Econometrica, Vol. 46, No. 6 (November, 1978)

TESTING FOR HIGHER ORDER SERIAL CORRELATION IN REGRESSION EQUATIONS WHEN THE REGRESSORS INCLUDE LAGGED DEPENDENT VARIABLES

BY L. G. GODFREY[1]

There has been increasing concern recently over the use of the simple first order Markov form to model error autocorrelation in regression analysis. The consequence of misspecifying the error model will be especially serious when the regressors include lagged values of the dependent variable. The purpose of this paper is to develop Lagrange multiplier tests of the assumed error model against specified ARMA alternatives. It is shown that all of the tests can be regarded as asymptotic tests of the significance of a coefficient of determination, and a table is provided which gives details of two general tests and several special cases.

1. INTRODUCTION

THIS PAPER IS a sequel to Godfrey [**6**] and considers the problem of testing the specification of the assumed model for the error term of a regression equation which includes lagged values of the dependent variable in the regressors. The approach adopted is to set up a hypothesis testing situation in which the assumed error model provides the null hypothesis and the alternative is obtained by increasing the order (and possibly generalizing the form) of the assumed process. The test procedures described below relate to autoregressive, moving average and mixed-autoregressive–moving-average processes, and a unified approach is obtained by employing Silvey's [**13**] Lagrange multiplier technique.

Tests of the type derived below are to be recommended to applied workers since error processes are often modelled by low order (usually first order) autoregressive schemes and if such models are inappropriate, then the estimators of the parameters of the dynamic regression model will be inconsistent. It should also be noted that the Durbin–Watson statistic and the Box–Pierce [**1**] Q statistic do not provide valid diagnostic checks in situations of the type considered in this paper, despite the fact that they are sometimes reported.

Although simultaneous equation models are not considered below, it is straightforward to generalize the test procedures in much the same way that Godfrey [**5**] generalized Durbin's [**2**] h test.

The plan of the paper is as follows. Section 2 contains details of the notation and assumptions employed other than those used in Godfrey [**6**]. Two general test procedures are derived in Sections 3 and 4, and a simple interpretation of these procedures is provided in Section 5. Section 5 also contains a tabular summary of the test procedures for the general cases and several special cases of interest to applied workers. Some concluding remarks are offered in Section 6.

[1] The author is grateful to Christine Godfrey and the referees for helpful comments.

2. THE MODEL

The notation and assumptions of Godfrey [**6**] are employed in this paper. The regression equation is then

$$(1) \qquad \underline{y} = X\underline{\beta} + \underline{u},$$

and the error schemes considered below are all special cases of the ARMA(p, q) model

$$(2) \qquad M(\underline{\rho})\underline{u} = N(\underline{\mu})\underline{\varepsilon}\ldots \text{ARMA}(p, q).$$

The most popular specification of (2) in applied work is the first order autoregressive (AR(1) or ARMA(1, 0)) model, although higher order autoregressive models are sometimes employed. A few studies have presented results based upon estimates derived under the assumption that the errors were generated by a pure moving average (MA(q) or ARMA$(0, q)$) process (e.g. see Trivedi [**14**]), but the general ARMA model is very rarely used. Misspecification of the error process is, of course, important and Engle [**3**] has shown that an AR(1) approximation to a true AR(2) process can actually be inferior to an AR(0) approximation.[2] It is therefore important to be able to check that the specification of the error model is consistent with the sample data and that there is no significant serial correlation in the estimates of the ε_t residuals.

