Distribution of the Residual Cross-Correlation in Univariate ARMA Time Series Models

A. IAN McLEOD*

Cross-correlations between univariate autoregressive moving average (ARMA) time series residuals are useful in the examination of relationships between time series (Pierce 1977a) and in the identification of dynamic regression models (Haugh and Box 1977). In this article, the asymptotic distribution of these residual cross-correlations is derived, and its application to the problem of testing for lagged relationships in the presence of instantaneous causality is discussed. Some results of a simulation study to investigate the accuracy of the asymptotic variances and covariances of the residual cross-correlations in finite samples are reported.

KEY WORDS: ARMA time series; Granger causality; Model identification; Relationships between time series; Residual cross-correlations.

1. INTRODUCTION

One approach to the problem of the elucidation of the relationship between two time series is to examine the cross-correlation function of the residuals of univariate models fitted to the two series. Interestingly enough, this approach seems to have first been suggested by Fisher (1921) with polynomial trend models. Of course, now it is understood that stochastic time series models, such as the ARMA model, are more realistic and perform better in applications such as forecasting (Box and Jenkins 1970, Ch. 1). Some recent applications of the univariate ARMA residual cross-correlation approach are mentioned in Sections 1.2 and 3. In this article, the large-sample distribution of the residual cross-correlations in univariate ARMA models is derived, and the finite sample accuracy of the derived asymptotic variances and covariances of the residual cross-correlations is investigated by simulation. An application to the problem of testing for lagged relationships in the presence of instantaneous causality is discussed and a brief economic example given.

1.1 Univariate ARMA Time Series Models

The theory and application of univariate ARMA models is discussed in a book by Box and Jenkins (1970).

Let $(w_{1,t}, w_{2,t})$, $-\infty < t < \infty$ be a discrete-time bivariate stationary Gaussian time series with mean zero. Suppose that $w_{h,t}$ can be represented as a univariate stationary and invertible ARMA time series of order (p_h, q_h):

$$\phi_h(B)w_{h,t} = \theta_h(B)a_{h,t} \qquad (1.1)$$

where

$$\phi_h(B) = 1 - \phi_{h1}B - \ldots - \phi_{hp_h}B^{p_h} ,$$
$$\theta_h(B) = 1 - \theta_{h1}B - \ldots - \theta_{hq_h}B^{q_h} ,$$

B is the backshift operator ($Bw_{h,t} = w_{h,t-1}$), and $a_{1,t}$ and $a_{2,t}$ are the individual innovation or white-noise series. The innovations $(a_{1,t}, a_{2,t})$ are then a bivariate Gaussian time series with mean zero and autocovariance function

$$\begin{aligned}\gamma_{a_h a_h}(l) &= \langle a_{h,t}a_{h,t+l}\rangle \\ &= \sigma_h^2 , \quad \text{if} \quad l = 0 , \quad h = 1,2 , \\ &= 0 , \quad \text{if} \quad l \neq 0 , \quad h = 1,2 , \end{aligned} \qquad (1.2)$$

where $\langle \cdot \rangle$ denotes mathematical expectation and σ_h^2 is the individual innovation variance for the time series $w_{h,t}$. The cross-covariance function of $a_{1,t}$ and $a_{2,t}$ is defined by

$$\gamma_{a_1a_2}(l) = \langle a_{1,t}a_{2,t+l}\rangle , \quad l = 0, \pm 1, \ldots, \qquad (1.3)$$

and it is assumed that

$$\sum_{l=-\infty}^{\infty} |l|\,|\gamma_{a_1a_2}(l)| < \infty . \qquad (1.4)$$

Given n observations, $w_{h,t}$, $t = 1, 2, \ldots, n$, from the time series, efficient univariate algorithms to estimate the model parameters $\boldsymbol{\beta}_h = (\phi_{h1}, \ldots, \phi_{hp_h}, \theta_{h1}, \ldots, \theta_{hq_h})$ have been described by Box and Jenkins (1970), Ljung and Box (1976), McLeod (1977b), and other researchers.

1.2 Cross-Correlations in Univariate ARMA Models

A number of authors (see Haugh and Box 1977 and references therein) have advocated the use of the cross-correlation function of $a_{1,t}$ and $a_{2,t}$,

$$\rho_{a_1a_2}(l) = \gamma_{a_1a_2}(l)/(\sigma_1\sigma_2) , \quad l = 0, \pm 1, \ldots, \qquad (1.5)$$

for elucidating the relationship between $w_{1,t}$ and $w_{2,t}$. To measure the strength of the relationship between $w_{1,t}$ and $w_{2,t}$, Pierce (1977a,b) has suggested the coefficient

$$R^2 = \sum_{l=-\infty}^{\infty} \rho_{a_1a_2}^2(l) . \qquad (1.6)$$

If $\rho_{a_1a_2}(l) \neq 0$ for some $l > 0$, Pierce and Haugh (1977) showed that $w_{1,t}$ is a useful predictor for $w_{2,t}$. Furthermore, if $\rho_{a_1a_2}(l) = 0$ for all $l < 0$ and $\rho_{a_1a_2}(l) \neq 0$ for some

* A. Ian McLeod is Assistant Professor, Statistics and Actuarial Science Group, Department of Mathematics, The University of Western Ontario, London, Canada N6A 5B9. The author wishes to thank the editor, an associate editor, and a referee for criticisms and suggestions that significantly improved the presentation of this article. The author also would like to thank M.E. Thompson for reading an earlier version of this article.

© Journal of the American Statistical Association
December 1979, Volume 74, Number 368
Theory and Methods Section

$l > 0$, then Haugh (1972a) and Haugh and Box (1977) have shown how a dynamic regression of $w_{2,t}$ on $w_{1,t}$ can be identified by using the innovation cross-correlation function $\rho_{a_1a_2}(\cdot)$.

