Testing Causality Between Two Vectors in Multivariate Autoregressive Moving Average Models

HAFIDA BOUDJELLABA, JEAN-MARIE DUFOUR, and ROCH ROY*

In the analysis of economic time series, a question often raised is whether a vector of variables causes another one in the sense of Granger. Most of the literature on this topic is concerned with bivariate relationships or uses finite-order autoregressive specifications. The purpose of this article is to develop a causality analysis in the sense of Granger for general vector autoregressive moving average (ARMA) models. We give a definition of Granger noncausality between vectors, which is a natural and simple extension of the notion of Granger noncausality between two variables. In our context, this definition is shown to be equivalent to a more complex definition proposed by Tjostheim. For the class of linear invertible processes, we derive a necessary and sufficient condition for noncausality between two vectors of variables when the latter do not necessarily include all the variables considered in the analysis. This result is then specialized to the class of stationary invertible ARMA processes. Further, relatively simple necessary and sufficient conditions are obtained for two important cases: (1) the case where the two vectors reduce to two variables inside a larger vector including other variables; and (2) the case where the two vectors embody all the variables considered. Test procedures for these necessary and sufficient conditions are discussed. Among other things, it is noted that the necessary and sufficient conditions for noncausality may involve singularities at which standard asymptotic regularity conditions do not hold. To deal with such situations, we propose a sequential approach that leads to bounds tests. Finally, the tests suggested are applied to Canadian money and income data. The tests are based on bivariate and trivariate models of changes in nominal income and two money stocks (M1 and M2). In contrast with the evidence based on bivariate models, we find from the trivariate model that money causes income unidirectionally.

KEY WORDS: Bounds test; Causality test; Granger causality; Invertible linear process; Sequential procedure.

In the analysis of economic time series, a question often raised is whether a vector of variables causes another one in the sense of Granger (1969). Most of the literature on this topic is concerned with bivariate relationships or uses finite-order autoregressive (AR) specifications; see the reviews of Geweke (1984a), Gouriéroux and Monfort (1990, chap. X), Newbold (1982), and Pierce and Haugh (1977). In particular, little attempt has been made to characterize and test causality in the context of multivariate autoregressive moving average (ARMA) models. Because an AR model can only approximate a moving average (MA) model and may require the estimation of a large number of parameters to do so, multivariate ARMA models can be considerably more parsimonious than AR models and thus lead to more powerful tests.

Characterizing and testing causality is, however, more complicated in multivariate ARMA models than in pure AR or MA models. Kang (1981) derived a necessary and sufficient condition for noncausality in a general bivariate ARMA model and suggested that a likelihood ratio test could be based on this condition, but did not discuss its implementation. Similarly, Newbold (1982) suggested using a likelihood ratio to test a sufficient (but not necessary) condition of noncausality in a bivariate ARMA model, and Eberts and Steece (1984), Newbold and Hotopp (1986) and Taylor (1989) studied Wald, likelihood ratio, and Rao score tests of the necessary and sufficient condition of noncausality again in a bivariate model. For the multivariate case with more than two variables, Tjostheim (1981) gave a formulation of the concept of Granger causality in a general multivariate situation and developed a test procedure for causality in multivariate autoregressions. Hsiao (1982) also proposed a generalization of the Granger notion of causality to make some provision for spurious and indirect causality that may arise in multivariate analysis; however, most of his results were limited to trivariate situations. Osborn (1984) examined Granger causality in multivariate ARMA models by rewriting the model so that the autoregressive polynomials were the same for all the variables. Causality tests were then based on MA coefficients only. This approach does not take into account all the restrictions implied by the vector ARMA specification and may require estimating an unduly large number of moving average and autoregressive coefficients. Similarly, even though noncausality hypotheses are in principle easy to test in the context of vector autoregressive models (VAR), like those suggested by Doan, Litterman, and Sims (1984), Litterman and Weiss (1985), and Sims (1980a,b), such models may require a large number of parameters to represent even simple vector ARMA models, especially when the MA coefficients are large. Note also that a subvector of a stationary VAR process is a vector ARMA process but not necessarily a VAR process. In a more general context, measures of linear dependence and feedback between multivariate time series were defined by Geweke (1982, 1984b).

The purpose of this article is to develop a causality analysis for general vector ARMA models. In Section 1 we give a simple definition of Granger causality between vectors and

* Hafida Boudjellaba is Assistant Professor, Department of Mathematics and Computer Sciences, Laurentian University, Ramsey Lake Road, Sudbury, Ontario, P3E 2C6. Jean-Marie Dufour is Professor, Département de sciences économiques and Centre de recherche et développement en économique (CRDE), Université de Montréal, C.P. 6128, succursale A, Montréal, Québec, H3C 3J7. Roch Roy is Professor, Département d'informatique et de recherche opérationnelle and CRDE, Université de Montréal, C.P. 6128, succursale A, Montréal, Québec, H3C 3J7. This work was supported by the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, and the Fonds FCAR (Government of Quebec). The authors thank the editor, John Galbraith, Eric Ghysels, Ian McLeod, Marc Moore, Saïd Nsiri, Franz Palm, Benedikt Pötscher, Jacques Raynauld, Eric Renault, and Kenneth G. Stewart for several useful comments and discussions.

Journal of the American Statistical Association
December 1992, Vol. 87, No. 420, Theory and Methods

point out its equivalence with the more complex alternative definition proposed by Tjostheim (1981). In Section 2 we consider the class of linear invertible processes and give a necessary and sufficient condition for noncausality between two vectors of variables when the latter do not necessarily include all the variables considered in the analysis. Even though this condition was stated by Kang (1981) for a bivariate process and has been used as a *sufficient* condition for noncausality (e.g., in analyses based on VAR models), a proof for the general case does not seem to be available elsewhere. In Section 3 we specialize the latter result to the class of stationary invertible ARMA processes. Further, we obtain simpler necessary and sufficient conditions for two important cases: (1) the case where the two vectors reduce to two variables inside a larger vector whose past belongs to the information set used to predict (Theorem 1), and (2) the case where the two vectors embody all the variables considered (Theorems 2 and 3). These simpler conditions are formulated in terms of determinants of matrices built from submatrices of the original matrices of AR and MA polynomials. These conditions can be considerably more tractable from the point of view of implementing tests. In Section 4 we discuss test procedures for the necessary and sufficient conditions previously obtained. Among other things, we note that the necessary and sufficient conditions for noncausality may easily involve singularities at which standard asymptotic regularity conditions do not hold. To deal with such situations, we propose a sequential approach that leads to bounds tests. Finally, in Section 5 we apply the tests proposed to Canadian money and income data previously studied by Hsiao (1979). The tests are based on bivariate and trivariate models of changes in nominal income and two money stocks (M1 and M2) specified using the methodology of Tiao and Box (1981). In contrast with the evidence based on bivariate models, we find from the trivariate model that money causes income unidirectionally.

1. CAUSALITY BETWEEN VECTORS

Let $\{\mathbf{X}_t : t \in Z\}$ and $\{\mathbf{Y}_t : t \in Z\}$ be two multivariate second-order stationary stochastic processes on the integers Z, suppose that the dimension of $\mathbf{X}_t$ is n, and write $\mathbf{X}_t = (X_{1t}, \ldots, X_{nt})'$. Let A_t be an "information set" containing $\mathbf{X}_t$ and $\mathbf{Y}_t$ and denote $\bar{A}_t = \{A_s : s < t\}$. For any information set I_t, the best mean square linear predictor of X_{it} is denoted $P(X_{it} \mid I_t)$, $\varepsilon_{it}(X_{it} \mid I_t) = X_{it} - P(X_{it} \mid I_t)$ is the corresponding prediction error, and $\sigma^2(X_{it} \mid I_t)$ is the variance of ε_{it}. The predictor $P(X_{it} \mid I_t)$ is the orthogonal projection of X_{it} on the Hilbert space spanned by the variables in I_t. For Gaussian processes, $P(X_{it} \mid I_t) = E[X_{it} \mid I_t]$, but this property does not hold in general; see, for example, Brockwell and Davis (1991, sec. 2.7) or Priestley (1981, chap. 10). The best linear predictor of $\mathbf{X}_t$ is the vector $P(\mathbf{X}_t \mid I_t) = (P(X_{1t} \mid I_t), \ldots, P(X_{nt} \mid I_t))'$, the corresponding vector of prediction errors is given by $\varepsilon_t(\mathbf{X}_t \mid I_t) = (\varepsilon_{1t}(X_{1t} \mid I_t), \ldots, \varepsilon_{nt}(X_{nt} \mid I_t))'$, and we denote $\Sigma(\mathbf{X}_t \mid I_t)$ the covariance matrix of ε_t.

In the sequel we will use the following definition of noncausality, which is a natural and simple extension of the notion of noncausality between two univariate stochastic processes. The set $\bar{A}_t - \bar{\mathbf{Y}}_t$ represents all the information in $\bar{A}_t$ apart from the information in $\bar{\mathbf{Y}}_t$.