The two general tests considered in Sections 3 and 4 involve the following combinations of null and alternative hypotheses: (i) H_0: the $\{u_t\}$ are ARMA (p, q) and H_1: the $\{u_t\}$ are ARMA$(p+r, q)$; and (ii) H_0: the $\{u_t\}$ are ARMA (p, q) and H_1: the $\{u_t\}$ are ARMA$(p, q+r)$.[3] (It is not possible to test H_0: the $\{u_t\}$ are ARMA(p, q) against H_1: the $\{u_t\}$ are ARMA$(p+r,\ q+s)$, $r,\ s>0$; see Hannan [**7**, pp. 388–389 and pp. 409–414].) Almost all applied econometric research involving estimation allowing for serial correlation is, however, based upon either the pure AR model or the pure MA model; so that various special cases obtained by putting either p or q equal to zero are considered in Section 5.

As in Godfrey [**6**], caret denotes the restricted maximum likelihood estimators (MLE) obtained when the null hypothesis is imposed, and tilde denotes unrestricted MLE, so that, e.g., $\hat{\underline{\beta}}$ of Sections 3 and 4 is the estimate of $\underline{\beta}$ obtained by maximizing the likelihood under the assumption that the $\{u_t\}$ are ARMA(p, q).

The following additional notation will be employed:

$$(3) \qquad U_j = (\underline{u}_1 \ldots \underline{u}_j), \quad \text{where} \quad \underline{u}_j = L^j\underline{u}, \quad 0<j<T;$$

$$(4) \qquad E_j = (\underline{\varepsilon}_1 \ldots \underline{\varepsilon}_j), \quad \text{where} \quad \underline{\varepsilon}_j = L^j\underline{\varepsilon}, \quad 0<j<T;$$

$$(5) \qquad \hat{E}_j = (\hat{\underline{\varepsilon}}_1 \ldots \hat{\underline{\varepsilon}}_j), \quad \text{where} \quad \hat{\underline{\varepsilon}}_j = L^j\hat{\underline{\varepsilon}}, \quad 0<j<T;$$

[2] There is some evidence that misspecification of the order of the error process is more important than misspecification of its form (AR or MA); see Hendry and Trivedi [**9**] and Hendry [**8**].

[3] It should be noted that if the regressors of (1) do not contain lagged dependent variables, then Pierce's [**11**] test is an asymptotically valid alternative to the large sample tests of this paper, provided that r is large.

(6) $K_u = (\underline{\hat{u}}_{p+1} \ldots \underline{\hat{u}}_{p+r});$

(7) $K_\varepsilon = (\underline{\hat{\varepsilon}}_{q+1} \ldots \underline{\hat{\varepsilon}}_{q+r});$

(8) $R(\underline{\alpha}) = \alpha_1 L^{p+1} + \ldots + \alpha_r L^{p+r};$

and

(9) $S(\underline{\alpha}) = \alpha_1 L^{q+1} + \ldots + \alpha_r L^{q+r}.$

3. TESTING ARMA(p, q) AGAINST ARMA$(p+r, q)$

The error models under null and alternative hypotheses can be written as

(2) $M(\underline{\rho})\underline{u} = N(\underline{\mu})\underline{\varepsilon} \ldots \text{ARMA}(p, q)$

and

(10) $[M(\underline{\rho}) + R(\underline{\alpha})]\underline{u} = N(\underline{\mu})\underline{\varepsilon} \ldots \text{ARMA}(p+r, q).$

The log likelihood $L_1(\underline{\theta})$, $\underline{\theta}' = (\underline{\beta}' \ \underline{\mu}' \ \underline{\rho}' \ \underline{\alpha}')$ (note the ordering of the subvectors of $\underline{\theta}$), for the models consisting of (1) and (10) is

(11) $L_1(\underline{\theta}) = -\frac{1}{2} \ln (\underline{\varepsilon}(\underline{\theta})'\underline{\varepsilon}(\underline{\theta})/T),$

where $\underline{\varepsilon}(\underline{\theta}) = N(\underline{\mu})^{-1}[M(\underline{\rho}) + R(\underline{\alpha})](\underline{y} - X\underline{\beta})$; and the asymptotic variance-covariance matrix of $\sqrt{T}(\underline{\tilde{\theta}} - \underline{\theta})$ is

