The Annals of Statistics
1981, Vol. 9, No. 5, 974–986

AN APPROACH TO TESTING LINEAR TIME SERIES MODELS

BY D. S. POSKITT AND A. R. TREMAYNE

University of York

The purpose of this paper is to develop diagnostic tests for open-loop transfer function models with autoregressive-moving average stochastic disturbances using the efficient scores procedure. The tests proposed are asymptotically equivalent to those based upon the likelihood ratio principle and have the advantage that they do not involve a heavy computational burden. Consideration is given both to the standard case of nonsingular information matrices and to the situation obtaining when there are identifiability problems and application of conventional large sample test procedures is not feasible. Relationships between score tests and portmanteau tests in time series analysis are also investigated. Some simulation evidence on the finite sample behaviour of the tests is presented and it is seen that the tests of this paper perform well.

1. Introduction. In this paper we are concerned with the linear, time-invariant, open-loop transfer function model

$$y(t) = \frac{\alpha(B)}{\beta(B)} x(t) + \eta(t), \qquad t = 0, \pm 1, \pm 2, \cdots, \tag{1.1}$$

where B is the backward shift operator. We assume that $\{x(t)\}$ is a zero mean, nondeterministic, stationary and ergodic process and that

$$\alpha(z) = \alpha_0 + \textstyle\sum_{i=1}^{q_\tau} \alpha_i z^i$$

and

$$\beta(z) = 1 - \textstyle\sum_{i=1}^{p_\tau} \beta_i z^i$$

have no common roots. The eta process is a stochastic disturbance independent of $\{x(t)\}$ specified as

$$\phi(B)\eta(t) = \theta(B)\epsilon(t)$$

where

$$\phi(z) = 1 - \textstyle\sum_{i=1}^{p} \phi_i z^i$$

and

$$\theta(z) = 1 - \textstyle\sum_{i=1}^{q} \theta_i z^i$$

have no roots in common and $\{\epsilon(t)\}$ is a Gaussian white noise process with zero mean and variance σ_ϵ^2, i.i.d. $N(0, \sigma_\epsilon^2)$. The Gaussian assumption can be replaced by weaker regularity conditions, as in Hannan and Heyde (1972), without affecting the results of the paper. It is further assumed that the polynomials $\alpha(z)$, $\beta(z)$, $\phi(z)$ and $\theta(z)$ satisfy the usual conditions for the model to be realisable, stationary and invertible, as discussed in Box and Jenkins (1976, Chapter 10). The polynomial ratio $\alpha(z)/\beta(z)$ is commonly referred to as the transfer function and hence we introduce the acronym ARMAT$((q_\tau, p_\tau)(p, q))$ for this model.

Received February 7, 1980; revised December 22, 1980.

AMS 1970 *subject classifications.* Primary 62M10; secondary 62F05.

Key words and phrases. Autoregressive-moving average, identifiability, pure significance test, score test, transfer function.

The ARMAT model has found application in a number of fields, for example control engineering and econometrics, where it is known as a rational distributed lag model. It has received attention in the statistical literature and the papers of Pierce (1972a,b) and Haugh and Box (1977) develop the familiar iterative process of identification, estimation and diagnostic checking for such dynamic-disturbance time series models. In particular, Pierce (1972b) extended the work of Box and Pierce (1970) and developed portmanteau type tests which can provide a useful diagnostic check of the adequacy of a fitted ARMAT model.

In this paper, tests of linear time series models based upon the score or Lagrangian multiplier procedure due to Rao (1948) and Silvey (1959) are developed. These are obtained by setting up a hypothesis testing framework in which a fitted model provides the null hypothesis and the alternative is obtained by generalising the specification. The tests possess similar asymptotic properties to those based upon the likelihood ratio principle. Not only is the situation where the information matrix is of full rank investigated, but also that obtaining when there are identifiability problems present and application of conventional large sample test procedures is not straightforward. Such considerations lead to an interpretation of the tests proposed as pure significance tests, Cox and Hinkley (1974, Chapter 3). Furthermore, in the third section of the paper the tests resulting from application of the score test principle are compared to those of the previous paragraph and some interesting equivalences are found.

The tests advocated have certain attractions for the practicing time series analyst in that the score test statistic can be easily computed and the researcher need only fit the most parsimonious model of interest. The practical usefulness of these tests is also contingent upon their performance in small and moderate samples and, in the final section of the paper, some simulation evidence on these matters is provided to assess the relative merits of score and portmanteau tests.

2. Hypothesis testing. Consider the problem of testing the null hypothesis $H((q_\tau, p_\tau)(p, q))$, that $\{y(t)\}$ and $\{x(t)\}$ are related by an ARMAT $((q_\tau, p_\tau)(p, q))$ model, against alternatives of the form $H((q_\tau + s_\tau, p_\tau + r_\tau)(p + r, q + s))$, $s_\tau, r_\tau, r, s \geq 0$. Assume that a realisation of T observations on $\{y(t)\}$ and $\{x(t)\}$ is available and denote the log likelihood function by $l(\cdot)$. In order to test the above hypothesis using the score test procedure, it is necessary to obtain the vector of efficient scores, say $T^{-1/2}\mathbf{d}$, and Fisher's measure of information per observation. The elements of $\mathbf{d}$ are:

(2.1) $$\partial l/\partial\sigma_\epsilon^2 = (2\sigma_\epsilon^4)^{-1}\textstyle\sum_{t=1}^T \epsilon^2(t) - T(2\sigma_\epsilon^2)^{-1};$$

(2.2) $$\partial l/\partial\alpha_i = \sigma_\epsilon^{-2}\textstyle\sum_{t=1}^T \epsilon(t)\omega(t-i) \qquad i = 0, \cdots, q_\tau + s_\tau;$$

(2.3) $$\partial l/\partial\beta_i = \sigma_\epsilon^{-2}\textstyle\sum_{t=1}^T \epsilon(t)\xi(t-i) \qquad i = 1, \cdots, p_\tau + r_\tau;$$

(2.4) $$\partial l/\partial\phi_i = \sigma_\epsilon^{-2}\textstyle\sum_{t=1}^T \epsilon(t)u(t-i) \qquad i = 1, \cdots, p + r;$$

(2.5) $$\partial l/\partial\theta_i = -\sigma_\epsilon^{-2}\textstyle\sum_{t=1}^T \epsilon(t)v(t-i) \qquad i = 1, \cdots, q + s.$$

In the above expressions $\omega(t) = \phi(B)x(t)/\theta(B)\beta(B)$, $\xi(t) = \alpha(B)\omega(t)/\beta(B)$, $u(t) = \epsilon(t)/\phi(B)$ and $v(t) = \epsilon(t)/\theta(B)$. All required pre-sample values are set equal to their unconditional expectation of zero. The effect of this assumption is simply to introduce a transient error of $O(\kappa^T)$, $0 \leq \kappa < 1$, $T \to \infty$.