Definition 1. The vector $\mathbf{Y}$ does not cause the vector $\mathbf{X}$ if

$$\sigma^2(X_{it} \mid \bar{A}_t) = \sigma^2(X_{it} \mid \bar{A}_t - \bar{\mathbf{Y}}_t), \qquad i = 1, \ldots, n. \tag{1.1}$$

According to this definition, the vector $\mathbf{Y}$ causes the vector $\mathbf{X}$ if and only if $\sigma^2(X_{it} \mid \bar{A}_t) < \sigma^2(X_{it} \mid \bar{A}_t - \bar{\mathbf{Y}}_t)$ for at least one value of i. On the other hand, Tjostheim (1981) proposed an apparently different definition of noncausality between the vectors, which is as follows.

Definition 2. The vector $\mathbf{Y}$ does not cause the vector $\mathbf{X}$ if

$$\Sigma(\mathbf{X}_t \mid \bar{A}_t) = \Sigma(\mathbf{X}_t \mid \bar{A}_t - \bar{\mathbf{Y}}_t). \tag{1.2}$$

In our context the two definitions are equivalent and the notion of noncausality can also be expressed directly in terms of projections. Indeed, by the uniqueness of orthogonal projections, (1.1) is equivalent to $P(X_{it} \mid \bar{A}_t) = P(X_{it} \mid \bar{A}_t - \bar{\mathbf{Y}}_t)$, $i = 1, \ldots, n$ or

$$P(\mathbf{X}_t \mid \bar{A}_t) = P(\mathbf{X}_t \mid \bar{A}_t - \bar{\mathbf{Y}}_t), \tag{1.3}$$

where the equality holds in the L_2 sense (hence with probability 1). Obviously, (1.3) implies (1.2). Conversely, noncausality in Tjostheim's sense implies noncausality in our sense, and (1.3) follows. Thus the formulations (1.1), (1.2), and (1.3) are equivalent.

2. CAUSALITY IN INVERTIBLE LINEAR PROCESSES

Let $\{\mathbf{X}_t : t \in Z\}$ be a second-order stationary purely indeterministic n-dimensional vector process. Without loss of generality, we can suppose that $E[\mathbf{X}_t] \equiv \mathbf{0}$. By Wold's decomposition theorem, the process $\{\mathbf{X}_t\}$ admits the following representation:

$$\mathbf{X}_t = \sum_{j=0}^{\infty} \mathbf{\Psi}_j \mathbf{a}_{t-j}, \qquad t \in Z, \tag{2.1}$$

where the $\mathbf{\Psi}_j$'s are $n \times n$ matrices such that $\sum_{j=1}^{\infty} |\mathbf{\Psi}_j|^2 < \infty$, $|\cdot|$ is the Euclidean norm matrix, $\mathbf{\Psi}_0 = \mathbf{I}$ is the identity matrix of order n, and $\{\mathbf{a}_t : t \in Z\}$ is the innovation process. The $\mathbf{a}_t$'s are uncorrelated random vectors with mean $\mathbf{0}$ and nonsingular covariance matrix $\mathbf{V}$. We suppose that the process (2.1) is invertible; that is, $\mathbf{X}_t$ admits a possibly infinite autoregressive representation

$$\mathbf{X}_t = \sum_{j=1}^{\infty} \mathbf{\Pi}_j \mathbf{X}_{t-j} + \mathbf{a}_t, \qquad t \in Z, \tag{2.2}$$

where the $\mathbf{\Pi}_j$'s are $n \times n$ matrices such that the series in (2.2) converges in quadratic mean. A characterization of the invertibility of a multivariate linear process in terms of its spectral density (matrix) function was given in Nsiri and Roy (1992). Stationary invertible multivariate ARMA processes are special cases of (2.2). Using the notation of the backward shift operator B, (2.2) is equivalent to

$$\mathbf{\Pi}(B)\mathbf{X}_t = \mathbf{a}_t, \qquad t \in Z, \tag{2.3}$$

where $\mathbf{\Pi}(B) = -\sum_{j=0}^{\infty} \mathbf{\Pi}_j B^j = (\Pi_{ij}(B))_{n \times n}$ is an $n \times n$ matrix of power series in the operator B and $\mathbf{\Pi}_0 = -\mathbf{I}_{n \times n}$.

Let us partition $\mathbf{X}_t$ and $\mathbf{a}_t$ into three subvectors as follows: $\mathbf{X}_t = (\mathbf{X}'_{1t}, \mathbf{X}'_{2t}, \mathbf{X}'_{3t})'$, $\mathbf{a}_t = (\mathbf{a}'_{1t}, \mathbf{a}'_{2t}, \mathbf{a}'_{3t})'$, where $\mathbf{X}_{it}$ and $\mathbf{a}_{it}$ are $n_i \times 1$ vectors; $i = 1, 2, 3$, with $n_1 \geq 1$, $n_2 \geq 1$, and $n_3 \geq 0$; and $n_1 + n_2 + n_3 = n$. When $n_3 = 0$, $\mathbf{X}_t$ is partitioned into two subvectors only. Using the corresponding partition of $\mathbf{\Pi}(B)$, (2.3) can be written as

$$\begin{pmatrix} \mathbf{\Pi}_{11}(B) & \mathbf{\Pi}_{12}(B) & \mathbf{\Pi}_{13}(B) \\ \mathbf{\Pi}_{21}(B) & \mathbf{\Pi}_{22}(B) & \mathbf{\Pi}_{23}(B) \\ \mathbf{\Pi}_{31}(B) & \mathbf{\Pi}_{32}(B) & \mathbf{\Pi}_{33}(B) \end{pmatrix} \begin{pmatrix} \mathbf{X}_{1t} \\ \mathbf{X}_{2t} \\ \mathbf{X}_{3t} \end{pmatrix} = \begin{pmatrix} \mathbf{a}_{1t} \\ \mathbf{a}_{2t} \\ \mathbf{a}_{3t} \end{pmatrix}. \quad (2.4)$$

For any operator $\mathbf{\Pi}(B)$, the corresponding function of the complex variable z is denoted by $\mathbf{\Pi}(z)$.

The following result is important for the sequel. It was stated by Kang (1981) for a bivariate process, but a proof for the general case does not seem to be available in the literature. The proofs of the proposition and theorems are given in the Appendix.

Proposition 1. In the linear invertible process (2.3) partitioned as in (2.4), $\mathbf{X}_1$ does not cause $\mathbf{X}_2$ if and only if $\mathbf{\Pi}_{21}(z) \equiv \mathbf{0}$.

By taking $n_1 = n_2 = 1$, we obtain a characterization of noncausality between any two components of $\mathbf{X}$.

Corollary 1. In the linear invertible process (2.3), for any i and j, the variable X_i does not cause the variable X_j if and only if $\Pi_{ji}(z) \equiv 0$.

From Proposition 1 and Corollary 1, we deduce the following useful characterization.

Corollary 2. In the linear invertible process (2.3) partitioned as in (2.4), $\mathbf{X}_1$ does not cause $\mathbf{X}_2$ if and only if X_i does not cause X_j, $i = 1, \ldots, n_1$, $j = n_1 + 1, \ldots, n_1 + n_2$.

3. CAUSALITY IN ARMA PROCESSES

Let the n-dimensional ARMA (p, q) process

$$\mathbf{\Phi}(B)\mathbf{X}_t = \mathbf{\Theta}(B)\mathbf{a}_t \quad (3.1)$$

be stationary and invertible, where $\mathbf{\Phi}(B) = \mathbf{I} - \mathbf{\Phi}_1 B - \cdots - \mathbf{\Phi}_p B^p$ and $\mathbf{\Theta}(B) = \mathbf{I} - \mathbf{\Theta}_1 B - \cdots - \mathbf{\Theta}_q B^q$. We also assume that the parameters in $\mathbf{\Phi}(B)$ and $\mathbf{\Theta}(B)$ are identified (uniquely defined) as functions of the autocovariance matrices of $\mathbf{X}_t$, so that $\mathbf{\Phi}(B)$ and $\mathbf{\Theta}(B)$ have no common factor. Using Corollary 1, we derive the following characterization of noncausality between any two components of $\mathbf{X}$. In the sequel, det $\mathbf{A}$ will denote the determinant of the matrix $\mathbf{A}$.

Theorem 1. In the stationary and invertible ARMA process (3.1), X_i does not cause X_j if and only if

$$\det(\mathbf{\Phi}_i(z), \mathbf{\Theta}_{(j)}(z)) \equiv 0, \quad (3.2)$$

where $\mathbf{\Phi}_i(z)$ is the ith column of the matrix $\mathbf{\Phi}(z)$ and $\mathbf{\Theta}_{(j)}(z)$ is the matrix $\mathbf{\Theta}(z)$ without its jth column.

If $\mathbf{X} = (\mathbf{X}'_1, \mathbf{X}'_2, \mathbf{X}'_3)'$, as in Section 2, then the following result follows immediately from Theorem 1 and Corollary 2.

Corollary 3. In the stationary and invertible ARMA process (3.1), $\mathbf{X}_1$ does not cause $\mathbf{X}_2$ if and only if (3.2) is satisfied for $i = 1, \ldots, n_1$, and $j = n_1 + 1, \ldots, n_1 + n_2$.