(12) $\Omega_1 = \sigma_\varepsilon^2 \operatorname{plim} T\{A_{ij}\}^{-1} \quad (i, j = 1, 3)$

where

$$A_{11} = X'[M(\underline{\rho}) + R(\underline{\alpha})]'(N(\underline{\mu})^{-1})'N(\underline{\mu})^{-1}[M(\underline{\rho}) + R(\underline{\alpha})]X,$$

$$A_{12} = A_{21}' = X'[M(\underline{\rho}) + R(\underline{\alpha})]'(N(\underline{\mu})^{-1})'N(\underline{\mu})^{-1}E_q,$$

$$A_{13} = A_{31}' = -X'[M(\underline{\rho}) + R(\underline{\alpha})]'(N(\underline{\mu})^{-1})'N(\underline{\mu})^{-1}U_{p+r},$$

$$A_{22} = E_q'(N(\underline{\mu})^{-1})'N(\underline{\mu})^{-1}E_q,$$

$$A_{23} = A_{32}' = -E_q'(N(\underline{\mu})^{-1})'N(\underline{\mu})^{-1}U_{p+r},$$

and

$$A_{33} = (U_{p+r})'(N(\underline{\mu})^{-1})'N(\underline{\mu})^{-1}U_{p+r}.$$

It can be shown that

(13) $\underline{\hat{\lambda}} = -(N(\underline{\hat{\mu}})^{-1}K_u)'\underline{\hat{\varepsilon}}/\underline{\hat{\varepsilon}}'\underline{\hat{\varepsilon}} = -(N(\underline{\hat{\mu}})^{-1}K_u)'\underline{\hat{\varepsilon}}/(T\hat{\sigma}_\varepsilon^2)$

where $\hat{\sigma}_\varepsilon^2 = \underline{\hat{\varepsilon}}'\underline{\hat{\varepsilon}}/T$ is an estimate of σ_ε^2 based upon the ARMA(p, q) residuals $\{\hat{\varepsilon}_t\}$, and that

(14) $W = \hat{\sigma}_\varepsilon^2 T[(N(\underline{\hat{\mu}})^{-1}K_u)'(I - Z_{AM}(Z_{AM}'Z_{AM})^{-1}Z_{AM}')(N(\underline{\hat{\mu}})^{-1}K_u)]^{-1}$

where $Z_{AM} = N(\underline{\hat{\mu}})^{-1}[M(\underline{\hat{\rho}})X : \hat{E}_q : \hat{U}_p]$.

The test statistic for testing (2) against (10) is obtained by computing the quadratic form $(\sqrt{T}\hat{\underline{\lambda}})'W(\sqrt{T}\hat{\underline{\lambda}})$, where $\hat{\underline{\lambda}}$ and W are given by (13) and (14), respectively. The matrix $(Z'_{AM}Z_{AM})^{-1}$ appearing in (14) can, of course, be replaced by $\hat{\sigma}_\varepsilon^{-2}V_{AM}$, where V_{AM} is the estimate of the asymptotic variance-covariance matrix of $(\hat{\underline{\beta}}'\ \hat{\underline{\mu}}'\ \hat{\underline{\rho}})'$.[4]

It is worth noting that it is not necessary to invert the T by T matrix $N(\hat{\underline{\mu}})$ when computing the test statistic since asymptotically valid approximations can be obtained by the use of a simple recurrence relationship (see Phillips [**10**]). The computational burden can be further reduced by noting that $N(\hat{\underline{\mu}})^{-1}\hat{U}_p$ and $N(\hat{\underline{\mu}})^{-1}K_u$ together make up $N(\hat{\underline{\mu}})^{-1}\hat{U}_{p+r}$ and that the jth column of $N(\hat{\underline{\mu}})^{-1}\hat{U}_{p+r}$ $(N(\hat{\underline{\mu}})^{-1}\hat{E}_q)$ is simply $N(\hat{\underline{\mu}})^{-1}\hat{\underline{u}}$ $(N(\hat{\underline{\mu}})^{-1}\hat{\underline{\varepsilon}})$ lagged j periods so that only this vector need be computed.