In what follows it will be useful to employ the following notation. For any two processes $\{\mu(t)\}$ and $\{\nu(t)\}$ $\gamma_{\mu\nu}(z)$ will denote the cross-covariance generating function. Then

(2.6) $$\gamma_{uu}(z) = \gamma_{\eta\eta}(z)/\theta(z)\theta(z^{-1}),$$

(2.7) $$\gamma_{uv}(z) = \gamma_{\eta\epsilon}(z)/\theta(z)\theta(z^{-1})$$

and

(2.8) $$\gamma_{vv}(z) = \sigma_\epsilon^2/\theta(z)\theta(z^{-1}).$$

Further, defining $\lambda(z) = \sigma_\epsilon^2/\beta(z)\beta(z^{-1})\gamma_{\eta\eta}(z)$ we have

$$\gamma_{\omega\omega}(z) = \lambda(z)\gamma_{xx}(z), \tag{2.9}$$

$$\gamma_{\omega\xi}(z) = \lambda(z)\gamma_{xx}(z)\alpha(z^{-1})/\beta(z^{-1}) \tag{2.10}$$

and

$$\gamma_{\xi\xi}(z) = \lambda(z)\gamma_{xx}(z)\alpha(z)\alpha(z^{-1})/\beta(z)\beta(z^{-1}). \tag{2.11}$$

Now, employing the linear transformation $\mathscr{T}$ defined in the Appendix, let

$$\sigma_\epsilon^{-2}\mathbf{\Sigma}_1 = \begin{bmatrix} \mathbf{I}_{\alpha\alpha} & \mathbf{I}_{\alpha\beta} \\ \mathbf{I}'_{\alpha\beta} & \mathbf{I}_{\beta\beta} \end{bmatrix}$$

where $\mathbf{I}_{\alpha\alpha} = \mathscr{T}_{(q_\tau+s_\tau+1)(q_\tau+s_\tau+1)}[\gamma_{\omega\omega}(z)/\sigma_\epsilon^2]$, $\mathbf{I}_{\alpha\beta} = \mathscr{T}_{(q_\tau+s_\tau+1)(p_\tau+r_\tau)}[\gamma_{\omega\xi}(z^{-1})z/\sigma_\epsilon^2]$ and $\mathbf{I}_{\beta\beta} = \mathscr{T}_{(p_\tau+r_\tau)(p_\tau+r_\tau)}[\gamma_{\xi\xi}(z)/\sigma_\epsilon^2]$. Similarly, let

$$\sigma_\epsilon^{-2}\mathbf{\Sigma}_2 = \begin{bmatrix} \mathbf{I}_{\phi\phi} & \mathbf{I}_{\phi\theta} \\ \mathbf{I}'_{\phi\theta} & \mathbf{I}_{\theta\theta} \end{bmatrix}$$

where $\mathbf{I}_{\phi\phi}$, $\mathbf{I}_{\phi\theta}$ and $\mathbf{I}_{\theta\theta}$ are defined in terms of $\mathscr{T}$ and (2.6)–(2.8) in an appropriate manner. After some manipulation the information matrix of the ARMAT$((q_\tau + s_\tau, p_\tau + r_\tau)(p + r, q + s))$ model can be shown to be

$$\mathbf{I}_{\delta\delta} = \sigma_\epsilon^{-2} \begin{bmatrix} (2\sigma_\epsilon^2)^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{\Sigma}_2 \end{bmatrix} \tag{2.12}$$

where $\boldsymbol{\delta} = (\sigma_\epsilon^2, \alpha_0, \cdots, \alpha_{q_\tau+s_\tau}, \beta_1, \cdots, \beta_{p_\tau+r_\tau}, \phi_1, \cdots, \phi_{p+r}, \theta_1, \cdots, \theta_{q+s})' = (\delta_1, \delta_2, \cdots, \delta_N)'$; compare Pierce (1972a).

The testing of $H((q_\tau, p_\tau)(p, q))$ against $H((q_\tau + s_\tau, p_\tau + r_\tau)(p + r, q + s))$ can be represented as a test of the restrictions that certain elements of $\boldsymbol{\delta}$ are equal to zero; i.e., $\delta_{j_0} = 0$, j = 1, $\cdots$, M, M = $s_\tau + r_\tau + r + s$. This leads to a consideration of

THEOREM 1. *If $\mathbf{I}_{\delta\delta}$ is positive definite and $\boldsymbol{\delta}$ satisfies the restrictions $\delta_{j_0} = 0$, $j = 1, \cdots, M$, then the statistic*

$$S_1 = T^{-1}\mathbf{d}^{*\prime}\mathbf{I}_{\delta\delta}^{*-1}\,\mathbf{d}^* \tag{2.13}$$

is asymptotically distributed as χ_M^2 (a chi-squared variate with M degrees of freedom).

The asterisk is used to denote evaluation at the point $\boldsymbol{\delta} = \boldsymbol{\delta}^*$, that value which maximises $l(\boldsymbol{\delta})$ subject to the constraints $\delta_{j_0} = 0$, j = 1, $\cdots$, M and the unrestricted elements of $\boldsymbol{\delta}^*$ are simply the usual (nonlinear) least-squares maximum likelihood estimates of the parameters of the ARMAT$((q_\tau, p_\tau)(p, q))$ model; see Pierce (1972a). A significantly large value of S_1 indicates that the restrictions imposed are not consistent with the sample data. It should be noted that the form of the test statistic of (2.13) and the structure of $\mathbf{I}_{\delta\delta}$ indicate that the transfer function and stochastic disturbance terms of the ARMAT model can be investigated separately. Hence the researcher may entertain the evaluation of a trio of score test statistics, one each for the transfer function and disturbance components individually and a third for joint misspecification of both parts of the model.

Autoregressive-moving average processes of order (p, q), ARMA(p, q), arise naturally as special cases of ARMAT models where $\{x(t)\} \equiv 0$. When the null hypothesis that $\{\eta(t)\}$ is ARMA(p, q) is to be tested against the general ARMA$(p + r, q + s)$ alternative, it is well known that an identifiability problem occurs and the matrix $\mathbf{\Sigma}_2$ will be singular. Using the notation established above, this can be clearly seen by recognizing that the equations

$$\gamma_{\eta\eta}(z)\phi(z^{-1}) = \gamma_{\eta\epsilon}(z)\theta(z^{-1})$$

$$\gamma_{\eta\epsilon}(z)\phi(z) = \sigma_\epsilon^2\theta(z),$$

together with (2.6)–(2.8) imply that $\boldsymbol{\Sigma}_2\mathbf{w}_2 = \mathbf{0}$ where

$$\mathbf{w}_2' = (1, -\boldsymbol{\phi}_p', \mathbf{0}_{r-1}', 1, -\boldsymbol{\theta}_q', \mathbf{0}_{s-1}'),$$

