Q_{2t} + \underset{(0.018)}{0.004}Q_{3t}, \qquad \hat{\rho} = \underset{(0.131)}{0.527}. \end{aligned}$$

This is preferred to more general schemes (12) and (13), according to likelihood ratio tests, and the quarterly dummy variables remain significant, so we have a further example of both seasonal variables and estimation subject to fourth order autocorrelation being required. A test of the non-linear restriction imposed in estimation amounts to a test of the hypothesis that the coefficient of a variable in the original specification is equal to $-\hat{\rho}$ times the coefficient of its four quarter lagged value, when such lagged values of all except the dummy variables are added into the equation. This gives a value of 6.15 for the $\chi^2(5)$ test statistic, which is not significant; hence, the original equations, autoregressive error specification, is preferred. We see that in the revised estimates, the coefficient of v_t becomes non-significant; hence, the original equation, autoregressive error specification, is reducing the estimated average lag from $8\frac{1}{2}$ to $7\frac{1}{4}$ months.

3.3. *Price Inflation in the United Kingdom*

Our second example is the price equation estimated from quarterly data by Lipsey and Parkin [**19**] in their study of inflation in the post-war United Kingdom. The least squares estimates are

$$\underset{}{p_t} = \underset{(0.548)}{1.374} + \underset{(0.102)}{0.562}w_t + \underset{(0.019)}{0.085}m_t - \underset{(0.042)}{0.145}q_t,$$

$$n = 76, \qquad R^2 = 0.699, \qquad d_1 = 0.945, \qquad d_4 = 2.550,$$

where the variables are proportionate rates of change of retail prices (p), wage rates (w), import prices lagged one quarter (m), and output per head (q). With both d_1 and d_4 significant, we might expect that an autoregressive error structure of the form (13) would be relevant, and this is indeed the case. The maximum likelihood estimates are

$$p_t = \underset{(0.725)}{1.821} + \underset{(0.144)}{0.445}w_t + \underset{(0.029)}{0.081}m_t - \underset{(0.053)}{0.142}q_t,$$

$$(1 - \underset{(0.137)}{0.685}L)(1 + \underset{(0.111)}{0.461}L^4)u_t = e_t.$$

In testing the autoregressive error structure, the hypothesis that $\alpha_5 = -\alpha_1\alpha_4$ in the model

$$u_t = \alpha_1 u_{t-1} + \alpha_4 u_{t-4} + \alpha_5 u_{t-5} + \varepsilon_t \tag{14}$$

is accepted, and more generally, in the model (12) with $r = 5$ the joint hypothesis $\alpha_2 = \alpha_3 = 0, \alpha_5 = -\alpha_1\alpha_4$ is accepted. Seasonal variables are not required at any stage. While a negative $\hat{\alpha}_4$ (corresponding to $d_4 > 2$) is rather unexpected, on turning to a test of the non-linear restriction imposed in estimation we find that the original specification, autoregressive error hypothesis, is rejected in favor of the more complicated lagged structural equation. This equation, containing one, four, and five quarter lagged values of all four variables and an independent error term, achieves a substantially lower residual sum of squares than the restricted form given above. However, on closer inspection it turns out that most of this reduction is achieved by relaxing the restriction on the transformed equation that coefficients of variables as of time t and $t-4$ should be equal to $-\alpha_1$ times the coefficients of the same variables as of time $t-1$ and $t-5$, respectively, and that removing the restriction implied by the use of the $(1-\alpha_4 L^4)$ transformation achieves little further reduction. This is confirmed in likelihood ratio tests, and so the preferred form, after dropping some non-significant variables, is

$$\begin{aligned} p_t &= \underset{(0.153)}{0.220} + \underset{(0.078)}{0.567}p_{t-1} + \underset{(0.076)}{0.364}w_{t-1} + \underset{(0.015)}{0.070}m_t - \underset{(0.023)}{0.041}q_t, \\ u_t &= -\underset{(0.103)}{0.459}u_{t-4} + e_t. \end{aligned}$$

While it appears that a dynamic price adjustment mechanism is in operation, the rate-of-change variables used by Lipsey and Parkin may be a contributing factor to the observed autocorrelation, for they are constructed according to $x_t = (X_{t+2} - X_{t-2})/\frac{1}{2}(X_{t+2} + X_{t-2})$ where X_t denotes the level of a given variable. Thus the observed fourth order effect is probably in part a reflection of the four-quarter differencing procedure; we note that four-period differences of a random series have a fourth order autocorrelation coefficient of -0.5.

