Testing Causality Between Two Vectors in Multivariate Autoregressive Moving Average Models

HAFIDA BOUDJELABA, 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.

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 économie (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, H3G 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, Said Nsiri, Franz Palm, Benedikt Pötscher, Jacques Raynauld, Eric Renault, and Kenneth G. Stewart for several useful comments and discussions.

© 1992 American Statistical Association
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 Tjøstheim (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 \{X_i : t \in Z\} and \{Y_i : t \in Z\} be two multivariate second-order stationary stochastic processes on the integers \(Z\), suppose that the dimension of \(X_i\) is \(n\), and write \(X_i = (X_{i1}, \ldots, X_{in})\). Let \(A_i\) be an “information set” containing \(X_i\) and \(Y_i\), and denote \(\tilde{A}_i = \{A_s : s < t\}\). For any information set \(I_t\), the best mean square linear predictor of \(X_{it}\) is denoted \(P(X_{it} | I_t)\), \(e_t(X_{it} | I_t) = X_{it} - P(X_{it} | I_t)\) is the corresponding prediction error, and \(\sigma^2(X_{it} | I_t)\) is the variance of \(e_t\). The predictor \(P(X_{it} | I_t)\) is the orthogonal projection of \(X_{it}\) on the Hilbert space spanned by the variables in \(I_t\). For Gaussian processes, \(P(X_{it} | I_t) = E[X_{it} | I_t]\), 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 \(X_i\) is the vector \(P(X_i | I_t) = (P(X_{i1} | I_t), \ldots, P(X_{in} | I_t))^\prime\), the corresponding vector of prediction errors is given by \(e_t(X_i | I_t) = (e_{i1}(X_{i1} | I_t), \ldots, e_{in}(X_{in} | I_t))^\prime\), and we denote \(\Sigma(X_i | I_t)\) the covariance matrix of \(e_t\).

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 \(\tilde{A}_t - \tilde{Y}_t\) represents all the information in \(\tilde{A}_t\) apart from the information in \(\tilde{Y}_t\).

**Definition 1.** The vector \(Y\) does not cause the vector \(X\) if

\[
\sigma^2(X_{it} | \tilde{A}_i) = \sigma^2(X_{it} | \tilde{A}_i - \tilde{Y}_i), \quad i = 1, \ldots, n. \tag{1.1}
\]

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

**Definition 2.** The vector \(Y\) does not cause the vector \(X\) if

\[
\Sigma(X_i | \tilde{A}_i) = \Sigma(X_i | \tilde{A}_i - \tilde{Y}_i). \tag{1.2}
\]

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 \(P(X_{it} | \tilde{A}_i) = P(X_{it} | \tilde{A}_i - \tilde{Y}_i), i = 1, \ldots, n\) or

\[
P(X_{it} | \tilde{A}_i) = P(X_{it} | \tilde{A}_i - \tilde{Y}_i), \tag{1.3}
\]

where the equality holds in the \(L_2\) sense (hence with probability 1). Obviously, (1.3) implies (1.2). Conversely, noncausality in Tjøstheim’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 \(\{X_i : t \in Z\}\) be a second-order stationary purely indeterministic \(n\)-dimensional vector process. Without loss of generality, we can suppose that \(E[X_i] = 0\). By Wold’s decomposition theorem, the process \(\{X_i\}\) admits the following representation:

\[
X_i = \sum_{j=0}^{\infty} \Psi_j a_{i-j}, \quad t \in Z, \tag{2.1}
\]

where the \(\Psi_j\)’s are \(n \times n\) matrices such that \(\sum_{j=0}^{\infty} |\Psi_j|^2 < \infty\), \(|\cdot|\) is the Euclidean norm matrix, \(\Psi_0 = I\) is the identity matrix of order \(n\), and \(\{a_i : t \in Z\}\) is the innovation process. The \(a_i\)’s are uncorrelated random vectors with mean \(0\) and nonsingular covariance matrix \(V\). We suppose that the process (2.1) is invertible; that is, \(X_i\) admits a possibly infinite autoregressive representation