From now on we will suppose that $\mathbf{X}$ is partitioned into two subvectors: $\mathbf{X} = (\mathbf{X}'_1, \mathbf{X}'_2)'$, where $\mathbf{X}_i$ is $n_i \times 1$, $i = 1, 2$, and $n_1 + n_2 = n$. In this case model (3.1) can be rewritten as

$$\begin{pmatrix} \mathbf{\Phi}_{11}(B) & \mathbf{\Phi}_{12}(B) \\ \mathbf{\Phi}_{21}(B) & \mathbf{\Phi}_{22}(B) \end{pmatrix} \begin{pmatrix} \mathbf{X}_{1t} \\ \mathbf{X}_{2t} \end{pmatrix} = \begin{pmatrix} \mathbf{\Theta}_{11}(B) & \mathbf{\Theta}_{12}(B) \\ \mathbf{\Theta}_{21}(B) & \mathbf{\Theta}_{22}(B) \end{pmatrix} \begin{pmatrix} \mathbf{a}_{1t} \\ \mathbf{a}_{2t} \end{pmatrix} \quad (3.3)$$

where $\mathbf{\Phi}_{ij}(B)$ and $\mathbf{\Theta}_{ij}(B)$ are $n_i \times n_j$ matrices, $i, j = 1, 2$. Then the condition of noncausality between $\mathbf{X}_1$ and $\mathbf{X}_2$ can be formulated in the following way.

Theorem 2. Suppose that the stationary ARMA process (3.3) is invertible, with $\det[\mathbf{\Theta}_{11}(z)] \neq 0$ for all $z \in C$ such that $|z| \leq 1$. Then, $\mathbf{X}_1$ does not cause $\mathbf{X}_2$ if and only if

$$\mathbf{\Phi}_{21}(z) - \mathbf{\Theta}_{21}(z)\mathbf{\Theta}_{11}(z)^{-1}\mathbf{\Phi}_{11}(z) \equiv \mathbf{0}. \quad (3.4)$$

If $n_1 = n_2 = 1$, then X_1 does not cause X_2 if and only if $\Theta_{11}(z)\Phi_{21}(z) - \Theta_{21}(z)\Phi_{11}(z) \equiv 0$, and we retrieve the condition of Kang (1981). For a pure MA process, $\mathbf{\Phi}_{21}(z) \equiv \mathbf{0}$, $\mathbf{\Phi}_{11}(z) \equiv \mathbf{I}$, and (3.4) reduces to $\mathbf{\Theta}_{21}(z) \equiv \mathbf{0}$.

The following characterization of the noncausality between $\mathbf{X}_1$ and $\mathbf{X}_2$ is more convenient to deal with applications. Write $\mathbf{\Phi}(z) = (\Phi_{ij}(z))_{n\times n}$ and $\mathbf{\Theta}(z) = (\Theta_{ij}(z))_{n\times n}$. To simplify the notation, we omit the argument z in $\Phi_{ij}(z)$ and $\Theta_{ij}(z)$.

Theorem 3. Suppose that the stationary ARMA process (3.3) is invertible, with $\det[\mathbf{\Theta}_{11}(z)] \neq 0$ for all $z \in C$ such that $|z| \leq 1$. Then, $\mathbf{X}_1$ does not cause $\mathbf{X}_2$ if and only if

$$\begin{vmatrix} \Phi_{1j} & \Theta_{11} & \Theta_{12} & \cdots & \Theta_{1r} \\ \Phi_{2j} & \Theta_{21} & \Theta_{22} & \cdots & \Theta_{2r} \\ \vdots & \vdots & \vdots & & \vdots \\ \Phi_{rj} & \Theta_{r1} & \Theta_{r2} & \cdots & \Theta_{rr} \\ \Phi_{r+i,j} & \Theta_{r+i,1} & \Theta_{r+i,2} & \cdots & \Theta_{r+i,r} \end{vmatrix} \equiv 0 \quad (3.5)$$

for $i = 1, \ldots, s$, and $j = 1, \ldots, r$, where $r = n_1$ and $s = n_2$.

Example 1. Let $\mathbf{X}_t$ be a three-dimensional ARMA process. From Corollary 3 it follows that X_1 does not cause $(X_2, X_3)'$ if and only if

$$\begin{vmatrix} \Phi_{11} & \Theta_{11} & \Theta_{13} \\ \Phi_{21} & \Theta_{21} & \Theta_{23} \\ \Phi_{31} & \Theta_{31} & \Theta_{33} \end{vmatrix} \equiv 0$$

and

$$\begin{vmatrix} \Phi_{11} & \Theta_{11} & \Theta_{13} \\ \Phi_{21} & \Theta_{21} & \Theta_{22} \\ \Phi_{31} & \Theta_{31} & \Theta_{32} \end{vmatrix} \equiv 0.$$

On the other hand, by Theorem 3, X_1 does not cause $(X_2, X_3)'$ if and only if

$$\begin{vmatrix} \Phi_{11} & \Theta_{11} \\ \Phi_{1+i,1} & \Theta_{1+i,1} \end{vmatrix} \equiv 0, \qquad i = 1, 2.$$

Thus Theorem 1 and Corollary 3 lead to the evaluation of 3×3 determinants, whereas Theorem 3 leads to the evaluation of 2×2 determinants.

More generally, the determinants involved in Corollary 3 are of dimension n, where n is the number of variables considered in the analysis, and the determinants involved in Theorem 3 are of dimension $r + 1$, where r is the number of variables in the subvector $\mathbf{X}_1$. In many situations, application of Theorem 3 can simplify the computations considerably.

4. TESTING CAUSALITY

4.1 General Test Procedure

Given a series of N observations of the vector $\mathbf{X} = (\mathbf{X}_1', \mathbf{X}_2')'$, we want to test the hypothesis of noncausality between $\mathbf{X}_1$ and $\mathbf{X}_2$. We propose a three-stage procedure:

1. Build a multivariate ARMA model for the series following the procedure of Tiao and Box (1981).
2. Using the results of Section 3, derive the noncausality conditions and express them in terms of the AR and MA parameters of the estimated model. Denoting β as the vector of all AR and MA parameters, the noncausality conditions lead to (possibly nonlinear) constraints on an $l \times 1$ subvector β_1 of β. We will denote these restrictions by

$$R_j(\beta_1) = 0, \qquad j = 1, \ldots, K, \tag{4.1}$$

where $K \le l$.

3. Choose a test criterion. We will consider Wald and likelihood ratio (LR) tests.

The Wald test is easier to apply, because it uses only the unconstrained maximum likelihood estimators (MLE's) of the parameters of the full model. It does not require estimation of the constrained model. Let $\mathbf{T}(\hat{\beta}_1)$ be the $l \times K$ matrix of derivatives

$$\mathbf{T}(\hat{\beta}_1) = \left(\frac{\partial R_j(\beta_1)}{\partial \beta_i} \bigg|_{\beta_1 = \hat{\beta}_1} \right)_{l \times K},$$

and let $\mathbf{V}(\beta_1)$ be the asymptotic covariance matrix of $\sqrt{N}(\hat{\beta}_1 - \beta_1)$. The Wald statistic is given by

$$\xi_W = N\mathbf{R}(\hat{\beta}_1)'[\mathbf{T}(\hat{\beta}_1)'\mathbf{V}(\hat{\beta}_1)\mathbf{T}(\hat{\beta}_1)]^{-1}\mathbf{R}(\hat{\beta}_1), \tag{4.2}$$

where $\mathbf{R}(\beta_1) = (R_1(\beta_1), \ldots, R_K(\beta_1))'$. In this definition $\mathbf{T}(\beta_1)$ must be of full rank. Because this is not always the case, we propose in Section 4.2, a sequential procedure for such situations.

Let $L(\beta, \mathbf{X})$ be the logarithm of the likelihood function of the N observations. The likelihood ratio statistic is given by

$$\xi_{LR} = 2(L(\hat{\beta}, \mathbf{X}) - L(\hat{\beta}^*, \mathbf{X})), \tag{4.3}$$

where $\hat{\beta}^*$ is the MLE of β under the constraints (4.1) and $\hat{\beta}$ is the unconstrained estimator.

Under the null hypothesis of noncausality, it is well known that ξ_W and ξ_{LR} are asymptotically equivalent and follow χ^2_K distributions; see Basawa and Koul (1979) and Basawa, Billard, and Srinivasan (1984). At the significance level α, we reject the hypothesis of noncausality if $\xi > \chi^2_{K;1-\alpha}$, where $\chi^2_{K;1-\alpha}$ is the $(1-\alpha)$-quantile of the chi-squared distribution with K degrees of freedom and ξ represents ξ_W or ξ_{LR}.