$\mathbf{0}_m$ being an m element column vector of zeros, $\boldsymbol{\phi}_p = (\phi_1, \phi_2, \cdots, \phi_p)'$ and $\boldsymbol{\theta}_q = (\theta_1, \theta_2, \cdots, \theta_q)'$; compare Hannan (1970, pages 413–414). Similar identifiability problems are also encountered with the transfer function component of the ARMAT model. Consider simultaneously increasing the order of both the numerator and denominator polynomials of the transfer function in the alternative hypothesis. Recall that (2.9)–(2.11) determine the elements of $\boldsymbol{\Sigma}_1$ and, setting $\{h(t)\} = \{y(t)\} - \{\eta(t)\}$ we obtain equations

$$\gamma_{xh}(z)\beta(z^{-1}) = \alpha(z^{-1})\gamma_{xx}(z) \tag{2.14}$$

and

$$\gamma_{hh}(z)\beta(z) = \alpha(z)\gamma_{xh}(z). \tag{2.15}$$

These expressions collectively imply that the vector

$$\mathbf{w}_1' = (0, \boldsymbol{\alpha}_{q_\tau}', \mathbf{0}_{s_\tau-1}, -1, \boldsymbol{\beta}_{p_\tau}', \mathbf{0}_{r_\tau-1}), \tag{2.16}$$

$\alpha_{q_\tau} = (\alpha_0, \alpha_1, \cdots, \alpha_{q_\tau})'$ and $\boldsymbol{\beta}_{p_\tau} = (\beta_1, \beta_2, \cdots, \beta_{p_\tau})'$, annihilates $\boldsymbol{\Sigma}_1$ when the null hypothesis is true.

The singularities present in the information matrix when considering alternative hypotheses in which both r and s are simultaneously non-zero and/or r_τ and $s_\tau > 0$ indicate that Theorem 1 is not applicable in these circumstances. However, Silvey (1959) has considered how to adapt the theory of the score test procedure in order to drop the assumption that $\mathbf{I}_{\delta\delta}$ be nonsingular and, in the case of ARMAT models, his modification leads to

THEOREM 2. *Assume that $\mathbf{I}_{\delta\delta}$ is singular of rank $N - g$ and that $\delta_{j_0} = 0, j = 1, \cdots, M$. If $\mathbf{G} = [\partial\delta_{j_0}/\partial\delta_i]$ $i = 1, \cdots, N$, $j = 1, \cdots, M$ and $\mathbf{G}_1$ is an appropriate $N \times g$ submatrix of $\mathbf{G}$ such that $\mathbf{I}_{\delta\delta} + \mathbf{G}_1\mathbf{G}_1'$ is positive definite, then the statistic*

$$S_2 = T^{-1}\mathbf{d}^{*\prime}(\mathbf{I}_{\delta\delta}^* + \mathbf{G}_1\mathbf{G}_1')^{-1}\,\mathbf{d}^* \tag{2.17}$$

is asymptotically distributed as χ^2_{M-g}.

The proofs of this and the preceding theorem follow from application of Billingsley's (1968, page 206) martingale central limit theorem to show the asymptotic normality of the score vector and Lemmas 5 and 7, respectively, of Silvey (1959).

In order to employ the statistic of Theorem 2 the rank of the information matrix under the null hypothesis must be known. Let ρ be the rank of $\boldsymbol{\Sigma}_1$. Eliminating the last r_τ rows and columns of $\boldsymbol{\Sigma}_1$ corresponding to the additional denominator polynomial coefficients of the transfer function in the alternative hypothesis, we obtain a submatrix of order $q_\tau + s_\tau + 1 + p_\tau$ of full rank. Deletion of the $q_\tau + 2$ to $(q_\tau + s_\tau + 1)$th rows and columns of $\boldsymbol{\Sigma}_1$ results in a nonsingular square matrix of order $q_\tau + 1 + p_\tau + r_\tau$. Hence $\rho \geq \max(q_\tau + s_\tau + 1 + p_\tau, q_\tau + 1 + p_\tau + r_\tau) = (q_\tau + s_\tau + 1 + p_\tau + r_\tau) - \min(s_\tau, r_\tau)$. Further, let $\mathbf{R} = \operatorname{diag}(\mathbf{C}_{q_\tau+s_\tau+1}, \mathbf{C}_{p_\tau+r_\tau})$ where $\mathbf{C}_m$ denotes the $m \times m$ circulant matrix with initial row $(0, 0, \cdots, 0, 1)$. Using expressions (2.14)–(2.16) and (2.9)–(2.11), it follows that each of the h linearly independent vectors

$$\mathbf{a}_j = \mathbf{R}^j\mathbf{w}_1 \qquad j = 0, \cdots, h - 1, h = \min(s_\tau, r_\tau),$$

provides an eigenvector of $\boldsymbol{\Sigma}_1$ corresponding to a zero eigenvalue. Therefore, $\rho = (q_\tau + s_\tau + 1 + p_\tau + r_\tau) - \min(s_\tau, r_\tau)$. An analogous argument can be employed to show that the rank of $\boldsymbol{\Sigma}_2$ is $(p + r + q + s) - \min(r, s)$, so that $g = \min(s_\tau, r_\tau) + \min(r, s)$.

The form of the restrictions employed here makes it clear that $\mathbf{G}_1$ will be null apart from g appropriately positioned values of unity. Although the selection of $\mathbf{G}_1$ must guarantee that $\mathbf{I}_{\delta\delta} + \mathbf{G}_1\mathbf{G}_1'$ is of full rank, various legitimate choices are available and this raises the question of the effect that any particular $\mathbf{G}_1$ has upon the test statistic. From Rao and Mitra (1971, Complement 5(b), page 40) $(\mathbf{I}_{\delta\delta} + \mathbf{G}_1\mathbf{G}_1')^{-1}$ is a generalized inverse of $\mathbf{I}_{\delta\delta}$, say $\mathbf{I}_{\delta\delta}^-$. Since $\mathbf{d}$ (asymptotically) belongs to $\mathscr{M}(\mathbf{I}_{\delta\delta})$, the vector space generated by the columns of $\mathbf{I}_{\delta\delta}$, it follows from Rao and Mitra (1971, Lemma 2.2.4(ii), page 21) that the statistic of Theorem 2 is, in fact, asymptotically invariant with respect to $\mathbf{G}_1$.

Aitchison and Silvey (1960) point out that the selection of a $\mathbf{G}_1$ effectively requires that g elements of $\boldsymbol{\delta}$ are zero under both null and alternative hypotheses in order to obtain the identifiability of the remaining $N - g$ parameters. For example, one possible choice of $\mathbf{G}_1$ when $s_\tau = r_\tau$ and $r = s$ corresponds to implicitly testing $H((q_\tau, p_\tau)(p, q))$ against the "restricted" alternative $H((q_\tau + s_\tau, p_\tau)(p, q + s))$ while another amounts to testing this null hypothesis against the "restricted" alternative $H((q_\tau, p_\tau + r_\tau)(p + r, q))$. Both of these tests could, of course, be conducted explicitly using (2.13). In addition, it may be possible to represent the null hypothesis using parametric restrictions other than simple exclusions, but a simple corollary of the above invariance is that, provided $\mathbf{I}_{\delta\delta}$ can be augmented by a $N \times g$ submatrix of the gradient matrix of those restrictions to obtain nonsingularity, the test statistic (2.17) remains unaffected asymptotically. This suggests that, when considering any particular alternative hypothesis, the researcher is actually implicitly testing the null hypothesis against a wider range of alternatives than is immediately apparent.