For any given parameter value, say $\dot{\beta}_h$, the corresponding estimated innovation series $\dot{a}_{h,t}$, $t = 1, \ldots, n$ may be directly calculated from (1.1) either by setting $w_{h,t}$ and $a_{h,t}$ for $t \leq 0$ equal to their expected values conditional on $w_{h,1}, \ldots, w_{h,n}$ as described in Box and Jenkins (1970) and Newbold (1974) or more approximately by setting $w_{h,t}$ and $a_{h,t}$ for $t \leq 0$ equal to their unconditional expected value of zero. Also, for any given $\dot{\beta}_h$, the sample innovation cross-covariance and cross-correlation functions of $\dot{a}_{1,t}$ and $\dot{a}_{2,t}$ at lag l are defined by

$$\dot{c}_{a_1a_2}(l) = n^{-1} \sum_{t=1}^{n-l} \dot{a}_{1,t}\dot{a}_{2,t+l} \tag{1.7}$$

and

$$\dot{r}_{a_1a_2}(l) = \dot{c}_{a_1a_2}(l)/[(\dot{c}_{a_1a_1}(0)\dot{c}_{a_2a_2}(0))]^{\frac{1}{2}} , \tag{1.8}$$

respectively. It is easily shown that the absolute error in $\dot{r}_{a_1a_2}(l)$ from using either of these methods to calculate $\dot{a}_{h,t}$, $h = 1, 2, t = 1, \ldots, n$ is $O(1/n)$. Hence, if the exact value of the model parameters were known to be β_1 and β_2, then the large-sample variances and covariances of the cross-correlations of the calculated innovations, $a_{1,t}$, $t = 1, \ldots, n$ and $a_{2,t}$, $t = 1, \ldots, n$, can be obtained from a formula of Bartlett (1966, p. 332). This formula yields

$$\begin{aligned} n\cdot\operatorname{cov}(r_{a_1a_2}(l), r_{a_1a_2}(k)) &= \rho_{a_2a_2}(k - l) \\ &+ \rho_{a_1a_2}(k)\rho_{a_1a_2}(l)\Big(\sum_{i=-\infty}^{\infty} \rho_{a_1a_2}{}^2(i) - 3\Big) \\ &+ \sum_{i=-\infty}^{\infty} \rho_{a_1a_2}(l - i)\rho_{a_1a_2}(k + i) \ . \end{aligned} \tag{1.9}$$

Let $\hat{\beta}_h$ be a univariate asymptotically efficient estimate of β_h for $h = 1, 2$ and let $\hat{a}_{h,t}$, $t = 1, \ldots, n$ and $\hat{r}_{a_1a_2}(l)$, $l = 0, \pm 1, \ldots$ be the corresponding residuals and residual cross-correlations. Box and Pierce (1970) obtained the large-sample distribution of the residual autocorrelations in univariate ARMA time series models that can be shown to be equivalent to the large-sample distribution of the residual cross-correlations when $\rho_{a_1a_2}(0) = 1$ and $\rho_{a_1a_2}(l) = 0$, $l \neq 0$ (see Section 3). It follows that the large-sample covariances of the residual cross-correlations are not given by Bartlett's formula in the general case. Nevertheless, Haugh (1972a,b, 1976) showed that if the series $w_{1,t}$ and $w_{2,t}$ are independent (so $\rho_{a_1a_2}(l) = 0$, $l = 0, \pm 1, \ldots$), then the large-sample distribution of the residual cross-correlations is jointly normal with covariance matrix determined by

$$\begin{aligned} \operatorname{cov}(\hat{r}_{a_1a_2}(l), \hat{r}_{a_1a_2}(k)) &= 1/n \ , \quad \text{if} \quad l = k \ , \\ &= 0 \ , \quad \text{if} \quad l \neq k \ . \end{aligned} \tag{1.10}$$

Several authors (Haugh and Box 1977; Pierce and Haugh 1977; Pierce 1977a) have remarked that it would be useful to know the distribution of the residual cross-correlations in the general case. This is derived in Section 2.

2. GENERAL RESULT

In this section, the asymptotic joint distribution of the residual cross-correlations in model (1.1) is derived. For any fixed $M \geq 0$, let

$$\begin{aligned} \dot{\mathbf{r}} = (\dot{r}_{a_1a_2}(-1), \ldots, \dot{r}_{a_1a_2}(-M) \ , \ \dot{r}_{a_1a_2}(0) \ , \\ \dot{r}_{a_1a_2}(1), \ldots, \dot{r}_{a_1a_2}(M)) \ , \end{aligned} \tag{2.1}$$

and let

$$\begin{aligned} \boldsymbol{\rho} = (\rho_{a_1a_2}(-1), \ldots, \rho_{a_1a_2}(-M), \rho_{a_1a_2}(0), \\ \rho_{a_1a_2}(1), \ldots, \rho_{a_1a_2}(M)) \ . \end{aligned} \tag{2.2}$$

Let $\mathbf{r}$ and $\hat{\mathbf{r}}$ denote the vector $\dot{\mathbf{r}}$ when $\dot{\beta} = \beta$ and $\dot{\beta} = \hat{\beta}$, respectively. Thus, $\mathbf{r}$ and $\hat{\mathbf{r}}$ are vectors of innovation and residual cross-correlations. The derivation of the asymptotic joint distribution of $\hat{\mathbf{r}}$ is based on the use of a Taylor series linearization of $\hat{\mathbf{r}}$ as a function of $(\hat{\beta}_1, \hat{\beta}_2, \mathbf{r})$ and the asymptotic joint distribution of $(\hat{\beta}_1, \hat{\beta}_2, \mathbf{r})$.

Lemma 1: The distribution of $\hat{\mathbf{r}}$ does not depend on σ_1 or σ_2.

Remark 1: It follows from Lemma 1 that, without loss of generality, it can be assumed that $\gamma_{a_1a_1}(0) = \gamma_{a_2a_2}(0) = 1$.

The following lemma is useful for simplifying double summations of theoretical cross-covariances.

Lemma 2:

$$n^{-1} \sum_{k=1}^{n} \sum_{l=1}^{n} \gamma_{a_1a_2}(k - l) = \sum_{k=-\infty}^{\infty} \gamma_{a_1a_2}(k) + O(1/n) \ . \tag{2.3}$$

Proof: This follows from the assumption stated in (1.4).