4.2 A Sequential Bounds Procedure for the Singular Case

We now describe a sequential approach to deal with singular cases. Consider the bivariate stationary and invertible ARMA (1, 1) model:

$$\begin{pmatrix} 1 - \phi_{11}B & -\phi_{12}B \\ -\phi_{21}B & 1 - \phi_{22}B \end{pmatrix} \begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix} = \begin{pmatrix} 1 - \theta_{11}B & -\theta_{12}B \\ -\theta_{21}B & 1 - \theta_{22}B \end{pmatrix} \begin{pmatrix} a_{1t} \\ a_{2t} \end{pmatrix}.$$

From (3.4), X_1 does not cause X_2

$$\begin{aligned} &\Leftrightarrow \Theta_{11}(z)\Phi_{21}(z) - \Theta_{21}(z)\Phi_{11}(z) \equiv 0, \\ &\Leftrightarrow (\phi_{21} - \theta_{21})z + (\theta_{11}\theta_{21} - \phi_{21}\theta_{11})z^2 \equiv 0, \\ &\Leftrightarrow \phi_{21} - \theta_{21} = 0 \quad \text{and} \quad \phi_{11}\theta_{21} - \phi_{21}\theta_{11} = 0. \end{aligned} \tag{4.4}$$

For the vector $\beta_1 = (\phi_{11}, \phi_{21}, \theta_{11}, \theta_{21})'$, the matrix

$$T(\beta_1) = \begin{pmatrix} 0 & \theta_{21} \\ 1 & -\theta_{11} \\ 0 & -\phi_{21} \\ -1 & \phi_{11} \end{pmatrix}$$

is not necessarily a full column-rank matrix under the null hypothesis $H_0: X_1$ does not cause X_2. To avoid this problem, rewrite the noncausality constraints (4.4) as follows:

$$\phi_{21} - \theta_{21} = 0$$

and

$$\phi_{21} = \theta_{21} = 0 \quad \text{or} \quad [\phi_{21} \ne 0 \text{ and } \phi_{11} - \theta_{11} = 0]. \tag{4.5}$$

Thus $\phi_{21} - \theta_{21} = 0$ is a necessary condition for H_0, $\phi_{21} = \theta_{21} = 0$ is a sufficient condition for H_0, and $\phi_{21} - \theta_{21} = 0$ and $\phi_{11} = \theta_{11}$ (taken jointly) are sufficient conditions for H_0. Consider the hypotheses: H_0^1: $\phi_{21} - \theta_{21} = 0$; H_0^2: $\phi_{21} = \theta_{21} = 0$; H_0^3: $\phi_{21} \ne 0$, $\phi_{21} - \theta_{21} = 0$, and $\phi_{11} - \theta_{11} = 0$; and $\tilde{H}_0^3$: $\phi_{11} - \theta_{11} = 0$. We have the following relations: $H_0^3 = \tilde{H}_0^3 \cap H_0^1$, $H_0^2 \subseteq H_0 \subseteq H_0^1$, $H_0^3 \subseteq H_0 \subseteq H_0^1$. Suppose now that we can test the three hypotheses H_0^1, H_0^2, and $\tilde{H}_0^3$ *separately*. For $\tilde{H}_0^3$, it will be sufficient to test $\phi_{11} - \theta_{11} = 0$ under the assumption $\phi_{21} \ne 0$. Then, for given significance level $\alpha (0 < \alpha < 1)$, we can test H_0 in the following way. Let $\alpha = \alpha_1 + \alpha_2$ and $0 < \alpha_i < 1$, $i = 1, 2$.

1. We first test H_0^1 at level α_1. If H_0^1 is rejected, then H_0 is rejected too and the procedure stops.
2. If H_0^1 is not rejected, then we test H_0^2 at level α_2. If H_0^2 is not rejected, then we cannot reject H_0 and we stop.
3. If we reject H_0^2, then we test $\tilde{H}_0^3$: $\phi_{11} - \theta_{11} = 0$ at level α_2. If $\tilde{H}_0^3$ is rejected, then H_0 is also rejected. If $\tilde{H}_0^3$ is not rejected, then H_0 is also not rejected.

If H_0 is rejected by this procedure, then we write $\psi(\alpha_1, \alpha_2) = 1$; if it is not rejected, then we write $\psi(\alpha_1, \alpha_2) = 0$. The procedure is summarized in Figure 1.

The procedure just described is conservative: Under H_0, we have $P[\psi(\alpha_1, \alpha_2) = 1] \le P[A] + P[B]$, where A represents the event "reject H_0^1" and B represents the event "do not reject H_0^1 and reject H_0^2 and $\tilde{H}_0^3$." For the event B there are two possible cases (under H_0):

1. $\phi_{21} = \theta_{21} = 0$, in which case we have $P[B] \le P[\text{reject } H_0^2] = \alpha_2$.

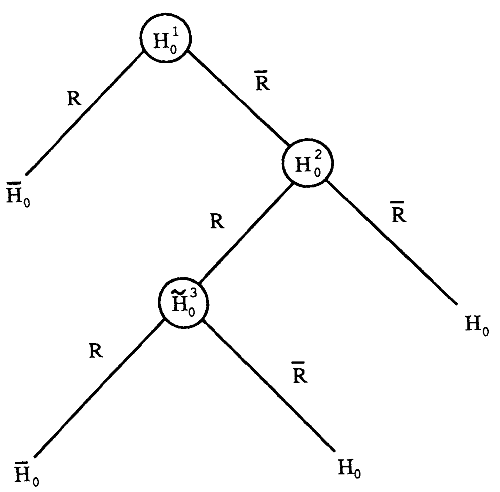

Figure 1. Sequential Procedure for Testing H_0. The symbols R and $\bar{R}$ mean "reject" and "do not reject," and $\bar{H}_0$ stands for "not H_0."

2. $\phi_{21} = \theta_{21} \neq 0$ and $\phi_{11} = \theta_{11}$, in which case $P[B] \leq P[\text{reject } \tilde{H}_0^3] = \alpha_2$.

Therefore, we have

$$P[\psi(\alpha_1, \alpha_2) = 1] \leq \alpha_1 + \alpha_2 \leq \alpha. \tag{4.5}$$

Because the critical region $\psi(\alpha_1, \alpha_2) = 1$ is conservative, it is better to view the test as inconclusive when $\psi(\alpha_1, \alpha_2) = 0$. However, it is possible to reduce the probability of an inconclusive test by observing that

$$P[\psi(\alpha_1, \alpha_2) = 0] \leq 1 - P[\text{reject } H_0^1] = 1 - \alpha_1. \tag{4.6}$$

Because $\psi(\alpha, \alpha_2) = 0$ implies $\psi(\alpha_1, \alpha_2) = 0$ (because $\alpha \geq \alpha_1$), this suggests the following bounds procedure:

1. Reject H_0 when $\psi(\alpha_1, \alpha_2) = 1$.
2. Accept H_0 when $\psi(\alpha, \alpha_2) = 0$.
3. Consider the test inconclusive otherwise. (4.7)

Clearly, under H_0, $P[\text{reject } H_0] \leq \alpha$ and $P[\text{accept } H_0] \leq 1 - \alpha$. For further discussion of bounds procedures, see Dufour (1989, 1990). It is easy to see that the approach just described can be adapted to other situations. Of course, the occurrence of such singularities depends on the structure of the ARMA model studied, and so a bounds procedure may not be needed.

5. CAUSALITY TESTS BETWEEN MONEY AND INCOME

Causality relations between money and income have been much debated in the economic literature; see Feige and Pearce (1979), Hsiao (1979), Osborn (1984) and Sims (1972, 1980b). To illustrate the causality conditions and tests given previously, we will now study causality between money and income in Canada. The data used are those of Hsiao (1979) and Osborn (1984). They consist of quarterly seasonally adjusted nominal GNP, M1, and M2 over the period 1955–1977 (92 observations). A listing of the data is available in the Appendix of Hsiao (1979).

The natural logarithm of all three variables was taken before modeling. Unit root tests, following the methods of Dickey and Fuller (1979, 1981) and Phillips and Perron (1988), led us to analyze the first difference of each series. In the following we will denote $y_t = (1 - B)\ln \text{GNP}_t$, $m_{1t} = (1 - B)\ln \text{M1}_t$, and $m_{2t} = (1 - B)\ln \text{M2}_t$. Using the approach of Tiao and Box (1981), we first performed bivariate analyses of $(y_t, m_{1t})'$, $(y_t, m_{2t})'$, and $(m_{1t}, m_{2t})'$ and then performed a trivariate analysis of $(y_t, m_{1t}, m_{2t})'$. The models obtained in this way appear in Figure 2. These models all satisfy the diagnostic checks suggested by Tiao and Box (1981) to ensure model adequacy. Further details on these analyses are available in Boudjellaba (1988).

Let us first consider the bivariate models. For $(y_t, m_{1t})'$, we see from Theorem 1 that

$$y_t \not\to m_{1t} \Leftrightarrow \begin{vmatrix} \Phi_{11}(z) & \Theta_{11}(z) \\ \Phi_{21}(z) & \Theta_{21}(z) \end{vmatrix} \equiv 0 \Leftrightarrow$$

$$\begin{vmatrix} 1 - \phi_{11}^{(1)} z & 1 - \theta_{11}^{(1)} z \\ -\phi_{21}^{(1)} z & 0 \end{vmatrix} \equiv 0 \Leftrightarrow \phi_{21}^{(1)} = 0 \quad \text{and}$$

$$\phi_{21}^{(1)} \theta_{11}^{(1)} = 0 \Leftrightarrow \phi_{21}^{(1)} = 0$$

Bivariate Models

$$\begin{bmatrix} 1 - .799B & -.203B \\ (.049) & (.040) \\ -.374B & 1 - .396B \\ (.101) & (.096) \end{bmatrix} \begin{bmatrix} y_t \\ m_{1t} \end{bmatrix} = \begin{bmatrix} .001 \\ (.001) \\ .002 \\ (.002) \end{bmatrix} + \begin{bmatrix} 1 - .861B & 0 \\ (.071) & \\ 0 & 1 - .508B^4 \\ & (.088) \end{bmatrix} \begin{bmatrix} a_{yt}^{(1)} \\ a_{1t}^{(1)} \end{bmatrix}$$

$$\begin{bmatrix} 1 - .876B & 0 \\ (.093) & \\ -.273B & 1 - .755B \\ (.068) & (.057) \end{bmatrix} \begin{bmatrix} y_t \\ m_{2t} \end{bmatrix} = \begin{bmatrix} .003 \\ (.002) \\ .000 \\ (.001) \end{bmatrix} + \begin{bmatrix} 1 - .657B & 0 \\ (.147) & \\ 0 & 1 - .717B^4 \\ & (.079) \end{bmatrix} \begin{bmatrix} a_{yt}^{(2)} \\ a_{2t}^{(2)} \end{bmatrix}$$

$$\begin{bmatrix} 1 & -.640B \\ & (.065) \\ -.208B^2 & 1 - .781B \\ (.064) & (.060) \end{bmatrix} \begin{bmatrix} m_{1t} \\ m_{2t} \end{bmatrix} = \begin{bmatrix} .002 \\ (.002) \\ .002 \\ (.001) \end{bmatrix} + \begin{bmatrix} 1 - .569B^4 & 0 \\ (.079) & \\ 0 & 1 - .616B^4 \\ & (.074) \end{bmatrix} \begin{bmatrix} a_{1t}^{(3)} \\ a_{2t}^{(3)} \end{bmatrix}$$