\[
X_i = \sum_{j=1}^{\infty} \Pi_j X_{i-j} + a_i, \quad t \in Z, \tag{2.2}
\]

where the \(\Pi_j\)’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

\[
\Pi(B) X_i = a_i, \quad t \in Z, \tag{2.3}
\]

where \(\Pi(B) = -\sum_{j=0}^{\infty} \Pi_j B^j\) is \((\Pi(B))_{n \times n}\) is an \(n \times n\) matrix of power series in the operator \(B\) and \(\Pi_0 = -I_{n \times n}\).
Let us partition \( X_i \) and \( a_i \) into three subvectors as follows:

\[
X_i = (X_{1i}, X_{2i}, X_{3i})', \quad a_i = (a_{1i}, a_{2i}, a_{3i})',
\]

where \( X_{ai} \) and \( a_{ai} \) 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 \( n_3 = 0, \) \( X_i \) is partitioned into two subvectors only. Using the corresponding partition of \( \Pi(B), (2.3) \) can be written as

\[
\begin{bmatrix}
\Pi_{11}(B) & \Pi_{12}(B) & \Pi_{13}(B) \\
\Pi_{21}(B) & \Pi_{22}(B) & \Pi_{23}(B) \\
\Pi_{31}(B) & \Pi_{32}(B) & \Pi_{33}(B)
\end{bmatrix} \begin{bmatrix}
X_{1i} \\
X_{2i} \\
X_{3i}
\end{bmatrix} = \begin{bmatrix}
a_{1i} \\
a_{2i} \\
a_{3i}
\end{bmatrix}.
\]  

(2.4)

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

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), \( X_i \) does not cause \( X_2 \) if and only if \( \Pi_{ij}(z) = 0 \).

By taking \( n_1 = n_2 = 1 \), we obtain a characterization of noncausality between any two components of \( X \).

**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 \( \Pi_{ij}(z) = 0 \).

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), \( X_i \) does not cause \( X_j \) if and only if \( \Pi_{ij}(z) = 0 \).

**3. CAUSALITY IN ARMA PROCESSES**

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

\[
\Phi(B)X_i = \Theta(B)a_i,
\]

(3.1)

be stationary and invertible, where \( \Phi(B) = I - \Phi_1B - \cdots - \Phi_pB^p \) and \( \Theta(B) = I - \Theta_1B - \cdots - \Theta_qB^q \). We also assume that the parameters in \( \Phi(B) \) and \( \Theta(B) \) are identified (uniquely defined) as functions of the autocovariance matrices of \( X_i \), so that \( \Phi(B) \) and \( \Theta(B) \) have no common factor. Using Corollary 1, we derive the following characterization of noncausality between any two components of \( X \). 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

\[
\det(\Phi_i(z), \Theta_{ij}(z)) = 0, \quad (3.2)
\]

where \( \Phi_i(z) \) is the \( i \)th column of the matrix \( \Phi(z) \) and \( \Theta_{ij}(z) \) is the matrix \( \Theta(z) \) without its \( j \)th column.

If \( X = (X_1', X_2', X_3')' \), 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), \( X_1 \) does not cause \( X_3 \) 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 \( X \) is partitioned into two subvectors: \( X = (X_1', X_2')' \), where \( X_i \) is \( n_i \times 1, i = 1, 2, \) and \( n_1 + n_2 = n. \) In this case model (3.1) can be rewritten as

\[
\begin{bmatrix}
\Phi_{11}(B) & \Phi_{12}(B) \\
\Phi_{21}(B) & \Phi_{22}(B)
\end{bmatrix} \begin{bmatrix}
X_{1i} \\
X_{2i}
\end{bmatrix} = \begin{bmatrix}
\Theta_{11}(B) & \Theta_{12}(B) \\
\Theta_{21}(B) & \Theta_{22}(B)
\end{bmatrix} \begin{bmatrix}
a_{1i} \\
a_{2i}
\end{bmatrix},
\]

(3.3)

where \( \Phi_{ij}(B) \) and \( \Theta_{ij}(B) \) are \( n_i \times n_j \) matrices, \( i, j = 1, 2. \)

Then the condition of noncausality between \( X_1 \) and \( X_2 \) can be formulated in the following way.