4. TESTING ARMA(p, q) AGAINST ARMA$(p, q+r)$

The error models under null and alternative hypotheses will be written as

$$(2)\qquad M(\underline{\rho})\underline{u} = N(\underline{\mu})\underline{\varepsilon}\ldots\text{ARMA}(p, q)$$

and

$$(15)\qquad M(\underline{\rho})\underline{u} = [N(\underline{\mu}) + S(\underline{\alpha})]\underline{\varepsilon}\ldots\text{ARMA}(p, q+r).$$

The log likelihood $L_2(\underline{\theta})$, $\underline{\theta}' = (\underline{\beta}'\ \underline{\rho}'\ \underline{\mu}'\ \underline{\alpha}')$ (note the ordering of the subvectors of $\underline{\theta}$) for the model consisting of (1) and (15) is

$$(16)\qquad L_2(\underline{\theta}) = -\tfrac{1}{2}\ln\left(\underline{\varepsilon}(\underline{\theta})'\underline{\varepsilon}(\underline{\theta})/T\right),$$

with $\underline{\varepsilon}(\underline{\theta}) = [N(\underline{\mu}) + S(\underline{\alpha})]^{-1}M(\underline{\rho})(\underline{y} - X\underline{\beta})$, and the asymptotic variance-covariance matrix of $\sqrt{T}(\tilde{\underline{\theta}} - \underline{\theta})$ is

$$(17)\qquad \Omega_2 = \sigma_\varepsilon^2 \operatorname{plim} T\{B_{ij}\}^{-1} \quad (i, j = 1, 3)$$

where

$$B_{11} = X'M(\underline{\rho})'([N(\underline{\mu}) + S(\underline{\alpha})]^{-1})'[N(\underline{\mu}) + S(\underline{\alpha})]^{-1}M(\underline{\rho})X,$$

$$B_{12} = B'_{21} = -X'M(\underline{\rho})'([N(\underline{\mu}) + S(\underline{\alpha})]^{-1})'[N(\underline{\mu}) + S(\underline{\alpha})]^{-1}U_p,$$

$$B_{13} = B'_{31} = X'M(\underline{\rho})'([N(\underline{\mu}) + S(\underline{\alpha})]^{-1})'[N(\underline{\mu}) + S(\underline{\alpha})]^{-1}E_{q+r},$$

$$B_{22} = U'_p([N(\underline{\mu}) + S(\underline{\alpha})]^{-1})'[N(\underline{\mu}) + S(\underline{\alpha})]^{-1}U_p,$$

$$B_{23} = B'_{32} = -U'_p([N(\underline{\mu}) + S(\underline{\alpha})]^{-1})'[N(\underline{\mu}) + S(\underline{\alpha})]^{-1}E_{q+r},$$

and

$$B_{33} = (E_{q+r})'([N(\underline{\mu}) + S(\underline{\alpha})]^{-1})'[N(\underline{\mu}) + S(\underline{\alpha})]^{-1}E_{q+r}.$$

[4] See Sargan [**12**] for a discussion of a large sample test for the $(q = 0, p = r = 1)$ case for equations estimated by instrumental variables, and Durbin [**2**] for the correction of a small error contained in Sargan's derivation of his test statistic.