We also have the interesting relationship between the test statistics of Theorems 1 and 2 enunciated in

THEOREM 3. *The test statistic of Theorem* 1, *when used to explicitly test a "restricted"* ARMAT *alternative implicit in the choice of a* $\mathbf{G}_1$ *is asymptotically identical to the test statistic of Theorem* 2.

A theorem of this kind is proved in Poskitt and Tremayne (1980) in the context of autoregressive-moving average processes and its extension to the models of the present paper is comparatively straightforward.

An implication of the results obtained above is that, when applied to linear time series models, the score test procedure assumes the role of a pure significance test, although the transfer function and stochastic disturbance components of the model may be considered separately. The test procedure advocated is to be interpreted as a general specification error test of the adequacy of a fitted model. In common with portmanteau tests, the score test does not provide a criterion for identifying the degrees of the polynomials in (1.1) and should be regarded simply as a device for diagnostic checking.

3. Score tests and residual correlations. The diagnostic tests previously proposed for ARMAT models have been of the portmanteau type and are structured in terms of estimated residual auto and cross-correlations. In view of the interpretation of the procedures of this paper as pure significance tests, it seems useful to investigate the relationship between score and portmanteau tests of linear time series models.

Consider the situation in which an ARMAT$((q_\tau, p_\tau)(p, q))$ model has been fitted and the adequacy of the specification of the stochastic disturbance is to be assessed. Pierce (1972b) has suggested that the Box and Pierce (1970) statistic $Q_D = T\mathbf{r}_\epsilon^{*\prime}\mathbf{r}_\epsilon^*$, $\mathbf{r}_\epsilon = (r_\epsilon(1), \cdots, r_\epsilon(K))'$ a K element vector of residual autocorrelations, may be employed in these circumstances. Q_D is approximately distributed χ^2_{K-p-q} for large K if the model is correct. If the disturbance component of the model has been incorrectly specified, then one would expect significant values of Q_D to be observed. However, Box and Jenkins (1976, Section 11.3.3) point out that it is possible that a large value of Q_D may arise because of a misspecification of the transfer function component. Pierce (1972b) developed a portmanteau statistic which is specifically designed to investigate the adequacy of this part of the model. Given the assumptions of Section 1, we may write

$$x(t) = \textstyle\sum_{j=0}^{\infty} \pi_j \zeta(t - j) \tag{3.1}$$

where $|\sum \pi_j| < \infty$ and $\{\zeta(t)\}$ is a non-autocorrelated process with variance σ_ζ^2 so that $\gamma_{xx}(z) = \pi(z)\pi(z^{-1})\sigma_\zeta^2$. Let $\mathbf{r}_{\epsilon\zeta} = (r_{\epsilon\zeta}(0), \cdots, r_{\epsilon\zeta}(K))'$ be a $K+1$ component vector of cross-correlations between $\{\epsilon(t)\}$ and $\{\zeta(t)\}$. Pierce shows that $Q_T = T\mathbf{r}_{\epsilon\zeta}^{*\prime}\mathbf{r}_{\epsilon\zeta}^*$, which is approximately distributed $\chi^2_{K-q_\tau-p_\tau}$, can provide a useful check of the specification of the transfer function.

Now assume that it is desired to test $H((q_\tau, p_\tau)(p, q))$ against $H((q_\tau + s_\tau, p_\tau)(p, q))$, $s_\tau = K + 1$ using the method of scores. The score test for testing the null hypothesis against the proposed alternative is structured in terms of the subvector $T^{-1/2}\,\mathbf{d}_\alpha = [T^{-1/2}\partial l/\partial\alpha_i]$, $i = q_\tau + 1, \cdots, q_\tau + s_\tau$ of the score vector $T^{-1/2}\,\mathbf{d}$. The asymptotic distribution of the subvector $T^{-1/2}\,\mathbf{d}_\alpha$ is Gaussian with zero mean vector and covariance matrix $\mathbf{C} - \mathbf{B}'\mathbf{A}^{-1}\mathbf{B}$, $N(\mathbf{0}, \mathbf{C} - \mathbf{B}'\mathbf{A}^{-1}\mathbf{B})$, where:

$$\mathbf{A} = \left[\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{12}' & \mathbf{A}_{22} \end{array}\right]$$

$$\mathbf{A}_{11} = \mathscr{T}_{p_\tau p_\tau}\left[\frac{\psi(z)\psi(z^{-1})\alpha(z)\alpha(z^{-1})\sigma_\zeta^2}{\beta^2(z)\beta^2(z^{-1})\sigma_\epsilon^2}\right], \quad \mathbf{A}_{12} = \mathscr{T}_{p_\tau(q_\tau+1)}\left[\frac{\psi(z)\psi(z^{-1})\alpha(z^{-1})\sigma_\zeta^2}{\beta(z)\beta^2(z^{-1})z\sigma_\epsilon^2}\right]$$

$$\mathbf{A}_{22} = \mathscr{T}_{(q_\tau+1)(q_\tau+1)}\left[\frac{\psi(z)\psi(z^{-1})\sigma_\zeta^2}{\beta(z)\beta(z^{-1})\sigma_\epsilon^2}\right];$$

$$\mathbf{B}' = [\mathbf{B}_1' : \mathbf{B}_2'],$$

$$\mathbf{B}_1 = \mathscr{T}_{p_\tau s_\tau}\left[\frac{z^{q_\tau}\psi(z)\psi(z^{-1})\alpha(z^{-1})\sigma_\zeta^2}{\beta(z)\beta^2(z^{-1})\sigma_\epsilon^2}\right], \quad \mathbf{B}_2 = \mathscr{T}_{(q_\tau+1)s_\tau}\left[\frac{z^{q_\tau+1}\psi(z)\psi(z^{-1})\sigma_\zeta^2}{\beta(z)\beta(z^{-1})\sigma_\epsilon^2}\right];$$

$$\mathbf{C} = \mathscr{T}_{s_\tau s_\tau}\left[\frac{\psi(z)\psi(z^{-1})\sigma_\zeta^2}{\beta(z)\beta(z^{-1})\sigma_\epsilon^2}\right]$$

and $\psi(z) = \pi(z)\phi(z)/\theta(z)$. Finally, set

$$\mathbf{D} = \mathscr{T}_{s_\tau s_\tau}\left[\frac{\psi(z^{-1})\sigma_\zeta}{z^{q_\tau+1}\beta(z^{-1})\sigma_\epsilon}\right] \quad \text{and} \quad \mathbf{E}' = [\mathbf{E}_1' : \mathbf{E}_2'],$$

$$\mathbf{E}_1 = \mathscr{T}_{p_\tau s_\tau}\left[\frac{\psi(z^{-1})\alpha(z^{-1})\sigma_\zeta}{z\beta^2(z^{-1})\sigma_\epsilon}\right] \quad \text{and} \quad \mathbf{E}_2 = \mathscr{T}_{(q_\tau+1)s_\tau}\left[\frac{\psi(z^{-1})\sigma_\zeta}{\beta(z^{-1})\sigma_\epsilon}\right].$$