**Theorem 2.** Suppose that the stationary ARMA process (3.3) is invertible, with \( \det(\Theta_{11}(z)) \neq 0 \) for all \( z \in C \) such that \( |z| \leq 1 \). Then, \( X_1 \) does not cause \( X_2 \) if and only if

\[
\Phi_{21}(z) - \Theta_{21}(z)\Theta_{11}(z)^{-1}\Phi_{11}(z) = 0. \quad (3.4)
\]

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

The following characterization of the noncausality between \( X_1 \) and \( X_2 \) is more convenient to deal with applications. Write \( \Phi(z) = (\Phi_{ij}(z))_{hoz} \) and \( \Theta(z) = (\Theta_{ij}(z))_{hoz}. \) To simplify the notation, we omit the argument \( z \) in the matrices \( \Phi_{ij}(z) \) and \( \Theta_{ij}(z). \)

**Theorem 3.** Suppose that the stationary ARMA process (3.3) is invertible, with \( \det(\Theta_{11}(z)) \neq 0 \) for all \( z \in C \) such that \( |z| \leq 1 \). Then, \( X_1 \) does not cause \( X_2 \) if and only if

\[
\begin{bmatrix}
\Phi_{ij} & \Theta_{11} & \Theta_{12} & \cdots & \Theta_{1r} \\
\Phi_{2j} & \Theta_{21} & \Theta_{22} & \cdots & \Theta_{2r} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\Phi_{rj} & \Theta_{r1} & \Theta_{r2} & \cdots & \Theta_{rr} \\
\Phi_{ri} & \Theta_{r1} & \Theta_{r2} & \cdots & \Theta_{r+s+r}
\end{bmatrix}
= 0
\]

(3.5)

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

**Example 1.** Let \( X_i \) 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

\[
\begin{bmatrix}
\Phi_{11} & \Theta_{11} & \Theta_{13} \\
\Phi_{21} & \Theta_{21} & \Theta_{23} \\
\Phi_{31} & \Theta_{31} & \Theta_{33}
\end{bmatrix}
= 0
\]

and

\[
\begin{bmatrix}
\Phi_{11} & \Theta_{11} & \Theta_{13} \\
\Phi_{21} & \Theta_{21} & \Theta_{22} \\
\Phi_{31} & \Theta_{31} & \Theta_{32}
\end{bmatrix}
= 0.
\]

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

\[
\begin{bmatrix}
\Phi_{11} & \Theta_{11} \\
\Phi_{1s+i} & \Theta_{1s+i}
\end{bmatrix}
= 0, \quad i = 1, 2.
\]

Thus Theorem 1 and Corollary 3 lead to the evaluation of \( 3 \times 3 \) determinants, whereas Theorem 3 leads to the evaluation of \( 2 \times 2 \) 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 \( X_i \). In many situations, appli-
cation 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 \(X = (X_1, X_2)\), we want to test the hypothesis of noncausality between \(X_1\) and \(X_2\). 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 \(\beta\) as the vector of all AR and MA parameters, the noncausality conditions lead to (possibly nonlinear) constraints on an \(l \times 1\) subvector \(\beta_1\) of \(\beta\). We will denote these restrictions by

\[
R_j(\beta_1) = 0, \quad j = 1, \ldots, K, \tag{4.1}
\]

where \(K \leq l\).