Lemma 3:

$$\begin{aligned} r_{a_1a_2}(l) - \rho_{a_1a_2}(l) = c_{a_1a_2}(l) - \tfrac{1}{2}\gamma_{a_1a_2}(l)(c_{a_1a_1}(0) \\ + c_{a_2a_2}(0)) + O_p(1/n) \ . \end{aligned} \tag{2.4}$$

Proof: This follows from the Taylor series expansion of $\dot{r}_{a_1a_2}(l)$ as a function of $(\dot{c}_{a_1a_2}(l), \dot{c}_{a_1a_1}(0), \dot{c}_{a_2a_2}(0))$ about

$$(\gamma_{a_1a_2}(l), \gamma_{a_1a_1}(0), \gamma_{a_2a_2}(0))$$

and evaluated at

$$(c_{a_1a_2}(l), c_{a_1a_1}(0), c_{a_2a_2}(0)) \ .$$

Lemma 4: Let $\hat{\beta}_h$ be an asymptotically efficient estimate of β_h in the univariate ARMA model. Then,

$$\hat{\beta}_h - \beta_h = \mathbf{I}_h{}^{-1}\mathbf{S}_h + O_p(1/n) \tag{2.5}$$

where $\mathbf{I}_h$ is the large-sample information matrix per observation given by

$$\mathbf{I}_h = \left[\begin{array}{c|c} \gamma_{v_hv_h}(i - j) & \gamma_{v_hu_h}(i - j) \\ \hline \gamma_{u_hv_h}(i - j) & \gamma_{u_hu_h}(i - j) \end{array}\right]\begin{array}{l} p_h \\ q_h \end{array} \tag{2.6}$$

$$\qquad\qquad p_h \qquad\qquad\qquad q_h$$

where the (i, j) entry in each partitioned matrix is indicated and the auxiliary time series $v_{h,t}$ and $u_{h,t}$ are defined by

$$\phi_h(B)v_{h,t} = -a_{h,t} \tag{2.7}$$

and

$$\theta_h(B)u_{h,t} = a_{h,t} \ . \tag{2.8}$$

$\mathbf{S}_h = (S_{h,1}, \ldots, S_{h,p_h+q_h})$ is the score function

$$S_{h,i} = -n^{-1} \sum_{t=1}^{n} a_{h,t} v_{h,t-i} \,, \quad \text{if} \quad i = 1, \ldots, p_h \,;$$
$$= -n^{-1} \sum_{t=1}^{n} a_{h,t} u_{h,t+p_h-i} \,, \quad \text{if} \quad i = p_h + 1, \ldots, p_h + q_h \,. \quad (2.9)$$

Proof: The likelihood function of $\dot{\boldsymbol{\beta}}_h$ in the univariate ARMA model is

$$\log L_h = -\tfrac{1}{2} \sum_{t=1}^{n} \dot{a}_{h,t}^2 + m_h + O(r^n) \,, \quad (2.10)$$

where m_h is a rational function of the elements of $\boldsymbol{\beta}_h$ that does not depend on n and $0 < r < 1$ (McLeod 1977b). Also, it can be shown (Box and Jenkins 1970, p. 237) that

$$\partial \dot{a}_{h,t} / \partial \phi_{hi} = v_{h,t-i} \quad (2.11)$$

and

$$\partial \dot{a}_{h,t} / \partial \theta_{hi} = u_{h,t-i} \,, \quad (2.12)$$

where the partial derivative of $\dot{a}_{h,t}$ with respect to ϕ_{hi} evaluated at ϕ_{hi} is denoted by $\partial \dot{a}_{ht} / \partial \phi_{hi}$ (and similar notation is used throughout the article). It follows that

$$\partial \log L_h / \partial \beta_{h,i} = n S_{h,i} + O(1) \,. \quad (2.13)$$

Also, it can be shown that

$$\partial \log L_h / (\partial \boldsymbol{\beta}_h \partial \boldsymbol{\beta}_h^T) = -n \mathbf{I}_h + O(\sqrt{n}) \,. \quad (2.14)$$

The lemma now follows from the Taylor series expansion of $\partial \log L_h / \partial \dot{\boldsymbol{\beta}}_h$ about $\boldsymbol{\beta}_h$ evaluated at $\hat{\boldsymbol{\beta}}_h$.

Remark 2: Lemmas 3 and 4 present linearizations that are useful for handling the asymptotics of expressions involving $\mathbf{r}$ and $\hat{\boldsymbol{\beta}}_h$.

Lemma 5: The asymptotic joint distribution of $\sqrt{n}(\hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1)$, $\sqrt{n}(\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}_2)$ is normal with mean vector zero and covariance matrix

$$\mathbf{V} = \left[\begin{array}{c|c} \mathbf{I}_1^{-1} & \mathbf{I}_1^{-1}\mathbf{A}\mathbf{I}_2^{-1} \\ \hline \mathbf{I}_2^{-1}\mathbf{A}^T\mathbf{I}_1^{-1} & \mathbf{I}_2^{-1} \end{array}\right] \begin{matrix} p_1 + q_1 \\ p_2 + q_2 \end{matrix} \quad (2.15)$$
$$\qquad p_1 + q_1 \qquad p_2 + q_2$$

where

$$\mathbf{A} = \left[\begin{array}{c|c} \mathbf{A}^{(v_1,v_2)} & \mathbf{A}^{(v_1,u_2)} \\ \hline \mathbf{A}^{(u_1,v_2)} & \mathbf{A}^{(u_1,u_2)} \end{array}\right] \begin{matrix} p_1 \\ q_1 \end{matrix} \quad (2.16)$$
$$\qquad p_2 \qquad q_2$$

where the (i, j) element of the submatrix $\mathbf{A}^{(c,d)}$ is

$$A_{ij}^{(c,d)} = \sum_{k=-\infty}^{\infty} \gamma_{a_1a_2}(k)\gamma_{cd}(k + i - j) + \gamma_{a_1d}(k - j)\gamma_{ca_2}(k + i) \,. \quad (2.17)$$

Proof: The matrix $\mathbf{A}$ is obtained by straightforward calculation using Lemmas 2 and 4. The joint asymptotic normality of the estimates follows from Lemma 4 because any linear function of both $\mathbf{S}_1$ and $\mathbf{S}_2$ is the average of a series of martingale differences and so by the martingale central limit theorem (Billingsley 1961) is asymptotically normal.