It can be shown that

$$\hat{\underline{\lambda}} = (N(\hat{\underline{\mu}})^{-1}K_\varepsilon)'\hat{\underline{\varepsilon}}/\hat{\underline{\varepsilon}}'\hat{\underline{\varepsilon}} = (N(\hat{\underline{\mu}})^{-1}K_\varepsilon)'\hat{\underline{\varepsilon}}/(T\hat{\sigma}_\varepsilon^2) \tag{18}$$

where $\hat{\sigma}_\varepsilon^2 = \hat{\underline{\varepsilon}}'\hat{\underline{\varepsilon}}/T$ is an estimate of σ_ε^2 based upon the ARMA(p, q) residuals $\{\hat{\varepsilon}_t\}$, and that

$$W = \hat{\sigma}_\varepsilon^2 T[(N(\hat{\underline{\mu}})^{-1}K_\varepsilon)'(I - Z_{AM}(Z'_{AM}Z_{AM})^{-1}Z'_{AM})(N(\hat{\underline{\mu}})^{-1}K_\varepsilon)]^{-1} \tag{19}$$

where $Z_{AM} = N(\hat{\underline{\mu}})^{-1}[M(\hat{\underline{\rho}})X : \hat{U}_p : \hat{E}_q]$. As noted above, $(Z'_{AM}Z_{AM})^{-1}$ can be replaced by $\hat{\sigma}^{-2}V_{AM}$.

It follows that in order to obtain a test asymptotically equivalent to the LR test of $H_0 : \underline{\alpha} = \underline{0}$, i.e. of the assumption that the $\{u_t\}$ errors of (1) are ARMA(p, q), rather than ARMA$(p, q+r)$, the expressions of (18) and (19) should be used to calculate the Lagrange multiplier test statistic $(\sqrt{T}\hat{\underline{\lambda}})'W(\sqrt{T}\hat{\underline{\lambda}})$.[5] (The matrix $N(\hat{\underline{\mu}})^{-1}$ appears in equations (18) and (19), and remarks similar to those in the last paragraph of the previous section apply in this case.)

5. SOME FURTHER RESULTS

As mentioned above, almost all empirical work involving the estimation of linear regression equations with autocorrelated errors is based upon either the pure AR model or the pure MA model, with the former being much more popular. It, therefore, seems worthwhile to present test procedures for the special cases obtained by putting either p or q equal to zero since these special cases may be of greater practical interest than the tests of Section 3 and 4. In order to obtain a simple presentation of the various tests, it will be useful to derive an alternative interpretation of the Lagrange multiplier tests of the previous two sections. The derivations are very similar for the two general tests and so only the first will be considered.

Let $P(Z_{AM}) = I - Z_{AM}(Z'_{AM}Z_{AM})^{-1}Z'_{AM}$, so that the W matrix of Section 3 can be expressed as

$$W = T\hat{\sigma}_\varepsilon^2[(N(\hat{\underline{\mu}})^{-1}K_u)'P(Z_{AM})(N(\hat{\underline{\mu}})^{-1}K_u)]^{-1}. \tag{20}$$

Now the first order conditions for constrained maximization of the relevant likelihood are:

$$[(\partial\underline{\varepsilon}(\underline{\theta})/\partial\underline{\beta})'\underline{\varepsilon}(\underline{\theta})]_{\underline{\theta}=\hat{\underline{\theta}}} = -[N(\hat{\underline{\mu}})^{-1}M(\hat{\underline{\rho}})X]'\hat{\underline{\varepsilon}} = \underline{0}, \tag{21}$$

$$[(\partial\underline{\varepsilon}(\underline{\theta})/\partial\underline{\rho})'\underline{\varepsilon}(\underline{\theta})]_{\underline{\theta}=\hat{\underline{\theta}}} = [N(\hat{\underline{\mu}})^{-1}\hat{U}_p]'\hat{\underline{\varepsilon}} = \underline{0}, \tag{22}$$

and

$$[(\partial\underline{\varepsilon}(\underline{\theta})/\partial\underline{\mu})'\underline{\varepsilon}(\underline{\theta})]_{\underline{\theta}=\hat{\underline{\theta}}} = -[N(\hat{\underline{\mu}})^{-1}\hat{E}_q]'\hat{\underline{\varepsilon}} = \underline{0}. \tag{23}$$