Using the lemmas of the Appendix we have, for s_τ sufficiently large $T^{-1/2}\,\mathbf{d}_\alpha = T^{1/2}\mathbf{D}\mathbf{r}_{\epsilon\zeta}$, $\mathbf{C} \doteqdot \mathbf{D}\mathbf{D}'$, $\mathbf{E}' \doteqdot \mathbf{D}^{-1}\mathbf{B}'$ and $\mathbf{A} \doteqdot \mathbf{E}\mathbf{E}'$. Therefore, if $s_\tau \to \infty$ such that $s_\tau/T \to 0$, $T^{1/2}\mathbf{r}_{\epsilon\zeta}$ converges in distribution to $N(\mathbf{0}, \mathbf{1}_{s_\tau} - \mathbf{E}'(\mathbf{E}\mathbf{E}')^{-1}\mathbf{E})$, $\mathbf{1}_m$ being an identity matrix of order m. This proves

PROPOSITION 1. *If $s_\tau \to \infty$ and $T \to \infty$ such that $s_\tau/T \to 0$ then the score test statistic and the portmanteau statistic Q_T are equivalent in the sense that the score vector approaches a nonsingular linear transformation of the residual cross-autocorrelations.*

REMARK 1. In view of the idempotency present as $s_\tau \to \infty$, the statistic of Theorem 1 is not in fact available, although one could compute $T^{-1}\,\mathbf{d}_\alpha^{*\prime}(\mathbf{D}^*\mathbf{D}^{*\prime})^{-1}\,\mathbf{d}_\alpha^* \doteqdot Q_T$.

This proposition and its proof provide a parallel to the arguments of Pierce (1972b, Sections 2.3 and 4). The relationship between Q_D and the test statistics of this paper ensues from considering the score test of $H((q_\tau, p_\tau)(p, q))$ against $H((q_\tau, p_\tau)(p, q + s))$, $s = K$ and is given by

PROPOSITION 2. *As $s \to \infty$ with T such that $s/T \to 0$ there is an equivalence between the score test statistic and the portmanteau statistic Q_D.*

That this proposition holds can be demonstrated in a similar manner to Proposition 1. This type of relationship has been noted previously in the context of dynamic regression models with no exogenous regressors by Breusch (1978).

Propositions 1 and 2, together with the invariance and asymptotic equivalence properties established in the previous section indicate that, in effect, both Q_D and Q_T do test against a wide range of alternative misspecifications. A contrast between portmanteau and score tests, however, is that the approximations of Box and Pierce (1970) and Pierce (1972b) rely upon $K \to \infty$, though at a slower rate than T, whereas the distributional results of this paper do not require the degrees of freedom to be large.

REMARK 2. It should be pointed out that if it cannot be assumed that the stochastic processes of interest are characterised by a finite number of parameters, so that there are no "true" degrees, then the consistency of the test procedure would require the degrees of freedom to increase without bound as $T \to \infty$.

4. Practical considerations. In Sections 2 and 3 some formal results regarding the large sample behaviour of score tests when applied to ARMAT models were obtained. The concluding part of this paper investigates and discusses the implementation of score and portmanteau tests in practice.

Some authors, see for example Breusch (1978) and Godfrey (1979), have presented score tests in terms of T times a squared multiple correlation coefficient obtained from an auxiliary regression. This convenient computational device can be exploited when testing hypotheses concerning ARMAT models using the statistics of Theorems 1 and 2. To demonstrate this let $\boldsymbol{\epsilon} = (\epsilon(1), \cdots, \epsilon(T))'$ and $\mathbf{X}' = [\partial\epsilon(t)/\partial\delta_i]$, $i = 2, \cdots, N$, $t = 1, \cdots, T$. Reorder and partition $\mathbf{X}$ into $[\mathbf{X}_1 : \mathbf{X}_2]$ where $\mathbf{X}_1$ is $T \times N - g - 1$ and $\mathbf{X}_2$ is $T \times g$, the subscript 2 referring to those columns of $\mathbf{X}$ corresponding to coefficients restricted under the alternative hypothesis so as to obtain identifiability. The information matrix of the $N - g - 1$ parameters whose derivative processes appear in $\mathbf{X}_1$ is now nonsingular and the statistic of Theorem 1 when used to explicitly test the "restricted" alternative can thus be structured as

$$S_1 = T\boldsymbol{\epsilon}^{*\prime}\mathbf{X}_1^*(\mathbf{X}_1^{*\prime}\mathbf{X}_1^*)^{-1}\mathbf{X}_1^{*\prime}\boldsymbol{\epsilon}^*/\boldsymbol{\epsilon}^{*\prime}\boldsymbol{\epsilon}^*. \tag{4.1}$$

The statistic of (2.17) may be expressed in a similar way by augmenting $\boldsymbol{\epsilon}$ and $\mathbf{X}$. Setting $\boldsymbol{\epsilon}_a' = [\boldsymbol{\epsilon}' : \mathbf{0}_g]$ and $\mathbf{X}_a' = [\mathbf{X}' : \mathbf{G}_a]$, where $\mathbf{G}_a' = [\mathbf{0} : \mathbf{1}_g]$ is a $g \times N - 1$ matrix, we have

$$S_2 = T\boldsymbol{\epsilon}_a^{*\prime}\mathbf{X}_a^*(\mathbf{X}_a^{*\prime}\mathbf{X}_a^*)^{-1}\mathbf{X}_a^{*\prime}\boldsymbol{\epsilon}_a^*/\boldsymbol{\epsilon}_a^{*\prime}\boldsymbol{\epsilon}_a^*. \tag{4.2}$$

Both (4.1) and (4.2) are, apart from an asymptotically negligible nonzero sample mean correction, T times the coefficient of determination resulting from the regression of $\boldsymbol{\epsilon}^*$ on $\mathbf{X}_1^*$ or $\boldsymbol{\epsilon}_a^*$ on $\mathbf{X}_a^*$. However, when formulated in this way, the sample quantities constituting the matrix and vector elements of S_1 and S_2 satisfy the algebraic relationships (2.6)–(2.11). Arguments analogous to those previously employed can, therefore, now be used to show that the finite sample behaviour of the score test parallels the asymptotic invariance and equivalence properties of Section 2, namely S_2 is invariant with respect to the choice of generalized inverse implicit in its construction and, moreover, is numerically identical to S_1. An important corollary of our results is that, in empirical situations, the researcher need only ever employ the statistic S_1, the precise formulation of which depending on $m_\tau = \max(s_\tau, r_\tau)$ and $m = \max(r, s)$ in order to test a fitted ARMAT model against quite a wide range of misspecification.