Lemma 6: The asymptotic joint distribution of $(\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2, \mathbf{r})$ is normal with mean vector $(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\rho})$ and covariance matrix

$$\frac{1}{n}\left[\begin{array}{c|c} \mathbf{V} & -\boldsymbol{\Delta} \\ \hline -\boldsymbol{\Delta}^T & \mathbf{E} \end{array}\right] \begin{matrix} p_1 + q_1 + p_2 + q_2 \\ 2M + 1 \end{matrix} \quad (2.18)$$
$$\qquad p_1 + q_1 + p_2 + q_2 \qquad 2M + 1$$

where $\mathbf{V}$ is defined in Lemma 5, $\mathbf{E}/n$ is the large-sample covariance matrix of $\mathbf{r}$ that is determined directly from Bartlett's formula (1.9), and the jth column, $j = 1, \ldots, 2M + 1$, of $\boldsymbol{\Delta}$ is $(\boldsymbol{\delta}_j^{(1)}, \boldsymbol{\delta}_j^{(2)})^T$ where

$$\boldsymbol{\delta}_j^{(h)} = -\lim_{n\to\infty} n \cdot \operatorname{cov}(\hat{\boldsymbol{\beta}}_h, r_{a_1a_2}(k)) = \mathbf{I}_h^{-1}(\mathbf{f}^{(v_h)}, \mathbf{f}^{(u_h)})^T \quad (2.19)$$

where

$$h = 1, 2 \,,$$
$$k = -j \,, \quad \text{if} \quad j = 1, \ldots, M \,,$$
$$= j - M - 1 \,, \quad \text{if} \quad j = M + 1, \ldots, 2M + 1 \,,$$

$$f_i^{(c)} = \gamma_{ca_2}(k + i) + \sum_{l=-\infty}^{\infty} \gamma_{a_1a_2}(l)[\gamma_{a_1c}(k - i - l) - \gamma_{a_1a_2}(k)\gamma_{ca_2}(l + i)] \,, \quad \text{if} \quad c = v_1, u_1 \,,$$
$$= \gamma_{a_1c}(k - i) + \sum_{l=-\infty}^{\infty} \gamma_{a_1a_2}(l)[\gamma_{a_2c}(l - k - i) - \gamma_{a_1a_2}(k)\gamma_{ca_1}(l + i)] \,, \quad \text{if} \quad c = v_2, u_2 \,. \quad (2.20)$$

Proof: The computation of $\mathbf{f}^{(c)}$ follows directly from Lemmas 2, 3, and 4. The asymptotic joint normality is proved, as in Lemma 5.

Theorem: The asymptotic distribution of $\hat{\mathbf{r}}$ is normal with mean vector $\boldsymbol{\rho}$ and covariance matrix

$$\operatorname{var}(\hat{\mathbf{r}}) = (\mathbf{E} + \mathbf{X}\mathbf{V}\mathbf{X}^T - \mathbf{X}\boldsymbol{\Delta} - \boldsymbol{\Delta}^T\mathbf{X}^T)/n \quad (2.21)$$

where $\mathbf{E}$, $\mathbf{V}$, and $\boldsymbol{\Delta}$ are defined in Lemma 6 and

$$\mathbf{X} = \left[\begin{array}{c|c|c|c} \tau_{v_1a_2}(-i, j) & \tau_{u_1a_2}(-i, j) & \tau_{v_2a_1}(i, j) & \tau_{u_2a_1}(i, j) \\ \hline \tau_{v_1a_2}(0, j) & \tau_{u_1a_2}(0, j) & \tau_{v_2a_1}(0, j) & \tau_{u_2a_1}(0, j) \\ \hline \tau_{v_1a_2}(i, j) & \tau_{u_1a_2}(i, j) & \tau_{v_2a_1}(-i, j) & \tau_{u_2a_1}(-i, j) \end{array}\right] \begin{matrix} M \\ 1 \\ M \end{matrix} \quad (2.22)$$
$$\qquad p_1 \qquad q_1 \qquad p_2 \qquad q_2$$

where the (i, j) element in each partitioned matrix is indicated;

$$\tau_{ca_k}(l, j) = \gamma_{ca_k}(l + j) - \tfrac{1}{2}\gamma_{a_ha_k}(l)\gamma_{ca_k}(j) \quad (2.23)$$

where

$$c = u_h, v_h \,, \quad h, k = 1, 2 \,, \quad -M \leq |l| \leq M \,, \quad j = 1, 2, \ldots \,.$$

Proof: Let $\dot{\boldsymbol{\beta}} = (\dot{\boldsymbol{\beta}}_1, \dot{\boldsymbol{\beta}}_2)$, and let $\dot{\beta}_i$ denote the ith element of $\dot{\boldsymbol{\beta}}$. Then by expanding $\dot{r}_{a_1a_2}(i)$ in a Taylor series about $\boldsymbol{\beta}$ and evaluating $\dot{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}$ it follows that

$$\hat{r}_{a_1a_2}(i) = r_{a_1a_2}(i) + \sum_{j=1}^{p_1+q_1+p_2+q_2} (\hat{\beta}_j - \beta_j)\partial\dot{r}_{a_1a_2}(i)/\partial\beta_j + O_p(1/n) \ . \quad (2.24)$$

Consider the $M \times p_1$ submatrix of $\mathbf{X}$ corresponding to $i = 1, \ldots, M$ and $j = 1, \ldots, p_1$. The (i, j) element of this submatrix is

$$X_{ij} = \gamma_{v_1a_2}(j - i) - \tfrac{1}{2}\gamma_{a_1a_2}(-i)\gamma_{v_1a_2}(j) \ . \quad (2.25)$$

It follows directly from (2.11) that

$$\frac{\partial\dot{r}_{a_1a_2}(-i)}{\partial\phi_j} = \frac{c_{v_1a_2}(j - i)}{[(c_{a_1a_1}(0)c_{a_2a_2}(0))]^{\frac{1}{2}}} - \frac{1}{2}r_{a_1a_2}(-i)\frac{c_{v_1a_2}(j)}{c_{a_1a_1}(0)} \cdot \quad (2.26)$$

To determine the large-sample mean and variance of

$$r_{a_1a_2}(-i)c_{v_1a_2}(j)/c_{a_1a_1}(0) \ ,$$

note that by Lemma 3

$$r_{a_1a_2}(-i) = \gamma_{a_1a_2}(-i) + c_{a_1a_2}(-i) - \tfrac{1}{2}\gamma_{a_1a_2}(-i)[c_{a_1a_1}(0) + c_{a_2a_2}(0)] + O_p(1/n) \ , \quad (2.27)$$

and also by a Taylor series expansion it can be shown that

$$c_{v_1a_2}(j)/c_{a_1a_1}(0) = \gamma_{v_1a_2}(j) + c_{v_1a_2}(j) - \gamma_{v_1a_2}(j)c_{a_1a_1}(0) + O_p(1/n) \ . \quad (2.28)$$