Equations (21)–(23) clearly imply that $Z'_{AM}\hat{\underline{\varepsilon}} = \underline{0}$, since $Z_{AM} = N(\hat{\underline{\mu}})^{-1}[M(\hat{\underline{\rho}})X : \hat{E}_q : \hat{U}_p]$, and so $P(Z_{AM})\hat{\underline{\varepsilon}} = \hat{\underline{\varepsilon}}$. It follows that the Lagrange

[5] See Fitts [**4**] for a discussion of an asymptotic test for the $(p = 0,\ q = r = 1)$ case based upon Durbin's [**2**] general method for tests of specification.

multiplier vector of (13) can be written as

$$\hat{\underline{\lambda}} = -(N(\hat{\underline{\mu}})^{-1}K_u)'P(Z_{AM})\hat{\underline{\varepsilon}}/\hat{\underline{\varepsilon}}'\hat{\underline{\varepsilon}} \tag{24}$$

and that the Lagrange multiplier test against ARMA$(p+r, q)$ errors is simply the explained sum of squares from the OLS regression of $P(Z_{AM})\hat{\underline{\varepsilon}}$ on $P(Z_{AM})N(\hat{\underline{\mu}})^{-1}K_u$ divided by $\hat{\underline{\sigma}}_{\varepsilon}^2$. If this explained sum of squares is denoted by S_1 and the test statistic is denoted by ϕ, then

$$\begin{aligned}\phi &= S_1/\hat{\sigma}_{\varepsilon}^2\\ &= T(S_1/\hat{\underline{\varepsilon}}'\hat{\underline{\varepsilon}})\\ &= T(\hat{\underline{\varepsilon}}'P(Z_{AM})\hat{\underline{\varepsilon}} - S_2)/\hat{\underline{\varepsilon}}'\hat{\underline{\varepsilon}}\\ &= T(\hat{\underline{\varepsilon}}'\hat{\underline{\varepsilon}} - S_2)/\hat{\underline{\varepsilon}}'\hat{\underline{\varepsilon}} \quad (\text{since } P(Z_{AM})\hat{\underline{\varepsilon}} = \hat{\underline{\varepsilon}})\\ &= T(1 - S_2/\hat{\underline{\varepsilon}}'\hat{\underline{\varepsilon}}),\end{aligned} \tag{25}$$

where S_2 is the error sum of squares from the regression of $P(Z_{AM})\hat{\underline{\varepsilon}}$ on $P(Z_{AM})(N(\hat{\underline{\mu}})^{-1}K_u)$. The Frisch–Waugh result implies that S_2 is also the error sum of squares of the regression of $\hat{\underline{\varepsilon}}$ on Z_{AM} and $N(\hat{\underline{\mu}})^{-1}K_u$, and so

$$\phi = T(R_{\phi}^2) \tag{26}$$

where R_{ϕ}^2 is the coefficient of determination for the OLS regression of $\hat{\underline{\varepsilon}}$ on $N(\hat{\underline{\mu}})^{-1}K_u$ and Z_{AM}, i.e. upon $N(\hat{\underline{\mu}})^{-1}[M(\hat{\underline{\rho}})X : \hat{E}_q : \hat{U}_{p+r}]$. The tests of Sections 3 and 4 can, therefore, be interpreted as asymptotic tests of the significance of a coefficient of determination with the regressor set for the ARMA$(p, q+r)$ alternative being

$$N(\hat{\underline{\mu}})^{-1}[M(\hat{\underline{\rho}})X : \hat{U}_p : \hat{E}_{q+r}]. \tag{27}$$