3.4. *Output Decisions in the United States Cement Industry*

A final example is provided by the production equation estimated from quarterly unadjusted data on the U.S. cement industry in the interwar years by Mills [**21**]. The least squares estimates are

$$z_t = \underset{(25.38)}{72.44} + \underset{(0.030)}{0.806}x_t - \underset{(0.107)}{0.266}I_{t-1} + \underset{(0.031)}{0.155}z_{t-1},$$

$$n = 57, \qquad R^2 = 0.964, \qquad h = -1.359, \qquad d_4 = 0.639,$$

where z_t is production, x_t is sales, and I_t is end-quarter inventory level. While h is not significant, d_4 is highly significant, despite the small bias caused by the presence of z_{t-1}. This corresponds to the "obvious seasonal pattern in the error terms" noted by Mills. On re-estimating subject to an autoregressive error term, the more general form (12) is preferred, in contrast to the previous two examples, and in likelihood ratio tests the hypothesis that $r = 4$ is accepted against other alternative values. The autoregressive error hypothesis is also accepted against the unrestricted transformed regression equation. The estimates are

$$\begin{aligned} z_t &= \underset{(20.42)}{68.08} + \underset{(0.053)}{0.805}x_t - \underset{(0.100)}{0.366}I_{t-1} + \underset{(0.056)}{0.248}z_{t-1} \\ u_t &= -\underset{(0.169)}{0.341}u_{t-1} - \underset{(0.158)}{0.346}u_{t-2} - \underset{(0.158)}{0.369}u_{t-3} + \underset{(0.166)}{0.414}u_{t-4} + e_t. \end{aligned}$$

The coefficient of z_{t-1} has increased, indicating slightly less rapid adjustment, and suggesting that the predominantly negative autocorrelation biased the least squares coefficient downwards. The estimated standard errors show a tendency to increase, but this is not unambiguous. The pattern of the autoregressive coefficients suggests, however, that seasonal dummy variables might be appropriate, and adding these to the original equation gives the least squares estimates

$$\begin{aligned} z_t = {} & \underset{(20.17)}{59.05} + \underset{(0.049)}{0.716x_t} - \underset{(0.097)}{0.337I_{t-1}} + \underset{(0.048)}{0.302z_{t-1}} - \underset{(8.34)}{0.35Q_{1t}} \\ & + \underset{(14.01)}{52.40Q_{2t}} - \underset{(10.11)}{2.05Q_{3t}}, \\ & R^2 = 0.981, \qquad h = 1.232, \qquad d_4 = 1.072. \end{aligned}$$

The joint test of the seasonal dummy variables gives a value of 14.59 for the $F(3, 50)$ test statistic, which is highly significant. Inclusion of these variables increases the value of d_4, but this remains significant, again neglecting the presence of z_{t-1}. More striking is the increase in the coefficient of z_{t-1}, which is approximately double its original value, and in the context of a partial adjustment model raises the estimated average lag to about half a quarter. Thus an alternative explanation of the relatively low value of this coefficient in the initial estimates is that these estimates are subject to a specification error, namely omission of the seasonal variables.

When the equation including seasonal variables is re-estimated subject to an autoregressive error, a similar ambiguity emerges. The appropriate order of the error autoregression appears to be $r = 3$; that is, including the seasonal variables leads to acceptance of the hypothesis that $\alpha_4 = 0$, $\hat{\alpha}_1$, $\hat{\alpha}_2$, and $\hat{\alpha}_3$ remaining significant and negative (the value of $\hat{\alpha}_2$ presumably contributing to the significance of d_4). If the hypothesis that $r = 3$ is maintained, and the seasonable variables tested, these turn out to be significant; if the hypothesis that $r = 4$ is maintained, however, the seasonal variables are not significant. Whereas we have previously presented examples in which both seasonal variables and estimation subject to a fourth order autoregressive error are required, this case provides a certain contrast in that the autoregressive parameter α_4 and the seasonal dummy variables are alternatives, and a likelihood ratio test between them is not possible. A conclusion which emerges more strongly from this example, and also from the previous two examples, is that more detailed specification of lag structures and adjustment behavior in the context of seasonal variation is required.