Because the variances and covariances of the sample cross-covariances in (2.27) and (2.28) are $O(1/n)$, it follows that

$$\operatorname{var}(r_{a_1a_2}(-i)c_{v_1a_2}(j)/c_{a_1a_1}(0)) = O(1/n) \quad (2.29)$$

and

$$\langle r_{a_1a_2}(-i)c_{v_1a_2}(j)/c_{a_1a_1}(0)\rangle = \gamma_{a_1a_2}(-i)\cdot\gamma_{v_1a_2}(j) + O(1/n) \ . \quad (2.30)$$

Similarly, it can be shown that

$$c_{v_1a_2}(j - i)/([c_{a_1a_1}(0)c_{a_2a_2}(0)])^{\frac{1}{2}}$$

has variance $O(1/n)$ and mean equal to $\gamma_{v_1a_2}(j - i) + O(1/n)$. Hence, it follows that

$$\operatorname{var}(\partial\dot{r}_{a_1a_2}(-i)/\partial\phi_j) = O(1/n) \quad (2.31)$$

and that

$$\langle\partial\dot{r}_{a_1a_2}(-i)/\partial\phi_j\rangle = \gamma_{v_1a_2}(j - i) - \tfrac{1}{2}\gamma_{a_1a_2}(-i)\cdot\gamma_{v_1a_2}(j) + O(1/n) \ . \quad (2.32)$$

It follows from Chebyshev's inequality that

$$\partial\dot{r}_{a_1a_2}(-i)/\partial\phi_j = \gamma_{v_1a_2}(j - i) - \tfrac{1}{2}\gamma_{a_1a_2}(-i)\cdot\gamma_{v_1a_2}(j) + O_p(1/\sqrt{n}) \ . \quad (2.33)$$

In general, it can be shown that

$$\partial\dot{r}_{a_1a_2}(l)/\partial\beta_j = X_{ij} + O_p(1/\sqrt{n}) \quad (2.34)$$

where $-M \le l \le M$, $j = 1, \ldots, p_1 + q_1 + p_2 + q_2$,

$$\begin{aligned} i &= |l| \ , && \text{if} \quad l < 0 \ , \\ &= M + l + 1 \ , && \text{if} \quad l \ge 0 \ , \end{aligned}$$

and X_{ij} is the (i, j) element of $\mathbf{X}$.

Because $(\hat{\beta}_j - \beta_j)$ is $O_p(1/\sqrt{n})$ it follows from the theorem of Mann and Wald (1943, Corollary 1) that

$$\hat{\mathbf{r}} = \mathbf{r} + \mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) + O_p(1/n) \ . \quad (2.35)$$

The theorem now follows directly from Lemma 6 and a theorem given by Rao (1974, (2c.4.12)).

Remark 3: If the assumption of joint normality of $a_{t,h}$; $t = 1, \ldots, n$; $h = 1, 2$ is invalid, (2.21) will involve fourth-order cumulants.

Remark 4: It is not difficult to show that the theorem also applies to the case of two time series with nonzero means if the series are corrected for their sample means.

3. AN APPLICATION

The general result of Section 2 can be simplified in certain special cases. Many economic time series have the property that the largest residual cross-correlation is at lag zero (Pierce 1977a). In this section, the distribution of the residual cross-correlations when only the lag-zero innovation cross-correlation is nonzero is obtained and is used to derive a test for lagged relationships between economic time series.

3.1 Instantaneous Causality Only

Suppose the time series $w_{1,t}$ and $w_{2,t}$ are generated by the model (1.1) and the cross-correlation function between $a_{1,t}$ and $a_{2,t}$ is given by

$$\begin{aligned} \rho_{a_1a_2}(l) &= \rho \ , && \text{if} \quad l = 0 \ , \\ &= 0 \ , && \text{if} \quad l \ne 0 \ . \end{aligned} \quad (3.1)$$

Then the relationship between $w_{1,t}$ and $w_{2,t}$ may be said to be one of instantaneous causality only (Pierce and Haugh 1977, 1979).[1]

Bartlett's formula (1.9) for the large-sample variances and covariances of the sample cross-correlations of $a_{1,t}$ and $a_{2,t}$ yields

$$\begin{aligned} \operatorname{var}(r_{a_1a_2}(l)) &= (1 - \rho^2)^2/n \ , && \text{if} \quad l = 0 \ , \\ &= 1/n \ , && \text{if} \quad l \ne 0 \ , \end{aligned} \quad (3.2)$$

and

$$\begin{aligned} &\operatorname{cov}(r_{a_1a_2}(l), r_{a_1a_2}(k)) \\ &\quad = 0 \ , \quad \text{if} \quad l \ne k \ , \quad l \ne -k \\ &\quad = \rho^2/n \ , \quad \text{if} \quad l = -k \ , \quad l \ne 0 \ . \end{aligned} \quad (3.3)$$

Let

$$\hat{\mathbf{r}}^{(1)} = (\hat{r}_{a_1a_2}(-1), \ldots, \hat{r}_{a_1a_2}(-M)) \quad (3.4)$$

and

$$\hat{\mathbf{r}}^{(2)} = (\hat{r}_{a_1a_2}(1), \ldots, \hat{r}_{a_1a_2}(M)) \ . \quad (3.5)$$

[1] Empirical methods for detecting causality relationships between time series have been discussed by Granger (1969) and Pierce and Haugh (1977).