The tests proposed in this paper are large sample ones and their practical usefulness depends upon their performance with small and moderate sample sizes. Godfrey (1979) has, in the context of autoregressive-moving average models, provided simulation evidence that the score test performs quite well with regard to both size and power characteristics in finite samples. In view of the more elaborate nature of the transfer function models considered in this paper, further investigation of the empirical performance of diagnostic tests on both components of the model seems in order.

To provide a comparative study of the performance of the statistics of this paper, a number of variants of portmanteau tests are included in the simulation experiments. Two versions of these tests are used to assess the adequacy of both the disturbance and transfer function specifications. The statistic Q_D and the modification to this statistic suggested by Ljung and Box (1978), hereafter QL_D, are used to test the lack of fit of the disturbance part of the model. The statistics Q_T and that with the adjustment also advocated by Ljung and Box are employed to test the transfer function. The notation QL_T is adopted for the latter statistic. In order to employ the variants of the portmanteau statistics under consideration, it is necessary to decide upon the number of terms to be used in their computation. In many applications of portmanteau tests the value of K has been as high as 20 or 30, even when the sample size used is by no means large, possibly less than 100. This usage appears contrary to the theoretical derivations which require not only that $K \to \infty$ with T, but also that $K/T \to 0$. With this point in mind, three values of K are chosen, $K = 20$, $K = [\sqrt{T}]$ and $K = \left[\left(1 + \frac{n}{2}\right)\ln(T)\right]$, n being the number of coefficients fitted to the part of the model under test. Similarly, in the case of score tests, values of m_τ and m must be selected but, by contrast with portmanteau tests, they do not require the degrees of freedom to be large. Furthermore, if the identification stage of the model building process has been carefully carried out, it seems unlikely that the applied worker will, on philosophical grounds, usually wish to consider m_τ or m at all large. Some further discussion in the light of experimental results is given below but these quantities are generally chosen as 1 or 2 for the purposes of estimating the size and power of the score tests.

The size computations are all based on the model

$$y(t) = \frac{0.5}{(1 - 0.5B)} x(t) + \epsilon(t) \qquad t = 0, \pm 1, \pm 2, \cdots, \tag{4.3}$$

which is the model 10*b* of Box and Jenkins (1976, Section 10.2.3) with $b = 0$. The input $\{x(t)\}$ is chosen as i.i.d. $N(0, \sigma_x^2)$, $T = 100$ and 6000 replications are used to ensure an upper bound of 0.0065 on the standard error of the observed proportion of rejections. It seems useful (in view, for example, of the common practice of quoting "prob-values" in hypothesis testing) to check that the empirical size of a test statistic follows its theoretical null distribution quite closely for a wide range of significance levels and not just at conventional values such as 0.05 and 0.1. Accordingly, a grid of 13 such values, 0.9 (0.1)0.1, 0.05, 0.033, 0.02, 0.01 is used. Table 1 presents the results for this model. This table contains no entries for $m_\tau = m = 2$ as the results for these cases are virtually indistinguishable from those obtained with $m_\tau = m = 1$. Furthermore, portmanteau statistics without the adjustments of Ljung and Box (1978) are also absent since there is no case in our experiments in which their size characteristics match up to those of QL_T and QL_D. This finding is consistent with that obtained by Davies, Triggs and Newbold (1977), for example, in the context of autoregressive-moving average models. Naturally, their power characteristics are also invariably inferior to those of the adjusted portmanteau statistics, compare Davies and Newbold (1979), and so there seems no reason to present results for these tests. On perusal of the body of the table, the score tests and portmanteau tests employing the square root and logarithmic rules for choice of K appear to perform quite satisfactorily. The final row presents values of chi-squared goodness-of-fit statistics for each test. These indicate in the case of portmanteau tests that $K = 20$ is not an appropriate choice and that the logarithmic rule is to be preferred. The score tests invariably fit the null distribution well.

Turning to an investigation of the empirical power of the tests, a situation in which $\beta(z)$ is misspecified is first considered. The fitted model is an ARMAT((0, 1)(1, 0)) and the alternatives generated involve $\phi(z) = 1 - 0.75z$, $\alpha_0 = 0.5$ and $\beta(z) = (1 - 0.5z)(1 - bz)$

TABLE 1
Empirical size characteristics of score and portmanteau tests for model (4.3) *with* $T = 100$

SIGNIFICANCE LEVEL	MODEL COMPONENT								
	Disturbance				Transfer Function				Both
	QL_D			S	QL_T			S	S
	K=20	K=10	K=4	m=1	K=20	K=10	K=6	m_τ=1	m_τ+m=2
.9	.821	.885	.893	.900	.849	.904	.904	.907	.900
.8	.695	.773	.795	.801	.719	.798	.803	.801	.803
.7	.578	.661	.689	.697	.600	.692	.701	.698	.708
.6	.481	.565	.589	.599	.494	.595	.595	.598	.607
.5	.380	.465	.480	.507	.392	.493	.500	.501	.507
.4	.298	.366	.380	.402	.303	.386	.394	.408	.406
.3	.224	.273	.277	.301	.219	.284	.291	.308	.306
.2	.151	.179	.181	.202	.135	.182	.189	.202	.205
.1	.080	.086	.085	.100	.064	.089	.093	.104	.100
.05	.047	.048	.047	.046	.028	.047	.043	.051	.048
.033	.033	.033	.029	.029	.017	.028	.029	.035	.027
.02	.021	.022	.018	.017	.009	.018	.017	.023	.018
.01	.013	.011	.010	.008	.004	.009	.009	.011	.009
χ^2	633.28	61.58	34.54	12.34	406.92	26.87	15.46	18.81	18.43

where b takes on 19 values $-0.9(0.1)0.9$. A summary of the simulation results, based upon 1000 replications, is provided by the power curves for the conventional 0.05 significance level given in Figure 1. The graph with the dot symbols relates to the portmanteau statistic utilising the logarithmic rule which, in all simulations conducted, provides the most powerful variant of these tests. The curve identified by asterisks is that of the score test with $m_\tau = 1$, the true value, and it is evident that this test is the more powerful for all nonzero values of b. When m_τ is incorrectly chosen as 2, the empirical power of the score test, which is not shown on the graph, is somewhat less than when $m_\tau = 1$ but is still noticeably higher than that of QL_T.

A significant feature of both power curves in Figure 1 is their declining ability to detect the misspecification of the fitted model as the absolute value of b is increased beyond about 0.7. This implies that the transfer function of an ARMAT ((0, 1)(1, 0)) model can better approximate that of an ARMAT ((0, 2)(1, 0)) for some combinations of the roots of $\beta(z)$ than for others. Examination of the Fourier coefficients of the transfer function of the alternative model reveals that this state of affairs will obtain whenever one root of $1 - \beta_1 z - \beta_2 z^2$ is of moderate size and the other is close to the unit circle, or trivially, near zero. Parallel results for the disturbance component of the model are implicit in Box and Jenkins (1976, Sections 3.2.3–3.2.4). Further Monte Carlo experiments show that diagnostic checks for this part of the model do exhibit similar empirical characteristics to those described for the transfer function.