Trivariate model

$$\begin{bmatrix} 1 - .610B^2 & -.095B - .299B^2 & 0 \\ (.077) & (.068) \quad (.072) & \\ 0 & 1 & -.627B \\ & & (.065) \\ 0 & -.224B^2 & 1 - .784B \\ & (.067) & (.059) \end{bmatrix} \begin{bmatrix} y_t \\ m_{1t} \\ m_{2t} \end{bmatrix} = \begin{bmatrix} .002 \\ (.001) \\ .002 \\ (.002) \\ .002 \\ (.001) \end{bmatrix}$$

$$+ \begin{bmatrix} 1 - .668B^2 & 0 & 0 \\ (.105) & & \\ -.148B^5 & 1 - .578B^4 & 0 \\ (.102) & (.080) & \\ 0 & 0 & 1 - .072B^2 - .594B^4 \\ & & (.083) \quad (.078) \end{bmatrix} \begin{bmatrix} a_{yt} \\ a_{1t} \\ a_{2t} \end{bmatrix}$$

Error covariance matrices ($\times 10^4$ and in the same order as the models)

$$\begin{bmatrix} 1.28 & .29 \\ .29 & 2.17 \end{bmatrix}, \quad \begin{bmatrix} 1.64 & .24 \\ .24 & 1.01 \end{bmatrix}, \quad \begin{bmatrix} 1.98 & .77 \\ .77 & 1.06 \end{bmatrix}, \quad \begin{bmatrix} 1.22 & .33 & .13 \\ .33 & 1.87 & .70 \\ .13 & .70 & 1.03 \end{bmatrix}.$$

Figure 2. Multivariate ARMA Models. NOTE: Estimated standard errors are given in parentheses.

and

$$m_{1t} \nrightarrow y_t \Leftrightarrow \begin{vmatrix} \Phi_{12}(z) & \Theta_{12}(z) \\ \Phi_{22}(z) & \Theta_{22}(z) \end{vmatrix} \equiv 0 \Leftrightarrow$$

$$\begin{vmatrix} -\phi_{12}^{(1)}z & 0 \\ 1-\phi_{22}^{(1)}z & 1-\theta_{22}^{(4)}z^4 \end{vmatrix} \equiv 0 \Leftrightarrow \phi_{12}^{(1)} = 0,$$

where $\nrightarrow$ means "does not cause," $\Phi_{ij}(B)$ and $\Theta_{ij}(B)$ refer to the corresponding lag polynomials in the model for $(y_t, m_{1t})'$, and $\phi_{ij}^{(k)}$ and $\theta_{ij}^{(k)}$ are the coefficients of B^k in $\Phi_{ij}(B)$ and $\Theta_{ij}(B)$. The Wald and LR statistics for testing "y_t does not cause m_{1t}" ($y_t \nrightarrow m_{1t}$) take the values 13.7 and 36.9. (Causality tests are summarized in Table 1). To test $m_{1t} \nrightarrow y_t$, the corresponding statistics are 25.8 and 64.6. Because the asymptotic null distribution of these test statistics is χ_1^2, both null hypotheses are strongly rejected at conventional significance levels, and we conclude that there is feedback between y_t and m_{1t} (i.e., $m_{1t} \leftrightarrow y_t$). Hsiao (1979) and Osborn (1984) reached the same conclusion.

For $(y_t, m_{2t})'$, Theorem 1 implies

$$y_t \nrightarrow m_{2t} \Leftrightarrow \begin{vmatrix} \Phi_{11}(z) & \Theta_{11}(z) \\ \Phi_{21}(z) & \Theta_{21}(z) \end{vmatrix} \equiv 0 \Leftrightarrow$$

$$\begin{vmatrix} 1-\phi_{11}^{(1)}z & 1-\theta_{11}^{(1)}z \\ -\phi_{21}^{(1)}z & 0 \end{vmatrix} \equiv 0 \Leftrightarrow \phi_{21}^{(1)} = 0$$

and

$$m_{2t} \nrightarrow y_t \Leftrightarrow \begin{vmatrix} \Phi_{12}(z) & \Theta_{12}(z) \\ \Phi_{22}(z) & \Theta_{22}(z) \end{vmatrix} \equiv 0 \Leftrightarrow$$

$$\begin{vmatrix} 0 & 0 \\ 1-\phi_{22}^{(1)}z & 1-\theta_{22}^{(4)}z^4 \end{vmatrix} \equiv 0,$$

where $\Phi_{ij}(B)$ and $\Theta_{ij}(B)$ now refer to lag polynomials in the model for $(y_t, m_{2t})'$. We see that the condition $m_{2t} \nrightarrow y_t$ is satisfied exactly by the model, whereas $y_t \nrightarrow m_{2t}$ is strongly rejected by the Wald and LR tests (see Table 1). Thus y_t causes m_{2t} unidirectionally ($y_t \rightarrow m_{2t}$). This conclusion is also in agreement with Hsiao (1979) and Osborn (1984).

For (m_{1t}, m_{2t}), we find in a similar way that $m_{1t} \nrightarrow m_{2t} \Leftrightarrow \phi_{21}^{(1)} = 0$ and $m_{2t} \nrightarrow m_{1t} \Leftrightarrow \phi_{12}^{(1)} = 0$. Both of these hypotheses are strongly rejected (especially the second one), so that we conclude that there is feedback between m_{1t} and m_{2t} ($m_{1t} \leftrightarrow m_{2t}$).

From the bivariate models we thus find $m_{1t} \leftrightarrow y_t$, $y_t \rightarrow m_{2t}$ and $m_{1t} \leftrightarrow m_{2t}$. The most striking results here is that money stock changes (m_{1t} and m_{2t}) do not cause unidirectionally nominal income changes (y_t), whereas y_t causes m_{2t} unidirectionally.

A bivariate causality analysis between y_t and m_{1t} or m_{2t} is unsatisfactory, however. Given the use of two money stock series, the hypothesis of interest is: Do money stock changes (m_{1t} and m_{2t}) cause nominal income changes (y_t)? To answer this question, a multivariate ARMA model incorporating at least three variables is required. Again using the approach of Tiao and Box (1981), we found the trivariate model given in Figure 2.

The most interesting relationship here is the one between y_t and the vector $(m_{1t}, m_{2t})'$. From Theorem 3, it follows that

$$y_t \nrightarrow (m_{1t}, m_{2t})' \Leftrightarrow \begin{vmatrix} \Phi_{11}(z) & \Theta_{11}(z) \\ \Phi_{21}(z) & \Theta_{21}(z) \end{vmatrix} \equiv 0 \quad \text{and}$$

$$\begin{vmatrix} \Phi_{11}(z) & \Theta_{11}(z) \\ \Phi_{31}(z) & \Theta_{31}(z) \end{vmatrix} \equiv 0 \quad \Leftrightarrow \begin{vmatrix} 1-\phi_{11}^{(2)}z^2 & 1-\theta_{11}^{(2)}z^2 \\ 0 & -\theta_{21}^{(5)}z^5 \end{vmatrix} \equiv 0$$

$$\text{and} \quad \begin{vmatrix} 1-\phi_{11}^{(2)}z^2 & 1-\theta_{11}^{(2)}z^2 \\ 0 & 0 \end{vmatrix} \equiv 0 \quad \Leftrightarrow \theta_{21}^{(5)} = 0.$$

Further, by Corollary 2, $(m_{1t}, m_{2t})' \nrightarrow y_t \Leftrightarrow m_{1t} \nrightarrow y_t$ and