Then it follows from the theorem in Section 2 that

$$\operatorname{var}(\hat{\mathbf{r}}^{(h)}) = \mathbf{P}_h/n \quad (3.6)$$

where

$$\mathbf{P}_h = \mathbf{1}_M - \rho^2 \mathbf{X}_h \mathbf{I}_h^{-1} \mathbf{X}_h^T \ , \quad h = 1, 2 \ , \quad (3.7)$$

where $\mathbf{1}_M$ is the $M \times M$ identity matrix, $\mathbf{I}_h$ is defined in (2.7), and

$$\mathbf{X}_h = (\underset{p_h}{-\pi_{h,i-j}} \,|\, \underset{q_h}{\psi_{h,i-j}})M \quad (3.8)$$

where the (i, j) element in each partitioned matrix is indicated, $\pi_{h,l} = -\gamma_{v_h a_h}(-l)$ and $\psi_{h,l} = \gamma_{u_h a_h}(-l)$. The coefficients $\pi_{h,l}$ and $\psi_{h,l}$ are easily calculated recursively (Box and Jenkins 1970, pp. 132–134) by using the identities $1/\phi_h(B) = \sum \pi_{h,k} B^k$ and $1/\theta_h(B) = \sum \psi_{h,k} B^k$. The elements of the information matrix $\mathbf{I}_h$ may be calculated by solving a set of linear equations as in McLeod (1975, 1977a). Also, from the theorem of Section 2,

$$\operatorname{cov}(\hat{\mathbf{r}}^{(h)}, \hat{r}_{a_1 a_2}(0)) = \mathbf{0} \ , \quad h = 1, 2 \quad (3.9)$$

and

$$\operatorname{cov}(\hat{\mathbf{r}}^{(1)}, \hat{\mathbf{r}}^{(2)}) = (\rho^2 \mathbf{J} + 2\mathbf{1}_M - \mathbf{P}_1 - \mathbf{P}_2 + \rho^2 \mathbf{X}_1 \mathbf{I}_1^{-1} \mathbf{A} \mathbf{I}_2^{-1} \mathbf{X}_2^T)/n \quad (3.10)$$

where $\mathbf{J}$ is the $M \times M$ matrix with 1's on the diagonal at right angles to the main diagonal and 0's elsewhere and

$$\mathbf{A} = \left[\begin{array}{c|c} \gamma_{v_1 v_2}(i-j) & \gamma_{v_1 u_2}(i-j) \\ \hline \gamma_{u_1 v_2}(i-j) & \gamma_{u_1 u_2}(i-j) \end{array}\right]\begin{matrix} p_1 \\ q_1 \end{matrix} . \quad (3.11)$$
$$\qquad\quad p_2 \qquad\qquad\qquad q_2$$

Finally, the large-sample variance of the estimate $\hat{\rho} = \hat{r}_{a_1 a_2}(0)$, of ρ is

$$\operatorname{var}(\hat{\rho}) = (1 - \rho^2)^2/n \ . \quad (3.12)$$

If $\rho = 0$, the asymptotic variances of the residual cross-correlations are all equal to $1/n$, but from (3.7) it can be shown that when ρ^2 is close to one the asymptotic variances of $\hat{r}_{a_1 a_2}(l)$ may be significantly less than $1/n$. In fact, when $\rho = 1$, the residual cross-correlations and residual autocorrelations have the same asymptotic distributions. McLeod (1977a, 1978) has shown that, for large n, any fixed $M \geq 1$ and $h = 1, 2$ the covariance matrix of the residual autocorrelations

$$(\hat{r}_{a_h a_h}(1), \ldots, \hat{r}_{a_h a_h}(M)) \quad (3.13)$$

is given by

$$(\mathbf{1}_M - \mathbf{X}_h \mathbf{I}_h^{-1} \mathbf{X}_h^T)/n \ . \quad (3.14)$$

It also can be shown (McLeod 1977a) that the covariance matrix (3.14) is exactly equivalent to that derived by Box and Pierce (1970) (also see Durbin 1970).

3.2 Test for Lagged Relationships

To test the null hypotheses

$$H_0^{(1)}: \rho_{a_1 a_2}(-1) = \ldots = \rho_{a_1 a_2}(-M) = 0$$

or

$$H_0^{(2)}: \rho_{a_1 a_2}(1) = \ldots = \rho_{a_1 a_2}(M) = 0$$

against the simple negation of $H_0^{(1)}$ or $H_0^{(2)}$, respectively, when $\rho \neq 0$, the following test statistic is suggested:

$$\hat{Q}_M^{(h)} = n(\hat{\mathbf{r}}^{(h)})^T \hat{\mathbf{P}}_h^{-1} \hat{\mathbf{r}}^{(h)} \ , \quad (3.15)$$

where $\hat{\mathbf{P}}_h$ denotes the matrix $\mathbf{P}_h$ in (3.7) calculated by using $\hat{\boldsymbol{\beta}}_h$ rather than $\boldsymbol{\beta}_h$. If both $H_0^{(1)}$ and $H_0^{(2)}$ are true, $\hat{Q}_M^{(h)}$ will be asymptotically χ^2-distributed on M degrees of freedom and large values of $\hat{Q}_M^{(h)}$ will provide evidence against $H_0^{(h)}$. This test of $H_0^{(h)}$ may be compared with the test suggested by Pierce (1977a) that is based on the statistic

$$Q_M^{(h)} = n(\hat{\mathbf{r}}^{(h)})^T \hat{\mathbf{r}}^{(h)} \ . \quad (3.16)$$

If $\rho = 0$ and $H_0^{(1)}$ and $H_0^{(2)}$ are both true, then $Q_M^{(h)}$ is also $\chi^2(M)$ for large n. The example given in the following paragraph shows that the test based on $\hat{Q}_M^{(h)}$ may be more sensitive than that using $Q_M^{(h)}$ when $\rho \neq 0$.

From the results of Davies, Triggs, and Newbold (1977) and Ljung and Box (1978) on the portmanteau significance test of Box and Pierce (1970), it may be expected that both these tests using $Q_M^{(h)}$ and $\hat{Q}_M^{(h)}$ may considerably underestimate the true significance level if M is fairly large. In the case when $\rho = 0$, Haugh (1976) provided simulation evidence that

$$\operatorname{var}(\hat{r}_{a_1 a_2}(l)) \doteq (n - l)/n^2 \quad (3.17)$$

and suggested a modified test using this result. A general alternative approach would be to use shrinked residual cross-correlation estimates, such as

$$\tilde{r}_{a_1 a_2}(l) = \hat{r}_{a_1 a_2}(l)([(n - |l|)/n])^{\frac{1}{2}} \ . \quad (3.18)$$

Estimated standard deviations of the residual cross-correlations may be obtained by using estimated values of ρ, $\boldsymbol{\beta}_1$, and $\boldsymbol{\beta}_2$ in (3.7) and (3.12).