The tests for the special cases obtained by putting either p or q equal to zero are now easy to obtain. For example, when testing H_0: the $\{u_t\}$ are AR(p) against H_1: the $\{u_t\}$ are ARMA(p, r), the regressor set is obtained by putting $N(\hat{\underline{\mu}}) = I$ and $q = 0$ in expression (27) and so is $[M(\hat{\underline{\rho}})X: \hat{U}_p: \hat{E}_r]$. The test statistic for this special case can be written as

$$\phi^* = \hat{\underline{\varepsilon}}'\hat{E}_r[\hat{E}_r'P(Z_A)\hat{E}_r]^{-1}\hat{E}_r'\hat{\underline{\varepsilon}}/\hat{\sigma}_{\varepsilon}^2 \tag{28}$$

where $Z_A = [M(\hat{\underline{\rho}})X : \hat{U}_p]$ (the matrix $M(\hat{\underline{\rho}})X$ is, of course, an estimated autoregressive transform of X).

The test procedures can then be summarized as follows. First, impose the null hypothesis and obtain the constrained estimates. Next, perform an OLS regression using the appropriate dependent variable and set of regressors (see Table I), and then compare T times the R^2 statistic to the selected critical value for a χ_r^2 variate. Significantly large values of the test statistic indicate that the null hypothesis is not consistent with the sample data. (Several of the regressor sets of Table I involve $N(\hat{\underline{\mu}})^{-1}$ and readers are reminded of the remarks contained in the last paragraph of section 3.) For the sake of completeness, tests of the null hypothesis of serial independence ($p = q = 0$) are included in Table I and the dependent variable for these tests is the vector of OLS residuals. It should be noted that Durbin [**2**, p. 420] proposed a procedure for testing the null hypothesis of ARMA(0, 0) errors against the ARMA(1, 0) alternative which also involved the regression of $\underline{\hat{u}}$ on $[X:\underline{\hat{u}}_1]$. Durbin's test was not based upon the coefficient of determination, but instead upon the significance of the partial correlation between $\underline{\hat{u}}$ and $\underline{\hat{u}}_1$, with the linear influence of X removed from both variables. There is, however, no real difference between the Lagrange multiplier procedure and Durbin's test since the variables of X are uncorrelated with $\underline{\hat{u}}$ and so the square of the partial correlation between $\underline{\hat{u}}$ and $\underline{\hat{u}}_1$ equals the R^2 statistic.

The test statistics can, of course, be calculated directly from the appropriate formulae, e.g. (28) gives the statistic for testing a simple autoregressive error model against a mixed autoregressive-moving average alternative.

TABLE I

SUMMARY OF LAGRANGE MULTIPLIER TESTS OF SPECIFICATION OF ERROR MODELS

Null Hypothesis	Alternative Hypothesis	Dependent Variable	Regressor Set
ARMA(p, q)	ARMA($p+r, q$)	$\underline{\hat{\varepsilon}}$	$N(\underline{\hat{\mu}})^{-1}[M(\underline{\hat{\rho}})X:\hat{E}_q : \hat{U}_{p+r}]$
ARMA(p, q)	ARMA($p, q+r$)	$\underline{\hat{\varepsilon}}$	$N(\underline{\hat{\mu}})^{-1}[M(\underline{\hat{\rho}})X:\hat{U}_p:\hat{E}_{q+r}]$
ARMA(p, 0)	ARMA($p+r$, 0)	$\underline{\hat{\varepsilon}}$	$[M(\underline{\hat{\rho}})X:\hat{U}_{p+r}]$
ARMA(p, 0)	ARMA(p, r)	$\underline{\hat{\varepsilon}}$	$[M(\underline{\hat{\rho}})X:\hat{U}_p:\hat{E}_r]$
ARMA(0, q)	ARMA(r, q)	$\underline{\hat{\varepsilon}}$	$N(\underline{\hat{\mu}})^{-1}[X:\hat{E}_q:\hat{U}_r]$
ARMA(0, q)	ARMA(0, $q+r$)	$\underline{\hat{\varepsilon}}$	$N(\underline{\hat{\mu}})^{-1}[X:\hat{E}_{q+r}]$
ARMA(0, 0)	ARMA(r, 0)	$\underline{\hat{u}}$	$[X:\hat{U}_r]$
ARMA(0, 0)	ARMA(0, r)	$\underline{\hat{u}}$	$[X:\hat{U}_r]$

TEST: Regress the dependent variable on the regressor set and obtain $\phi = (T$ times R^2 statistic from regression). The variable ϕ is asymptotically distributed as χ^2_r under H_0 and the sample value of ϕ should be compared to the pre-selected critical value.