Table 1. Causality Tests

	Null hypothesis (noncausality)	Parametric representation	Wald statistic	Likelihood ratio statistic	Degrees of freedom
Bivariate models	$y_t \nrightarrow m_{1t}$	$\phi_{21}^{(1)} = 0$	13.7[a]	36.8[a]	1
	$m_{1t} \nrightarrow y_t$	$\phi_{12}^{(1)} = 0$	25.8[a]	64.6[a]	1
	$y_t \nrightarrow m_{2t}$	$\phi_{21}^{(1)} = 0$	16.1[a]	23.3[a]	1
	$m_{2t} \nrightarrow y_t$[b]	—	—	—	—
	$m_{1t} \nrightarrow m_{2t}$	$\phi_{21}^{(2)} = 0$	10.5[a]	20.4[a]	1
	$m_{2t} \nrightarrow m_{1t}$	$\phi_{12}^{(1)} = 0$	96.9[a]	45.6[a]	1
Trivariate models	$y_t \nrightarrow (m_{1t}, m_{2t})'$	$\theta_{21}^{(5)} = 0$	2.1	3.5	1
	$(m_{1t}, m_{2t})' \nrightarrow y_t$	$\phi_{12}^{(1)} = \phi_{12}^{(2)} = 0$	69.3[a]	62.7[a]	2
	$y_t \nrightarrow m_{1t}$[c]	$\theta_{21}^{(5)} = 0$	2.1	3.5	1
	$m_{1t} \nrightarrow y_t$[c]	$\phi_{12}^{(1)} = \phi_{12}^{(2)} = 0$	69.3[a]	62.7[a]	2
	$y_t \nrightarrow m_{2t}$[b]	—	—	—	—
	$m_{2t} \nrightarrow y_t$[b]	—	—	—	—
	$m_{1t} \nrightarrow m_{2t}$	$\phi_{32}^{(2)} = 0$	11.2[a]	17.2[a]	1
	$m_{2t} \nrightarrow m_{1t}$	$\phi_{23}^{(1)} = 0$	93.0[a]	46.2[a]	1

[a] Significant at level .05.
[b] The condition of noncausality is satisfied exactly by the model.
[c] Because the structure of the model implies that $y_t \nrightarrow m_{2t}$ and $m_{2t} \nrightarrow y_t$, the statistics for testing noncausality between m_{1t} and y_t are identical to those between (m_{1t}, m_{2t}) and y_t.

$m_{2t} \nrightarrow y_t$ and, by Theorem 1,

$$m_{1t} \nrightarrow y_t \Leftrightarrow \begin{vmatrix} \Phi_{12}(z) & \Theta_{12}(z) & \Theta_{13}(z) \\ \Phi_{22}(z) & \Theta_{22}(z) & \Theta_{23}(z) \\ \Phi_{32}(z) & \Theta_{32}(z) & \Theta_{33}(z) \end{vmatrix} \equiv 0 \Leftrightarrow$$

$$\begin{vmatrix} -\phi_{12}^{(1)}z - \phi_{12}^{(2)}z^2 & 0 & 0 \\ 1 & 1 - \theta_{22}^{(4)}z^4 & 0 \\ -\phi_{32}^{(2)}z^2 & 0 & 1 - \theta_{33}^{(2)}z^2 - \theta_{33}^{(4)}z^4 \end{vmatrix} \equiv 0 \Leftrightarrow$$

$$(\phi_{12}^{(1)}z + \phi_{12}^{(2)}z^2)(1 - \theta_{22}^{(4)}z^4) \times (1 - \theta_{33}^{(2)}z^2 - \theta_{33}^{(4)}z^4) \equiv 0 \Leftrightarrow \phi_{12}^{(1)} = \phi_{12}^{(2)} = 0$$

and

$$m_{2t} \nrightarrow y_t \Leftrightarrow \begin{vmatrix} \Phi_{13}(z) & \Theta_{12}(z) & \Theta_{13}(z) \\ \Phi_{23}(z) & \Theta_{22}(z) & \Theta_{23}(z) \\ \Phi_{33}(z) & \Theta_{32}(z) & \Theta_{33}(z) \end{vmatrix} \equiv 0 \Leftrightarrow$$

$$\begin{vmatrix} 0 & 0 & 0 \\ \Phi_{23}(z) & \Theta_{22}(z) & \Theta_{23}(z) \\ \Phi_{33}(z) & \Theta_{32}(z) & \Theta_{33}(z) \end{vmatrix} \equiv 0.$$

We see that the condition for $m_{2t} \nrightarrow y_t$ holds exactly in this model, so that $(m_{1t}, m_{2t})' \nrightarrow y_t \Leftrightarrow m_{1t} \nrightarrow y_t \Leftrightarrow \phi_{12}^{(1)} = \phi_{12}^{(2)} = 0$. Using these conditions, we see in Table 1 that $(m_{1t}, m_{2t})' \nrightarrow y_t$ is strongly rejected, but $y_t \rightarrow (m_{1t}, m_{2t})'$ is accepted (at level .05). Causality appears to be unidirectional from $(m_{1t}, m_{2t})'$ to y_t. This result agrees with the one obtained by Osborn (1984) using a different methodology.

By Theorem 1 we also have $m_{1t} \nrightarrow m_{2t} \Leftrightarrow \phi_{32}^{(2)} = 0$, $m_{2t} \nrightarrow m_{1t} \Leftrightarrow \phi_{23}^{(1)} = 0$, $y_t \nrightarrow m_{1t} \Leftrightarrow \theta_{21}^{(5)} = 0$, and $m_{1t} \nrightarrow y_t \Leftrightarrow \phi_{12}^{(1)} = \phi_{12}^{(2)} = 0$, and the conditions for $m_{2t} \nrightarrow y_t$ and $y_t \nrightarrow m_{2t}$ are satisfied exactly by the model. We thus find that $m_{1t} \rightarrow y_t$, whereas y_t and m_{2t} do not cause each other. Further, the hypotheses $\phi_{32}^{(2)} = 0$ and $\phi_{23}^{(1)} = 0$ are strongly rejected, suggesting feedback between m_{1t} and m_{2t}.

Thus the causality structure that emerges from the trivariate model of $(y_t, m_{1t}, m_{2t})'$ is $m_{2t} \leftrightarrow m_{1t} \rightarrow y_t$. There is no direct link between m_{2t} and y_t and no causality running from m_{1t} and m_{2t} towards y_t. Due to the presence of feedback between m_{1t} and m_{2t}, the causality relationships suggested by the bivariate models now appear to be spurious. These results are, of course, quite consistent with a monetarist interpretation of the relation between money and nominal income.

6. CONCLUDING REMARKS

As illustrated by the previous example, the conclusions of a causality analysis obtained with bivariate models do not necessarily coincide with the ones obtained from a multivariate model (dimension $n > 2$). Therefore, it appears important when analyzing the relationships between two variables or two set of variables to work with a model that embody all the variables in the study. Multivariate ARMA models provide a natural and parsimonious framework for such an analysis. Further, the necessary and sufficient conditions established in this article allow one to test hypotheses of noncausality by considering directly a multivariate ARMA model.

APPENDIX: PROOFS

Proposition 1. From (2.4) it follows that $\sum_{i=1}^{3} \Pi_{2i}(B)\mathbf{X}_{it} = \mathbf{a}_{2t}$. Writing $\Pi_{2i}(B) = \delta_{2i}\mathbf{I} - \sum_{k=1}^{\infty} \Pi_{2i(k)}B^k$, $i = 1, 2, 3$, where δ_{ij} denotes the Kronecker delta and $\mathbf{I}$ stands for the identity matrix, we have

$$\mathbf{X}_{2t} = \sum_{i=1}^{3} \mathbf{Z}_{it} + \mathbf{a}_{2t}, \tag{A.1}$$

where

$$\mathbf{Z}_{it} = \sum_{k=1}^{\infty} \Pi_{2i(k)}\mathbf{X}_{i,t-k}, \qquad i = 1, 2, 3. \tag{A.2}$$

From (A.1) we have

$$P(\mathbf{X}_{2t} | \bar{\mathbf{X}}_{1t}, \bar{\mathbf{X}}_{2t}, \bar{\mathbf{X}}_{3t}) = \sum_{i=1}^{3} \mathbf{Z}_{it}, \tag{A.3}$$

because $\mathbf{a}_t$ is orthogonal to $\bar{\mathbf{X}}_t$. Further,

$$P(\mathbf{X}_{2t} | \bar{\mathbf{X}}_{2t}, \bar{\mathbf{X}}_{3t}) = P(\mathbf{Z}_{1t} | \bar{\mathbf{X}}_{2t}, \bar{\mathbf{X}}_{3t}) + \mathbf{Z}_{2t} + \mathbf{Z}_{3t}. \tag{A.4}$$

If $\Pi_{21}(z) \equiv \mathbf{0}$, then we have $\mathbf{Z}_{1t} = \mathbf{0}$, so that $P(\mathbf{Z}_{1t} | \bar{\mathbf{X}}_{2t}, \bar{\mathbf{X}}_{3t}) = \mathbf{0}$. From (A.3) and (A.4), it follows that $P(\mathbf{X}_{2t} | \bar{\mathbf{X}}_{1t}, \bar{\mathbf{X}}_{2t}, \bar{\mathbf{X}}_{3t}) = P(\mathbf{X}_{2t} | \bar{\mathbf{X}}_{2t}, \bar{\mathbf{X}}_{3t}) = \mathbf{Z}_{2t} + \mathbf{Z}_{3t}$. By the equivalence between (1.1) and (1.3), the "if" part of the proof is completed.