3.3 Example

Haugh (1976, p. 383) found that the first differences of two quarterly interest rate time series could be modeled by

$$w_{1,t} = .069 + (1 + .55B)\hat{a}_{1,t} \quad (3.19)$$

and

$$(1 - .76B + .39B^2)w_{2,t} = \hat{a}_{2,t} \ , \quad (3.20)$$

where $w_{h,t}$ and $\hat{a}_{h,t}$ are, respectively, the first differences and estimated innovations of series h for $h = 1, 2$. Also, in this example, $n = 71$ and $\hat{\rho} = .64$. The residual cross-correlations and their estimated standard errors calculated from (3.7) and (3.12) are shown in Table 1.

Note, that in Table 1, $\hat{r}_{a_1 a_2}(1)$ is significant at 5 percent, although it is not significant at 5 percent when compared with the benchmark standard deviation of $n^{-\frac{1}{2}} = .119$. The statistics for testing $H_0^{(1)}$ and $H_0^{(2)}$ have the following values:

h	1	2
$\hat{Q}_4^{(h)}$	1.34	14.06
$Q_4^{(h)}$	1.28	8.85

1. Residual Cross-Correlations and Estimated Standard Deviations

l	−4	−3	−2	−1	0	1	2	3	4
$\hat{r}_{a_1a_2}(l)$	.04	.10	.00	−.08	.64	.20	−.13	−.26	.01
Estimated standard deviation	.118	.117	.113	.100	.070	.096	.102	.109	.117

Thus, $H_0^{(1)}$ is not significant at 5 percent for both tests. $H_0^{(2)}$ is significant at 5 percent if the test statistic $\hat{Q}_4^{(2)}$ is used, and it is not significant at 5 percent if the less-sensitive test based on $Q_4^{(2)}$ is used.

4. SIMULATION STUDY

A simulation study was done to examine the accuracy of the asymptotic variances and covariances of the residual cross-correlations in the case of two first-order autoregressions,

$$(1 - \phi_1 B)w_{1,t} = a_{1,t} \qquad (4.1)$$

and

$$(1 - \phi_2 B)w_{2,t} = a_{2,t} \ , \qquad (4.2)$$

where $t = 1, \ldots, n$, $\rho_{a_1a_2}(0) = \rho$, $\rho_{a_1a_2}(l) = 0$ if $l \neq 0$, and $a_{1,t}$ and $a_{2,t}$ are Gaussian white noise with unit variance. The theoretical large-sample covariances of the residual cross-correlations are shown in Table 2.

A total of 225 models corresponding to the parameter settings $\phi_1, \phi_2 = 0, \pm.5, \pm.9$, $\rho = .3, .6, .9$ and $n = 50$, 200, 400 were included, and for each model 1,000 simulations were done. A multiplicative congruential random-number generator with modulus 2^{35} and multiplier 5^{15} (recommended by Conveyou and MacPherson 1967) was used in conjunction with the method of Marsaglia and Bray (1964) to generate independent normal pseudorandom numbers. The method of generating initial values of the time series can be quite important (McLeod and Hipel 1978). The following technique was used:

1. Initial values in the time series were generated by using the covariance matrix of $(w_{1,1}, w_{2,1})$.
2. A linear transformation was made on successive pairs of independent normal random numbers to obtain the simulated bivariate white-noise series $(a_{1,t}, a_{2,t})$, $t = 2, 3, \ldots, n$.
3. The remaining values of the time series were calculated recursively by using (4.1) and (4.2).

2. Asymptotic Covariances of the Residual Cross-Correlations in Two First-Order Autoregressions With Instantaneous Causality Only

l,k	$n \cdot \text{cov}(\hat{r}_{a_1a_2}(l), \hat{r}_{a_1a_2}(k))$
$l = 0$, any k	$\delta_{0,k}(1 - \rho^2)^2$
$l,k < 0$	$\delta_{l,k} - \rho^2\phi_1^{\lvert l+k\rvert-2}(1 - \phi_1^2)$
$l,k > 0$	$\delta_{l,k} - \rho^2\phi_2^{l+k-2}(1 - \phi_2^2)$
$l < 0, k > 0$	$\rho^2\delta_{-l,k} - \rho^2\phi_1^{k-l-2}(1 - \phi_1^2) - \rho^2\phi_2^{k-l-2}(1 - \phi_2^2)$ $+ \rho^4\phi_1^{\lvert l\rvert-1}\phi_2^{k-1}(1 - \phi_1^2)(1 - \phi_2^2)/(1 - \phi_1\phi_2)$

NOTE: $\delta_{l,k} = 1$, if $l = k$
$= 0$, if $l \neq k$.

The parameters ϕ_1 and ϕ_2 were estimated by using the lag-one sample autocorrelation coefficients and the residual cross-correlation vector, $\hat{\mathbf{r}}$, defined in (2.1) with $M = 2$ was calculated. For each model, the sample covariance matrix of $\hat{\mathbf{r}}$ was calculated by correcting the sample second moment of $\hat{\mathbf{r}}$ by its sample mean.

Let $\bar{\mathbf{C}}$ denote the sample estimate, based on 1,000 simulations, of the covariance matrix of the 5×1 vector $\hat{\mathbf{r}}$, and let $\mathbf{C}$ be the corresponding asymptotic covariance matrix of $\hat{\mathbf{r}}$ determined from Table 2. Then, assuming that $\mathbf{C}$ provides a good finite-sample approximation, $\bar{\mathbf{C}}$ has an asymptotic normal distribution (Anderson 1958, p. 75) with mean $\mathbf{C}$ and covariance matrix determined by

$$\text{cov}(\bar{C}_{ij}, \bar{C}_{kl}) = (C_{ik}C_{jl} + C_{il}C_{jk})/1{,}000 \ , \qquad (4.3)$$

where $\bar{C}_{ij}$ and C_{ij} denote the (i, j) element of $\bar{\mathbf{C}}$ and $\mathbf{C}$. Let $\bar{\boldsymbol{\varepsilon}}$ and $\boldsymbol{\varepsilon}$ be the 15×1 vectors corresponding to the elements of $\bar{\mathbf{C}}$ and $\mathbf{C}$ on or above the main diagonal taken in lexicographical order and let $\boldsymbol{\Xi}$ be the asymptotic covariance matrix of $\bar{\boldsymbol{\varepsilon}}$ determined by (4.3) and Table 2. If the covariance matrix of $\hat{\mathbf{r}}$ is well approximated by $\mathbf{C}$, T^2 should be $\chi^2(15)$, where

$$T^2 = (\bar{\boldsymbol{\varepsilon}} - \boldsymbol{\varepsilon})^T\boldsymbol{\Xi}^{-1}(\bar{\boldsymbol{\varepsilon}} - \boldsymbol{\varepsilon}) \ . \qquad (4.4)$$