London School of Economics

Manuscript received May, 1970; revision received April, 1971.

APPENDIX

In this appendix we derive the results presented in Section 2.4 for the model

$$y_t = \alpha y_{t-1} + u_t,$$
$$u_t = \rho u_{t-4} + \varepsilon_t \qquad (t = 1, \ldots, n),$$

where $|\alpha|, |\rho| < 1$, and the ε_t are independent $N(0, \sigma^2)$ random variables. Eliminating u_t from the regression equation gives

$$y_t - \alpha y_{t-1} - \rho y_{t-4} + \alpha\rho y_{t-5} = \varepsilon_t.$$

Writing the autocovariance $E(y_t y_{t-r}) = \gamma_r$, the autocovariance generating function $\Sigma \gamma_r z^r$ is given by $\sigma^2/A(z)A(z^{-1})$ where $A(z) = (1 - \alpha z)(1 - \rho z^4)$; hence γ_r is obtained by calculating the coefficient of z^r in the function

$$\sum \gamma_r z^r = \sigma^2(1 - \alpha z)^{-1}(1 - \alpha z^{-1})^{-1}(1 - \rho z^4)^{-1}(1 - \rho z^{-4})^{-1}$$
$$= \frac{\sigma^2}{(1 - \alpha^2)(1 - \rho^2)} \sum_{-\infty}^{\infty} \alpha^{|j|} z^j \sum_{-\infty}^{\infty} \rho^{|k|} z^{4k}.$$

Thus

$$\gamma_0 = \frac{\sigma^2(1 + \alpha^4\rho)}{(1 - \alpha^2)(1 - \rho^2)(1 - \alpha^4\rho)}, \qquad \gamma_1 = \frac{\sigma^2\alpha(1 + \alpha^2\rho)}{(1 - \alpha^2)(1 - \rho^2)(1 - \alpha^4\rho)},$$

and

$$\operatorname{plim} \hat{\alpha} = \operatorname{plim} \frac{(1/n)\sum y_t y_{t-1}}{(1/n)\sum y_{t-1}^2} = \frac{\gamma_1}{\gamma_0} = \frac{\alpha + \alpha^3\rho}{1 + \alpha^4\rho}.$$

We also require three further terms, as follows:

$$\frac{\gamma_3}{\gamma_0} = \frac{\alpha(\alpha^2 + \rho)}{1 + \alpha^4\rho}, \qquad \frac{\gamma_4}{\gamma_0} = \frac{\alpha^4 + \rho}{1 + \alpha^4\rho}, \qquad \frac{\gamma_5}{\gamma_0} = \frac{\alpha^5 + \rho(\alpha + \alpha^3\rho - \alpha^5\rho)}{1 + \alpha^4\rho}.$$

An estimate of ρ is calculated from the least squares residuals $z_t = y_t - \hat{\alpha} y_{t-1}$, as $\hat{\rho} = (\Sigma z_t z_{t-4})/(\Sigma z_{t-4}^2)$. Thus

$$\operatorname{plim} \hat{\rho} = \operatorname{plim} \frac{(1/n)\Sigma (y_t - \hat{\alpha} y_{t-1})(y_{t-4} - \hat{\alpha} y_{t-5})}{(1/n)\Sigma (y_t - \hat{\alpha} y_{t-1})^2} = \operatorname{plim} \frac{(1 + \hat{\alpha}^2)\gamma_4 - \hat{\alpha}(\gamma_3 + \gamma_5)}{(1 - \hat{\alpha}^2)\gamma_0}$$

which yields, after substitution for plim $\hat{\alpha}$ and the γ's and some tedious manipulation,

$$\operatorname{plim} \hat{\rho} = \rho - \frac{\rho\alpha^6(1 - \alpha^2)(1 - \rho^2)}{(1 + \alpha^4\rho)(1 - \alpha^6\rho^2)}.$$

REFERENCES

[1] ANDERSON, R. L.: "Distribution of the Serial Correlation Coefficient," *Annals of Mathematical Statistics*, 13 (1942), 1–13.

[2] ANDERSON, T. W.: "On the Theory of Testing Serial Correlation," *Skandinavisk Aktuarietidskrift*, 31 (1948), 88–116.

[3] DURBIN, J.: "Testing for Serial Correlation in Systems of Simultaneous Regression Equations," *Biometrika*, 44 (1957), 370–377.