Conversely, if $\mathbf{X}_1$ does not cause $\mathbf{X}_2$, then $P(\mathbf{X}_{2t} | \bar{\mathbf{X}}_{1t}, \bar{\mathbf{X}}_{2t}, \bar{\mathbf{X}}_{3t}) = P(\mathbf{X}_{2t} | \bar{\mathbf{X}}_{2t}, \bar{\mathbf{X}}_{3t})$. From (A.3) and (A.4), it follows that $\mathbf{Z}_{1t} = P(\mathbf{Z}_{1t} | \bar{\mathbf{X}}_{2t}, \bar{\mathbf{X}}_{3t})$; that is, the components of $\mathbf{Z}_{1t}$ are contained in the closed span of $\bar{\mathbf{X}}_{2t} \cup \bar{\mathbf{X}}_{3t}$. Therefore, we can find sequences of matrices

$$\{\Pi_{22(k)}^{(T)} : k = 1, \ldots, T\}_{T=1}^{\infty}, \qquad \{\Pi_{23(k)}^{(T)} : k = 1, \ldots, T\}_{T=1}^{\infty}$$

such that

$$\sum_{k=1}^{T} \Pi_{22(k)}^{(T)}\mathbf{X}_{2,t-k} + \sum_{k=1}^{T} \Pi_{23(k)}^{(T)}\mathbf{X}_{3,t-k} \xrightarrow{\text{q.m.}} -\mathbf{Z}_{1t}, \tag{A.5}$$

where $\xrightarrow{\text{q.m.}}$ refers to convergence in quadratic mean (as $T \rightarrow \infty$). From (A.2) for $i = 1$ and (A.5), we get

$$\sum_{k=1}^{T} \{\Pi_{21(k)}\mathbf{X}_{1,t-k} + \Pi_{22(k)}^{(T)}\mathbf{X}_{2,t-k} + \Pi_{23(k)}^{(T)}\mathbf{X}_{3,t-k}\} \xrightarrow{\text{q.m.}} \mathbf{0}$$

or, equivalently,

$$\sum_{k=1}^{T} \mathbf{D}_k^{(T)}\mathbf{X}_{t-k} \xrightarrow{\text{q.m.}} \mathbf{0}, \tag{A.6}$$

where $\mathbf{D}_k^{(T)}$ is the $n_2 \times n$ matrix defined by

$$\mathbf{D}_k^{(T)} = (\Pi_{21(k)}, \Pi_{22(k)}^{(T)}, \Pi_{23(k)}^{(T)}), \qquad k = 1, \ldots, T, \qquad T \geq 1. \tag{A.7}$$

Multiplying the left side of (A.6) by $\mathbf{a}_{t-1}'$, we have $\sum_{k=1}^{T} \mathbf{D}_k^{(T)} \times \mathbf{X}_{t-k}\mathbf{a}_{t-1}' \xrightarrow{L_1} 0$, where $\xrightarrow{L_1}$ means convergence in the L_1 norm (as $T \rightarrow \infty$). Hence $\sum_{k=1}^{T} \mathbf{D}_k^{(T)}E[\mathbf{X}_{t-k}\mathbf{a}_{t-1}'] = \mathbf{D}_1^{(T)}\mathbf{V} \rightarrow 0$, because

$$E[\mathbf{X}_{t-k}\mathbf{a}_{t-1}'] = \mathbf{V}, \qquad k = 1,$$
$$= \mathbf{0}, \qquad k > 1.$$

The matrix $\mathbf{V}$ being nonsingular, we have $\mathbf{D}_1^{(T)} \rightarrow \mathbf{0}$; hence $\Pi_{21(1)} = \mathbf{0}$. Similarly, multiplying the left side of (A.7) by $\mathbf{a}_{t-2}'$, we find $\mathbf{D}_1^{(T)}E\{\mathbf{X}_{t-1}\mathbf{a}_{t-2}'\} + \mathbf{D}_2^{(T)}\mathbf{V} \rightarrow \mathbf{0}$, so that $\mathbf{D}_2^{(T)}\mathbf{V} \rightarrow \mathbf{0}$, and $\Pi_{21(2)} = \mathbf{0}$, and so on for $k = 3, 4, \ldots$. Thus $\Pi_{21(k)} = \mathbf{0}$, $k \geq 1$, and $\Pi_{21}(z) \equiv \mathbf{0}$.

Theorem 1. Because the process $\{\mathbf{X}_t\}$ is invertible, it can be expressed in a pure autoregressive form: $\Theta(B)^{-1}\Phi(B)\mathbf{X}_t = \mathbf{a}_t$ or, equivalently,

$$[\det \Theta(B)]^{-1}\Theta^*(B)\Phi(B)\mathbf{X}_t = \mathbf{a}_t, \tag{A.8}$$

where $\boldsymbol{\Theta}^*(z)$ denotes the adjoint matrix of $\boldsymbol{\Theta}(z)$. The matrix $\boldsymbol{\Theta}^*(z)$ can be written as $\Theta^*(z) = ((-1)^{i+j}A_{ij}(z))'$, where A_{ij} is the minor of $\boldsymbol{\Theta}(z)$ associated with the (i, j) element. Writing

$$\boldsymbol{\Pi}^*(z) = \boldsymbol{\Theta}^*(z)\boldsymbol{\Phi}(z) = (\Pi^*_{ij}(z)), \qquad \text{(A.9)}$$

it follows from Corollary 1 and (A.8) that X_i does not cause X_j $\Leftrightarrow \Pi^*_{ji}(z) \equiv 0$, because det $\boldsymbol{\Theta}(z) \neq 0$ for all $z \in C$ such that $|z| \leq 1$. We see from (A.9) that

$$\Pi^*_{ji}(z) = \sum_{k=1}^{n} (-1)^{k+j}\Phi_{ki}(z)A_{kj}(z) = \pm\left(\sum_{k=1}^{n} (-1)^{k+1}\Phi_{ki}(z)A_{kj}(z)\right)$$

$$= \pm\det(\boldsymbol{\Phi}_i(z), \boldsymbol{\Theta}_{(j)}(z)),$$

and the result is proved.

Theorem 2. The process (3.3) being invertible, it can be expressed as an (infinite) autoregressive process:

$$\boldsymbol{\Pi}(B)\begin{pmatrix} \mathbf{X}_{1t} \\ \mathbf{X}_{2t} \end{pmatrix} = \begin{pmatrix} \mathbf{a}_{1t} \\ \mathbf{a}_{2t} \end{pmatrix},$$

where

$$\boldsymbol{\Pi}(z) = \begin{pmatrix} \boldsymbol{\Pi}_{11}(z) & \boldsymbol{\Pi}_{12}(z) \\ \boldsymbol{\Pi}_{21}(z) & \boldsymbol{\Pi}_{22}(z) \end{pmatrix} = \boldsymbol{\Theta}(z)^{-1}\boldsymbol{\Phi}(z), \qquad |z| \leq 1.$$

Using the inverse of a partitioned matrix as given by Searle and Hausman (1970, p. 113) and omitting the argument z for simplicity, we have, for $|z| \leq 1$,

$$\boldsymbol{\Theta}^{-1} = \begin{pmatrix} \boldsymbol{\Theta}^{-1}(\mathbf{I} + \boldsymbol{\Theta}_{12}\mathbf{D}^{-1}\boldsymbol{\Theta}_{21}\boldsymbol{\Theta}_{11}^{-1}) & -\boldsymbol{\Theta}_{11}^{-1}\boldsymbol{\Theta}_{12}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\boldsymbol{\Theta}_{21}\boldsymbol{\Theta}_{11}^{-1} & \mathbf{D}^{-1} \end{pmatrix},$$

where $\mathbf{I}$ denotes the identity matrix and $\mathbf{D} = \boldsymbol{\Theta}_{22} - \boldsymbol{\Theta}_{21}\boldsymbol{\Theta}_{11}^{-1}\boldsymbol{\Theta}_{12}$. The existence of $\mathbf{D}^{-1}$ follows from the assumption that $\boldsymbol{\Theta}_{11}^{-1}$ exists and from the invertibility of model (3.3), because det $\boldsymbol{\Theta}$ = det$(\boldsymbol{\Theta}_{11}) \times$ det$(\boldsymbol{\Theta}_{22} - \boldsymbol{\Theta}_{21}\boldsymbol{\Theta}_{11}^{-1}\boldsymbol{\Theta}_{12})$; see, for example, Searle and Hausman (1970, p. 111). Then, from Proposition 1, $\mathbf{X}_1$ does not cause $\mathbf{X}_2$ $\Leftrightarrow \boldsymbol{\Pi}_{21} = -\mathbf{D}^{-1}\boldsymbol{\Theta}_{21}\boldsymbol{\Theta}_{11}^{-1}\boldsymbol{\Phi}_{11} + \mathbf{D}^{-1}\boldsymbol{\Phi}_{21} \equiv \mathbf{0} \Leftrightarrow -\Theta_{21}\boldsymbol{\Theta}_{11}^{-1}\boldsymbol{\Phi}_{11} + \boldsymbol{\Phi}_{21} \equiv \mathbf{0}$, and (3.4) follows.