Large values of T^2 will tend to indicate that the asymptotic covariances given in Table 2 are not valid. The test statistic T^2 was evaluated by using double-precision arithmetic and the Cholesky decomposition method (Maindonald 1976). For each of the 25 models corresponding to a fixed value of n and ρ, the number of values of T^2 significant at the 5 and 1 percent levels and the mean T^2 value were calculated and the results are shown

3. Summary of T^2 Tests

n	Summary Result	ρ = .3	.6	.9
50	Mean	24.72	37.38	621.63
	>25.00[a]	9	20	25
	>35.58	5	16	25
200	Mean	15.12	19.43	57.70
	>25.00	3	5	24
	>30.58	1	2	20
400	Mean	14.76	16.70	25.80
	>25.00	2	2	11
	>30.58	0	2	7

[a] The 5% and 1% critical values are, respectively, 25.00 and 30.58.
NOTE: For each value of n and ρ there are 25 models. Each T^2 is calculated from 1,000 simulations of one model.

in Table 3. The main conclusion to be drawn from these results is that the accuracy of the asymptotic approximation depends not only on n but also on ρ. For example, if $|\rho| \leq .6$, the asymptotic approximation appears to be quite good if $n \geq 400$ (because the probability of four or more rejections at the 5 percent level is .24), but for larger values of $|\rho|$, larger values of n are required.

[*Received March 1977. Revised April 1979.*]

REFERENCES

Anderson, T.W. (1958), *An Introduction to Multivariate Statistical Analysis*, New York: John Wiley & Sons.

Bartlett, M.S. (1966), *Stochastic Processes* (2nd ed.), Cambridge, England: Cambridge University Press.

Billingsley, Patrick (1961), "The Lindeberg-Levy Theorem for Martingales," *Proceedings of the American Mathematical Society*, 12, 788–792.

Box, George E.P., and Jenkins, Gwilym M. (1970), *Time Series Analysis Forecasting and Control*, San Francisco: Holden-Day.

———, and Pierce, David A. (1970), "Distribution of the Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models," *Journal of the American Statistical Association*, 65, 1509–1526.

Conveyou, R.R., and MacPherson, R.D. (1967), "Fourier Analysis of Uniform Random Generators," *Journal of the Association of Computing Machinery*, 14, 100–119.

Davies, N., Triggs, C.M., and Newbold, P. (1977), "Significance Levels of the Box-Pierce Portmanteau Statistic in Finite Samples," *Biometrika*, 64, 517–522.

Durbin, James (1970), "Testing for Serial Correlation in Least-Squares Regression When Some of the Regressors Are Lagged Dependent Variables," *Econometrica*, 38, 410–421.

Fisher, R.A. (1921), "Contribution to the Discussion of the Paper 'On the Time-Correlation Problem,' by G.U. Yule," *Journal of the Royal Statistical Society*, 84, 534–536.

Granger, C.W.J. (1969), "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," *Econometrica*, 37, 424–438.

Haugh, Larry D. (1972a), "The Identification of Time Series Interrelationships With Special Reference to Dynamic Regression," unpublished PhD thesis, Dept. of Statistics, University of Wisconsin, Madison.

——— (1972b), "Checking Time Series Interrelationships and Identifying Dynamic Regression Models for Short-Term Forecasting," *Proceedings of the American Statistical Association*, Business and Economic Statistics Section, 325–330.

——— (1976), "Checking the Independence of Two Covariance-Stationary Time Series: A Univariate Residual Cross-Correlation Approach," *Journal of the American Statistical Association*, 71, 378–385.

———, and Box, George E.P. (1977), "Identification of Dynamic Regression (Distributed Lag) Models Connecting Two Time Series," *Journal of the American Statistical Association*, 72, 121–130.

Ljung, G.M., and Box, G.E.P. (1976), "Maximum Likelihood Estimation in the Autoregressive Moving Average Model," Technical Report No. 476, Dept. of Statistics, University of Wisconsin, Madison.

——— (1978), "On a Measure of Lack of Fit in Time Series Models," *Biometrika*, 65, 297–303.

Maindonald, J.H. (1976), "Least Squares Programs—A Second Look," *The American Statistician*, 30, 202–203.

Mann, Henry B., and Wald, Abraham (1943), "On Stochastic Limit and Order Relationships," *Annals of Mathematical Statistics*, 14, 217–226.

Marsaglia, G., and Bray, T.A. (1964), "A Convenient Method for Generating Normal Variables," *Society for Industrial and Applied Mathematics Review*, 6, 260–264.

McLeod, A. Ian (1975), "Derivation of the Theoretical Autocovariance Function of Autoregressive-Moving Average Time Series," *Applied Statistics*, 24, 255–256.

——— (1977a), "Topics in Time Series and Intervention Analysis," unpublished PhD thesis, Dept. of Statistics, University of Waterloo.

——— (1977b), "Improved Box-Jenkins Estimators," *Biometrika*, 64, 531–534.

——— (1978), "On the Distribution of Residual Autocorrelations in Box-Jenkins Models," *Journal of the Royal Statistical Society*, 40, 296–302.

———, and Hipel, Keith W. (1978), "Simulation Procedures for Box-Jenkins Models," *Water Resources Research*, 14, 969–975.

Newbold, Paul (1974), "The Exact Likelihood Function for a Mixed Autoregressive-Moving Average Process," *Biometrika*, 61, 423–426.

Pierce, David A. (1977a), "Relationships—and Lack Thereof—Between Economic Time Series, With Special Reference to Money and Interest Rates," *Journal of the American Statistical Association*, 72, 11–21.

——— (1977b), "R^2-Measures for Time Series," Special Studies Paper, Federal Reserve Board, Washington, D.C.

———, and Haugh, Larry D. (1977), "Causality in Temporal Systems: Characterizations and a Survey," *Journal of Econometrics*, 5, 265–293.

——— (1979), "The Characterization of Instantaneous Causality: A Comment," *Journal of Econometrics*, 10, 257–259.

Rao, C.R. (1974), *Linear Statistical Inference* (2nd ed.), New York: John Wiley & Sons.