In view of the theoretical content of this paper, it seems natural to consider a case in which the information matrix of the alternative model is singular when the null hypothesis is true. Experimental results are, therefore, presented using $y(t) = \alpha_0 x(t) + \epsilon(t)$ for the fitted model with the observations being generated by

$$y(t) = \alpha_0 x(t) + \frac{(1 - \theta_1 B)}{(1 - \phi_1 B)} \epsilon(t), \tag{4.4}$$

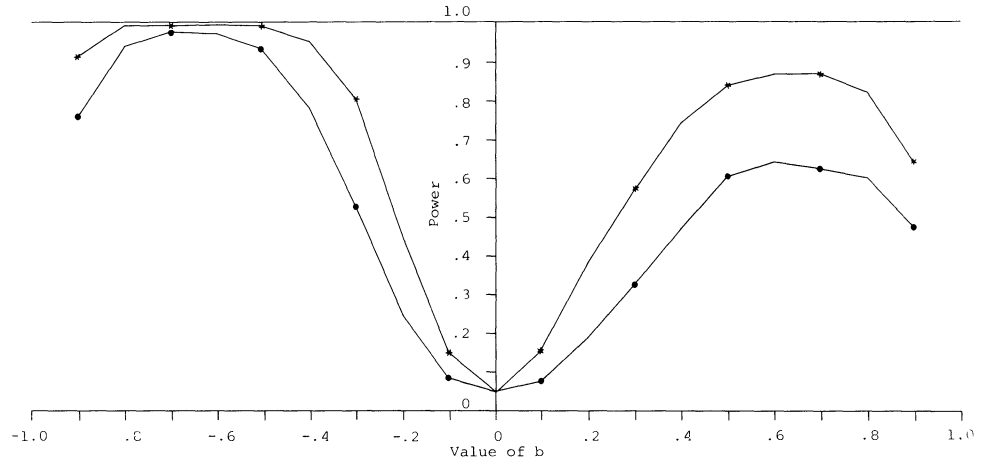

FIG. 1. *Empirical power curves of score and portmanteau tests when $\beta(z)$ is underspecified*: $T = 100$

$\alpha_0 = 0.5$, for various values of ϕ_1 and θ_1. In this case the equivalence established in Proposition 2 holds exactly in finite samples, so that, apart from the adjustment of Ljung and Box (1978), the score and portmanteau test statistics would be identical if m were equal to K. The principal difference between the two approaches to diagnostic testing in this instance lies, therefore, in the determination of the degrees of freedom. For the particular case of interest here $m = 1$ and the score test statistic equals $Tr_\epsilon^{*2}(1)$. It is worthy of note that Hannan (1980) discusses the likelihood ratio test for this situation and shows that there are circumstances when this may also be near to $Tr_\epsilon^{*2}(1)$. As might be expected, all test procedures considered here are substantially more powerful when the two parameters ϕ_1 and θ_1 are highly distinct. However, no matter how these values are selected, assuming no parameter redundancy, the empirical power of the portmanteau test is invariably inferior to that of the appropriate score test at all conventional significance levels. A typical example of the performance of the tests is given in Table 2 where $\phi_1 = 0.3$

TABLE 2
Observed power of diagnostic tests when both ϕ (z) and θ (z) are underspecified: T = 100

Significance Level	QL_D K=4	S m=1
.9	.979	.988
.8	.958	.967
.7	.922	.945
.6	.877	.918
.5	.819	.895
.4	.761	.856
.3	.671	.812
.2	.573	.737
.1	.437	.612
.05	.321	.482
.033	.268	.413
.02	.219	.340
.01	.155	.259

TABLE 3
Observed power of diagnostic tests when $\beta(z)$ and $\theta(z)$ are underspecified; T = 50

SIGNIFICANCE LEVEL	MODEL COMPONENT							
	Disturbance			Transfer Function			Both	
	QL_D	S		QL_T	S		S	
	K=3	m=1	m=2	K=7	m_τ=1	m_τ=3	m_τ+m=2	m_τ+m=5
.9	.975	.980	.975	.980	.978	.977	.992	.993
.8	.941	.954	.959	.944	.952	.941	.983	.983
.7	.902	.932	.922	.901	.917	.903	.973	.966
.6	.858	.893	.876	.847	.878	.845	.953	.940
.5	.806	.859	.824	.778	.835	.789	.935	.914
.4	.736	.805	.761	.714	.786	.720	.900	.870
.3	.644	.742	.693	.614	.720	.633	.855	.803
.2	.515	.648	.597	.491	.634	.502	.798	.706
.1	.366	.516	.436	.328	.482	.342	.689	.566
.05	.249	.387	.309	.222	.361	.219	.541	.423
.033	.199	.318	.244	.171	.292	.170	.467	.335
.02	.151	.260	.188	.128	.224	.129	.398	.245
.01	.107	.176	.115	.092	.156	.079	.298	.151

and $\theta_1 = 0.1$. This provides an alternative interpretation of the power of a test as the cumulative distribution of "prob-values", with one such distribution for each null and alternative. See Cox and Hinkley (1974, Chapters 3 and 4).

There has been no discussion of the effects of variations in certain of the experimental design parameters; for example, generalising the stochastic structure of the input process or altering the ratio of the variances of the signal to noise components of the model. In fact, such variations are incorporated in the collection of simulation experiments conducted, but they do not change any of the basic conclusions reached in a material way. All results so far quoted use $T = 100$ but the sample size can be reduced quite considerably without appreciable loss of empirical power or unusual size characteristics arising.

In practice, an a priori choice must be made for m_τ and m in the absence of knowledge about their true values. As the degree of misspecification of the polynomials in (1.1) has hitherto always been one, it is probably not surprising that the score tests prove effective for choice of $m_\tau = m = 1$, the correct values. In order to investigate the situation when this is not so, an alternative model is considered in which $m_\tau = r_\tau = 3$, $m = s = 2$ and $s_\tau = r = 0$ in truth. Table 3 relates to this case with $\beta(z) = 1 - 0.73z + 0.168z^2 - 0.0067z^3 + 0.00038z^4$ and $\theta(z) = 1 - 0.3z + 0.02z^2$ and where the fitted model is ARMAT ((0, 1)(0, 0)). The sample size used is 50 and $\{x(t)\}$ is generated as a first order autoregression with parameter 0.3. From this table, it can be seen that there appear to be no gains obtainable from setting m_τ and m equal to their true values, as opposed to following the simple rule of fixing them at unity, thus emphasizing the role of the score test procedure as a pure significance test. One caveat to the application of this rule occurs when some or other of the intermediate coefficients of the polynomials are zero, as may happen, for example, in seasonal models. It is straightforward to tailor the theoretical results of this paper so as to obtain score test appropriate for such models and also, of course, for models exhibiting dead time or delay. An additional observation to be made from the table is that the joint score test generally has higher power than any of the tests on individual components. Further investigation of the replications in a case where both parts of the model are misspecified shows that there are realisations for which the joint test correctly rejects the fitted null model at conventional significance levels when no other test does.