Theorem 3. We first suppose that $n_2 = 1$. The generalization to $n_2 = s$, an arbitrary integer, is straightforward. For simplicity, let

$$\mathbf{A} = \boldsymbol{\Theta}_{11} = \begin{pmatrix} \Theta_{11} & \Theta_{12} & \cdots & \Theta_{1r} \\ \Theta_{21} & \Theta_{22} & \cdots & \Theta_{2r} \\ \vdots & \vdots & & \vdots \\ \Theta_{r1} & \Theta_{r2} & & \Theta_{rr} \end{pmatrix},$$

$$\mathbf{B} = \boldsymbol{\Phi}_{11} = \begin{pmatrix} \Phi_{11} & \Phi_{12} & \cdots & \Phi_{1r} \\ \Phi_{21} & \Phi_{22} & \cdots & \Phi_{2r} \\ \vdots & \vdots & & \vdots \\ \Phi_{r1} & \Phi_{r2} & & \Phi_{rr} \end{pmatrix},$$

$$\mathbf{C} = \boldsymbol{\Phi}_{21} = (\Phi_{r+1,1}, \Phi_{r+1,2}, \ldots, \Phi_{r+1,r}),$$

and

$$\mathbf{D} = \boldsymbol{\Theta}_{21} = (\Theta_{r+1,1}, \Theta_{r+1,2}, \ldots, \Theta_{r+1,r}).$$

From Theorem 2, $\mathbf{X}_1$ does not cause $\mathbf{X}_2 \Leftrightarrow \mathbf{C} - \mathbf{DA}^{-1}\mathbf{B} \equiv \mathbf{0} \Leftrightarrow \Delta\mathbf{C} - \mathbf{DA}^*\mathbf{B} \equiv \mathbf{0}$, where $\mathbf{A}^*$ and Δ denote the adjoint matrix and the determinant of $\mathbf{A}$: $\mathbf{A}^* = [(-1)^{i+j}A_{ij}]'$. The jth component of the line vector $\mathbf{DA}^*\mathbf{B}$ is $\alpha_j = \sum_{i=1}^{r} \Theta_{r+1,i}(\sum_{k=1}^{r} (-1)^{k+i}\Phi_{kj}A_{ki})$, and the jth component of the vector $\Delta\mathbf{C}$ is $\beta_j = \Phi_{r+1,j}\Delta$. Hence $\Delta\mathbf{C} - \mathbf{DA}^*\mathbf{B} = \mathbf{0} \Leftrightarrow E_j = \beta_j - \alpha_j = 0$, $j = 1, \ldots, r$. But we can write $\alpha_j = -\sum_{i=1}^{r} (-1)^i\Theta_{r+1,i}(\sum_{k=1}^{r} (-1)^{k+1}\Phi_{kj}A_{ki})$ and, consequently, E_j is $(\pm)$ the determinant of the matrix

$$\begin{pmatrix} \Phi_{1j} & \Theta_{11} & \cdots & \Theta_{1r} \\ \Phi_{2j} & \Theta_{21} & \cdots & \Theta_{2r} \\ \vdots & \vdots & & \vdots \\ \Phi_{rj} & \Theta_{r1} & \cdots & \Theta_{rr} \\ \Phi_{r+1,j} & \Theta_{r+1,1} & \cdots & \Theta_{r+1,r} \end{pmatrix},$$

where $j = 1, \ldots, r$. When $n_2 = s > 1$,

$$\mathbf{C} = \begin{pmatrix} \Phi_{r+1,1} & \cdots & \Phi_{r+1,r} \\ \Phi_{r+2,1} & \cdots & \Phi_{r+2,r} \\ \vdots & & \\ \Phi_{r+s,1} & \cdots & \Phi_{r+s,r} \end{pmatrix},$$

$$\mathbf{D} = \begin{pmatrix} \Theta_{r+1,1} & \cdots & \Theta_{r+1,r} \\ \Theta_{r+2,1} & \cdots & \Theta_{r+2,r} \\ \vdots & & \\ \Theta_{r+s,1} & \cdots & \Phi_{r+s,r} \end{pmatrix}$$

and

$$\mathbf{C} - \mathbf{DA}^{-1}\mathbf{B} = \mathbf{0} \Leftrightarrow [\Phi_{r+i,1}, \ldots, \Phi_{r+i,r}]$$

$$- [\Theta_{r+i,1}, \ldots, \Theta_{r+i,r}]\mathbf{A}^{-1}\mathbf{B} = \mathbf{0},\ i = 1, \ldots, s \Leftrightarrow$$

(3.5) is satisfied for $i = 1, \ldots, s$ and $j = 1, \ldots, r$.

[*Received August 1990. Revised November 1991.*]

REFERENCES

Basawa, I. V., Billard, L., and Srinivasan, R. (1984), "Large Sample Tests of Homogeneity for Time Series Models," *Biometrika*, 71, 203–206.

Basawa, I. V., and Koul, H. L. (1979), "Asymptotic Tests for Composite Hypotheses for Non-Ergodic Type Stochastic Processes," *Stochastic Processes and Their Applications*, 9, 291–305.

Boudjellaba, H. (1988), "Tests de Causalité Dans les Modèles Autorégressifs et Moyenne Mobile Multivariés," unpublished Ph.D. thesis, Université de Montréal, Département d'Informatique et de Recherche Opérationnelle.

Brockwell, P. J., and Davis, R. A. (1991), *Time Series: Theory and Methods* (2nd ed.), New York: Springer-Verlag.

Dickey, D. A., and Fuller, W. A. (1979), "Distribution of the Estimators for Autoregressive Time Series With a Unit Root," *Journal of the American Statistical Association*, 74, 427–431.

——— (1981), "Likelihood Ratio Statistics for Autoregressive Time Series With Unit Root," *Econometrica*, 49, 1057–1072.

Doan, T. A., Litterman, R., and Sims, C. (1986), "Forecasting and Conditional Projection Using Realistic Prior Distributions," *Econometric Reviews*, 3, 1–100.

Dufour, J.-M. (1989), "Nonlinear Hypotheses, Inequality Restrictions and Non-Nested Hypotheses: Exact Simultaneous Tests in Linear Regressions," *Econometrica*, 57, 335–355.

——— (1990), "Exact Tests and Confidence Sets in Linear Regressions With Autocorrelated Errors," *Econometrica*, 58, 475–494.

Eberts, R. W., and Steece, B. M. (1984), "A Test for Granger-Causality in a Multivariate ARMA Model," *Empirical Economics*, 9, 51–58.

Feige, E. L., and Pearce, D. K. (1979), "The Causal Relationship Between Money and Income: Some Caveats for Time Series Analysis," *Review of Economics and Statistics*, 61, 521–533.

Geweke, J. (1982), "Measurement of Linear Dependence and Feedback Between Multiple Time Series" (with comments), *Journal of the American Statistical Association*, 77, 304–324.

——— (1984a), "Inference and Causality in Economic Time Series Models," in *Handbook of Econometrics, Volume II*, eds. Z. Griliches and M. Intrilligator, Amsterdam: North-Holland, pp. 1102–1144.

——— (1984b), "Measures of Conditional Linear Dependence and Feedback Between Time Series," *Journal of the American Statistical Association*, 79, 907–915.

Gouriéroux, C., and Monfort, A. (1990), *Séries Temporelles et Modèles Dynamiques*, Paris: Economica.

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

Hsiao, C. (1979), "Autoregressive Modeling of Canadian Money and Income Data," *Journal of the American Statistical Association*, 74, 553–560.

——— (1982), "Autoregressive Modeling and Causal Ordering of Economic Variables," *Journal of Economic Dynamics and Control,* 4, 243–259.
Kang, H. (1981), "Necessary and Sufficient Conditions for Causality Testing in Multivariate ARMA Models," *Journal of Time Series Analysis,* 2, 95–101.
Litterman, R., and Weiss, L. (1985), "Money, Real Interest Rates and Output: A Reinterpretation of Postwar U.S. Data," *Econometrica,* 53, 129–156.
Newbold, P. (1982), "Causality Testing in Economics," in *Time Series Analysis: Theory and Practice 1,* ed. O. D. Anderson, Amsterdam: North-Holland, pp. 701–716.
Newbold, P., and Hotopp, M. (1986), "Testing Causality Using Efficiently Parametrized Vector ARMA Models," *Applied Mathematics and Computation,* 20, 329–348.
Nsiri, S., and Roy, R. (in press), "On the Invertibility of Multivariate Linear Processes," *Journal of Time Series Analysis.*
Osborn, D. R. (1984), "Causality Testing and Its Implications for Dynamic Econometric Models," *The Economic Journal,* 94, 82–95.
Phillips, P. C. B., and Perron, P. (1988), "Testing for a Unit Root in Time Series Regression," *Biometrika,* 75, 335–346.
Pierce, D. A., and Haugh, L. D. (1977), "Causality in Temporal Systems: Characterizations and Survey," *Journal of Econometrics,* 5, 265–293.
Priestley, M. B. (1981), *Spectral Analysis and Time Series* (2 vols.), New York: Academic Press.
Searle, S. R., and Hausman, W. H. (1970), *Matrix Algebra for Business and Economics,* New York: John Wiley.
Sims, C. (1972), "Money, Income and Causality," *American Economic Review,* 62 540–552.
——— (1980a), "Macroeconomics and Reality," *Econometrica,* 48, 1–48.
——— (1980b), "Comparison of Interwar and Postwar Business Cycles: Monetarism Reconsidered," *American Economic Review,* 70, 250–257.
Taylor, S. A. (1989), "A Comparison of Classical Tests of Criterion When Determining Granger Causality With a Bivariate ARMA Models," *Empirical Economics,* 14, 257–271.
Tiao, G. C., and Box, G. E. P. (1981), "Modeling Multiple Time Series With Applications," *Journal of the American Statistical Association,* 76, 802–816.
Tjostheim, D. (1981), "Granger-Causality in Multiple Time Series," *Journal of Econometrics,* 17, 157–176.