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 \(T(\hat{\beta})\) be the \(l \times K\) matrix of derivatives

\[
T(\hat{\beta}) = \left( \frac{\partial R(\hat{\beta})}{\partial \beta} \right)_{\hat{\beta}} \tag{4.2}
\]

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

\[
\xi_W = NR(\beta)^T[T(\hat{\beta})]^{-1}R(\beta), \tag{4.3}
\]

where \(R(\beta) = (R_1(\beta), \ldots, R_K(\beta))^T\). In this definition \(T(\beta)\) 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 \(L(\beta, X)\) be the logarithm of the likelihood function of the \(N\) observations. The likelihood ratio statistic is given by

\[
\xi_{LR} = 2[L(\hat{\beta}, X) - L(\hat{\beta}^*, X)], \tag{4.4}
\]

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

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

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 \(X_2\)

\[
\Leftrightarrow \Phi_{11}(z)\Phi_{21}(z) - \Phi_{12}(z)\Phi_{11}(z) = 0,
\]

\[
\Leftrightarrow (\phi_{21} - \theta_{21})z + (\theta_{11}\phi_{21} - \phi_{22}\phi_{11})z^2 = 0,
\]

\[
\Leftrightarrow \phi_{21} - \theta_{21} = 0 \quad \text{and} \quad \phi_{11}\theta_{21} - \phi_{22}\phi_{11} = 0. \tag{4.4}
\]

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

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

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:

\[
H_0^1: \phi_{21} - \theta_{21} = 0; \quad H_0^2: \phi_{21} = \theta_{21} = 0; \quad H_0^3: \phi_{21} \neq 0, \quad \phi_{21} - \theta_{21} = 0, \quad \text{and} \quad \phi_{11} - \theta_{11} = 0;
\]

and \(H_0^4: \phi_{11} - \theta_{11} = 0\). We have the following relations:

\[
H_0^1 \cap H_0^2 \subseteq H_0^3 \subseteq H_0^4 \subseteq H_0^5 \subseteq H_0^6 \subseteq H_0^7. \tag{4.6}
\]

Suppose now that we can test the three hypotheses \(H_0^1, H_0^2, H_0^3\) separately. For \(H_0^3\), it will be sufficient to test \(\phi_{21} - \theta_{21} = 0\) under the assumption \(\phi_{21} \neq 0\). Then, for given significance level \(\alpha(0 < \alpha < 1)\), we can test \(H_0^3\) in the following way. Let \(\alpha = \alpha_1 + \alpha_2\) and \(0 < \alpha_i < 1, i = 1, 2\).

1. We first test \(H_0^3\) at level \(\alpha_1\). If \(H_0^3\) is rejected, then \(H_0^3\) is rejected too and the procedure stops.
2. If \(H_0^3\) is not rejected, then we test \(H_0^2\) at level \(\alpha_2\). If \(H_0^2\) is not rejected, then we cannot reject \(H_0\) and we stop.
3. If we reject \(H_0^3\), then we test \(H_0^3\). If \(H_0^3\) is rejected, then \(H_0\) is also rejected. If \(H_0^3\) 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 \(P[\psi(\alpha_1, \alpha_2) = 1] \leq P[A] + P[B]\), where \(A\) represents the event “reject \(H_0^3\) and \(B\) represents the event “do not reject \(H_0^3\) and reject \(H_0^2\).” For the event \(B\) there are two possible cases (under \(H_0\)):

1. \(\phi_{21} = \theta_{21} = 0\), in which case we have \(P[B] \leq P[\text{reject} H_0^2] = \alpha_2\).
fore 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 \( y_i = (1 - B) \ln \text{GNP}_i, m_{1i} = (1 - B) \ln M_1, \) and \( m_{2i} = (1 - B) \ln M_2. \) Using the approach of Tiao and Box (1981), we first performed bivariate analyses of \( (y_i, m_{1i}), (y_i, m_{2i}), \) and \( (m_{1i}, m_{2i}) \) and then performed a trivariate analysis of \( (y_i, m_{1i}, m_{2i}). \) 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 \( (y_i, m_{1i}), \) we see from Theorem 1 that

\[
\begin{align*}
   y_i \Rightarrow m_{1i} &\iff \begin{bmatrix} \Phi_{11}(z) & \Phi_{12}(z) \\ \Phi_{21}(z) & \Phi_{22}(z) \end{bmatrix} = 0 \\
   &\iff \begin{bmatrix} 1 - \phi_{11}^1 z & 1 - \theta_{11}^1 z \\ -\phi_{12}^1 z & 0 \end{bmatrix} = 0 \iff \phi_{21}^1 = 0 \text{ and } \\
   &\iff \phi_{11}^1 \theta_{11}^1 = 0 \iff \phi_{21}^1 = 0.
\end{align*}
\]