Little reference is made above to the existence, or otherwise, of contamination of diagnostic tests for one part of the model when that component is actually correctly specified, though the other is not. This is because it does not occur in the situations discussed, nor, in fact in most of the other cases for which we have simulation results. However, there are conditions when contamination does arise. (For example, when an ARMAT ((0, 1)(0, 0)) model is fitted to data generated from an ARMAT ((0, 2)(0, 0)) model with the roots of $\beta(z)$ close to the unit circle.) In such circumstances, the joint score test always seems to have substantially more empirical power than any of the individual test statistics. To sum up, therefore, we would recommend that all three score test statistics should be computed when applying diagnostic checks to open-loop transfer function models in practice.

APPENDIX

Let $\mathscr{R}[z]$ denote the field of rational power series with real coefficients

$$\mathscr{R}[z] = \{p(z) = \textstyle\sum_{i=-\infty}^{\infty} p_i z^i \mid p_i \in \mathscr{R}, \mid \sum_{i=-\infty}^{\infty} p_i \mid < \infty\}$$

and $\mathscr{R}^+[z] = \{p(z) \in \mathscr{R}[z] | p_i = 0 \text{ for } i < 0\}$; see Birkhoff and MacLane (1977). Consider the mapping $\mathscr{T}: \mathscr{R}[z] \to \mathscr{B}$ where, for each $p(z) \in \mathscr{R}[z]$ and each $m, n \in \mathscr{N}$, $\mathscr{T}_{mn}[p(z)] = \mathbf{P}$, a $m \times n$ matrix whose ijth element is given by the coefficient of z^{i-j} in $p(z)$. $\mathscr{T}$ defines a linear transformation from $\mathscr{R}[z]$ to the class of band matrices $\mathscr{B}$.

For the purposes of this paper the following results are useful.

LEMMA A1. *For all* $p(z) \in \mathscr{R}[z]$, *if* $\mathscr{T}_{mn}[p(z)] = \mathbf{P}$ *then* $\mathscr{T}_{nm}[p(z^{-1})] = \mathbf{P}'$.

LEMMA A2. *Let* $p(z), q(z) \in \mathscr{R}^+[z]$. *If* $\mathscr{T}_{mm}[p(z)] = \mathbf{P}$ *and* $\mathscr{T}_{mr}[q(z)] = \mathbf{Q}$ *then* $\mathscr{T}_{mr}[p(z)q(z)] = \mathbf{PQ}$.

LEMMA A3. *If* $p(z) \in \mathscr{R}^+[z]$ *is such that* $1/p(z) \in \mathscr{R}^+[z]$ *then, for any* m, $\mathscr{T}_{mm}[p(z)] = \mathbf{P}$ *implies* $\mathscr{T}_{mm}[1/p(z)] = \mathbf{P}^{-1}$.

LEMMA A4. *If* $p(z), q(z) \in \mathscr{R}^+[z]$ *and* $\mathscr{T}_{m\infty}[p(z)] = \lim_{r\to\infty} \mathscr{T}_{mr}[p(z)]$, $\mathscr{T}_{\infty n}[q(z)] = \lim_{r\to\infty} \mathscr{T}_{rn}[q(z)]$, *then*: $\mathscr{T}_{nm}[p(z^{-1})q(z)] = \mathscr{T}_{m\infty}[p(z^{-1})]\mathscr{T}_{\infty n}[q(z)] = \lim_{r\to\infty} \mathscr{T}_{mr}[p(z^{-1})] \cdot \mathscr{T}_{rn}[q(z)]$.

Simple algebraic proofs of these lemmas, which are implicit in the results of Grenander and Szegö (1958), are available from the authors on request.

Acknowledgments. We are grateful to L. G. Godfrey for helpful discussions and suggestions. The referees are also to be thanked for useful comments and guidance on the presentation of the paper.

REFERENCES

AITCHISON, J. and SILVEY, S. D. (1960). Maximum-likelihood estimation procedures and associated tests of significance. *J. Roy. Statist. Soc. Ser. B* **22** 154–171.

BILLINGSLEY, P. (1968). *Convergence of Probability Measure.* Wiley, New York.

BIRKHOFF, G. and MACLANE, S. (1977). *A Survey of Modern Algebra*, 4th ed. Macmillan, New York.

BOX, G. E. P. and JENKINS, G. M. (1976). *Time Series Analysis: Forecasting and Control*, 2nd ed. Holden-Day, San Francisco.

BOX, G. E. P. and PIERCE, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *J. Amer. Statist. Assoc.* **65** 1509–26.

BREUSCH, T. S. (1978). Testing for autocorrelation in dynamic linear models. *Australian Economic Papers* **17** 334–355.

COX, D. R. and HINKLEY, D. V. (1974). *Theoretical Statistics.* Chapman and Hall, London.

DAVIES, N. and NEWBOLD, P. (1979). Some power studies of a portmanteau test of time series model specification. *Biometrika* **66** 153–155.

DAVIES, N., TRIGGS, C. M. and NEWBOLD, P. (1977). Significance levels of the Box-Pierce portmanteau statistic in finite samples. *Biometrika* **64** 517–522.

GODFREY, L. G. (1979). Testing the adequacy of a time series model. *Biometrika* **66** 67–72.

GRENANDER, U. and SZEGÖ, G. (1958). *Toeplitz Forms and Their Applications.* Univ. California Press, Berkeley.

HANNAN, E. J. (1970). *Multiple Time Series.* Wiley, New York.

HANNAN, E. J. (1980). Testing for autocorrelation and Akaike's criterion. Unpublished paper, Australian National University.

HANNAN, E. J. and HEYDE, C. C. (1972). On limit theorems for quadratic functions of discrete time series. *Ann. Math. Statist.* **43** 2058–2066.

HAUGH, L. D. and BOX, G. E. P. (1977). Identification of dynamic regression (distributed lag) models connecting two time series. *J. Amer. Statist. Assoc.* **72** 121–130.

LJUNG, G. M. and BOX, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika* **65** 297–303.

PIERCE, D. A. (1972a). Least squares estimation in dynamic-disturbance time series models. *Biometrika* **59** 73–78.

PIERCE, D. A. (1972b). Residual correlations and diagnostic checking in dynamic-disturbance time series models. *J. Amer. Statist. Assoc.* **67** 636–640.

POSKITT, D. S. and TREMAYNE, A. R. (1980). Testing the specification of a fitted autoregressive-moving average model. *Biometrika* **67** 359–363.

RAO, C. R. (1948). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Proc. Camb. Phil. Soc.* **44** 50–57.

RAO, C. R. and MITRA, S. K. (1971). *Generalized Inverse of Matrices and its Applications.* Wiley, New York.

SILVEY, S. D. (1959). The Lagrangian multiplier test. *Ann. Math. Statist.* **30** 389–407.

DEPARTMENT OF ECONOMICS & RELATED STUDIES
UNIVERSITY OF YORK
HESLINGTON YORK Y01 5DD
ENGLAND