[4] ———: "Tests for Serial Correlation in Regression Analysis Based on the Periodogram of Least Squares Residuals," *Biometrika*, 56 (1969), 1–15.

[5] ———: "Testing for Serial Correlation in Least Squares Regression When Some of the Regressors Are Lagged Dependent Variables," *Econometrica*, 38 (1970), 410–421.

[6] DURBIN, J., AND G. S. WATSON: "Testing for Serial Correlation in Least Squares Regression I," *Biometrika*, 37 (1950), 409–428.

[7] ———: "Testing for Serial Correlation in Least Squares Regression II," *Biometrika*, 38 (1951), 159–178.

[8] DURBIN, J., AND G. S. WATSON: "Testing for Serial Correlation in Least Squares Regression III," *Biometrika*, 58 (1971), 1–19.
[9] FISHER, G. R.: "Quarterly Dividend Behaviour," in *The Econometric Study of the United Kingdom*, Ed. by K. Hilton and D. F. Heathfield. London: Macmillan, 1970, pp. 149–181.
[10] FULLER, W. A., AND J. E. MARTIN: "The Effects of Autocorrelated Errors on the Statistical Estimation of Distributed Lag Models," *Journal of Farm Economics*, 43 (1961), 71–82.
[11] GRILICHES, Z.: "Distributed Lags: a Survey," *Econometrica*, 35 (1967), 16–49.
[12] HALL, R. E., J. BRODE, AND F. C. RIPLEY: "Time Series Processor User's Manual," Discussion Paper 24, Harvard Institute of Economic Research, Harvard University, 1968.
[13] HANNAN, E. J., AND R. D. TERRELL: "Testing for Serial Correlation after Least Squares Regression," *Econometrica*, 36 (1968), 133–150.
[14] HENDRY, D. F.: "Users' Manual for the Estimation of Linear Equations with Lagged Dependent Variables and Autoregressive Errors," unpublished.
[15] ———: "The Estimation of Economic Models with Autoregressive Errors," unpublished Ph.D. thesis, London University, 1970.
[16] HENSHAW, R. C.: "Testing Single-Equation Least Squares Regression Models for Autocorrelated Disturbances," *Econometrica*, 34 (1966), 646–660.
[17] IMHOF, J. P.: "Computing the Distribution of Quadratic Forms in Normal Variables," *Biometrika*, 48 (1961), 419–426.
[18] KOERTS, J., AND A. P. J. ABRAHAMSE: *On the Theory and Application of the General Linear Model.* Rotterdam: Rotterdam University Press, 1969.
[19] LIPSEY, R. G., AND J. M. PARKIN: "Incomes Policy: A Re-appraisal," *Econometrica*, 37 (1970), 115–138.
[20] MALINVAUD, E.: *Statistical Methods of Econometrics.* Amsterdam: North-Holland, 1966.
[21] MILLS, E. S.: *Price, Output and Inventory Policy.* New York: Wiley, 1962.
[22] NADIRI, M. I.: "The Determinants of Real Cash Balances in the U.S. Total Manufacturing Sector," *Quarterly Journal of Economics*, 83 (1969), 173–196.
[23] NERLOVE, M., AND K. F. WALLIS: "Use of the Durbin-Watson Statistic in Inappropriate Situations," *Econometrica*, 34 (1966), 235–238.
[24] VON NEUMANN, J.: "Distribution of the Ratio of the Mean Square Successive Difference to the Variance," *Annals of Mathematical Statistics*, 12 (1941), 367–395.
[25] POWELL, M. J. D.: "A Method for Minimising a Sum of Squares of Non-Linear Functions Without Calculating Derivatives," *Computer Journal*, 7 (1965), 303–307.
[26] SARGAN, J. D.: "Wages and Prices in the United Kingdom: a Study in Econometric Methodology," in *Econometric Analysis for National Economic Planning*, Ed. by P. E. Hart et al. (*Proceedings of the 16th Symposium of the Colston Research Society*). London: Butterworths, 1964, pp. 25–54.
[27] THEIL, H., AND A. L. NAGAR: "Testing the Independence of Regression Disturbances," *Journal of the American Statistical Association*, 63 (1961), 242–251.
[28] THOMAS, J. J., AND K. F. WALLIS: "Seasonal Variation in Regression Analysis," *Journal of the Royal Statistical Society A*, 134 (1971), 54–72.