6. CONCLUDING REMARKS

The purpose of this paper has been to provide a unified approach to the large sample theory of testing the specification of the error process in dynamic regression equations. There are, however, certain practical considerations which have not been previously mentioned. Firstly, although the test statistics derived above are more complicated than those for testing the null hypothesis of serial independence (compare Godfrey [**6**]), once appropriate subroutines have been added to estimation programs, the marginal cost of computing them is negligible. Secondly, the tests above cannot be used to choose between an AR(p) model and a MA(q) model, since the corresponding hypotheses are non-nested (an attempt to make such a choice might be made by testing both of these models against an ARMA(p, q) model, but this procedure can lead to inconclusive results). Thirdly, r, the number of parameters constrained to be zero under the null hypothesis, must be selected. If too small a value is assumed, then the test may lack power, while if a very high value is adopted then a correspondingly large sample will probably be required to justify appeal to the asymptotic theory. Finally, the behavior of the Lagrange multiplier tests in small samples merits investigation.

University of York

Manuscript received May, 1977; revision received December, 1977.

REFERENCES

[1] BOX, G. E. P., AND D. A. PIERCE: "Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models," *Journal of the American Statistical Association*, 65 (1970), 1509-1526.

[2] DURBIN, J.: "Testing for Serial Correlation in Least Squares Regression When Some of the Regressors are Lagged Dependent Variables," *Econometrica*, 38 (1970), 410-421.

[3] ENGLE, R. F.: "Specification of the Disturbance Term for Efficient Estimation," *Econometrica*, 42 (1974), 135-146.

[4] FITTS, J.: "Testing for Autocorrelation in the Autoregressive Moving Average Error Model," *Journal of Econometrics*, 1 (1973), 363-376.

[5] GODFREY, L. G.: "Testing for Serial Correlation in Dynamic Simultaneous Equation Models," *Econometrica*, 44 (1976), 1077-1084.

[6] ———: "Testing Against General Autoregressive and Moving Average Error Models When the Regressors Include Lagged Dependent Variables," *Econometrica*, 46 (1978), 1293-1301.

[7] HANNAN, E. J.: *Multiple Time Series*. New York: Wiley, 1970.

[8] HENDRY, D. F.: "On the Time Series Approach to Econometric Model Building," Discussion Paper No. A8, London School of Economics, 1976.

[9] HENDRY, D. F., AND P. K. TRIVEDI: "Maximum Likelihood Estimation of Difference Equations with Moving Average Errors: A Simulation Study," *The Review of Economic Studies*, 39 (1972), 117-145.

[10] PHILLIPS, A. W.: "Estimation of Systems of Difference Equations with Moving Average Disturbances," paper presented to the San Francisco Meeting of the Econometric Society, 1966.

[11] PIERCE, D. A.: "Distribution of Residual Autocorrelations in the Regression Model with Autoregressive-Moving Average Errors," *Journal of the Royal Statistical Society*, Series B, 33 (1971), 140-146.

[12] 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, G. Mills and J. K. Whitaker. London: Butterworths Scientific Publications, 1964.

[13] SILVEY, S. D.: "The Lagrange Multiplier Test," *Annals of Mathematical Statistics*, 30 (1959), 389-407.

[14] TRIVEDI, P. K.: "Retail Inventory Investment Behavior," *Journal of Econometrics*, 1 (1973), 61-80.