**Bivariate Models**

\[
\begin{bmatrix}
   1 - .799B & -203B \\
   .049 & .040
\end{bmatrix}
= \begin{bmatrix}
   1.861B & 0 \\
   0 & 1 - .5088B^4
\end{bmatrix}
\]

\[
\begin{bmatrix}
   1 - .876B \\
   .093
\end{bmatrix}
= \begin{bmatrix}
   1.657B & 0 \\
   0 & 1 - .717B^4
\end{bmatrix}
\]

\[
\begin{bmatrix}
   1 - .640B \\
   .065
\end{bmatrix}
= \begin{bmatrix}
   1.569B & 0 \\
   0 & 1 - .616B^4
\end{bmatrix}
\]

**Trivariate model**

\[
\begin{bmatrix}
   1 - .610B^2 & -0.95B - 299B^2 \\
   .077 & .068
\end{bmatrix}
= \begin{bmatrix}
   1 & - .627B \\
   .065 & .065
\end{bmatrix}
\]

\[
\begin{bmatrix}
   0 & -224B^2 \\
   .067
\end{bmatrix}
= \begin{bmatrix}
   1 & - .784B \\
   .059 & .059
\end{bmatrix}
\]

\[
\begin{bmatrix}
   1 - .668B^2 & 0 \\
   .105
\end{bmatrix}
= \begin{bmatrix}
   1 - .448B^3 & 1 - 578B^4 \\
   .102 & .080
\end{bmatrix}
\]

\[
\begin{bmatrix}
   0 & 0 \\
   .083
\end{bmatrix}
= \begin{bmatrix}
   1 - 0.72B^2 - 594B^4 \\
   .083 & .078
\end{bmatrix}
\]

**Error covariance matrices (× 10^4 and in the same order as the models)**

\[
\begin{bmatrix}
   1.28 & .29 \\
   .29 & 2.17
\end{bmatrix}
\]

\[
\begin{bmatrix}
   1.64 & .24 \\
   .24 & 1.01
\end{bmatrix}
\]

\[
\begin{bmatrix}
   1.98 & .77 \\
   .77 & 1.06
\end{bmatrix}
\]

\[
\begin{bmatrix}
   1.22 & .33 \quad .13 \\
   .33 & 1.87 \quad .70 \\
   .13 & .70 \quad 1.03
\end{bmatrix}
\]

**Figure 2. Multivariate ARMA Models. NOTE:** Estimated standard errors are given in parentheses.
and
\[ m_{1t} \rightarrow y_t \iff \begin{vmatrix} \Phi_{12}(z) & \Theta_{12}(z) \\ \Phi_{22}(z) & \Theta_{22}(z) \end{vmatrix} = 0 \iff \begin{vmatrix} -\phi_{12}^{(1)} z & 0 \\ 1 - \phi_{22}^{(4)} z^2 & 1 - \theta_{22}^{(4)} z^2 \end{vmatrix} = 0 \iff \phi_{12}^{(1)} = 0, \]
where → means “does not cause,” \( \Phi_{ij}(B) \) and \( \Theta_{ij}(B) \) refer to the corresponding lag polynomials in the model for \( y_t \) and \( m_{1t} \), 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 \( m_{1t} \)” \( y_t \rightarrow m_{1t} \) take the values 13.7 and 36.9. (Causality tests are summarized in Table 1). To test \( m_{1t} \rightarrow y_t \), the corresponding statistics are 25.8 and 64.6. Because the asymptotic null distribution of these test statistics is \( \chi^2_1 \), 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 \( y_t \rightarrow m_{2t} \), Theorem 1 implies
\[ y_t \rightarrow m_{2t} \iff \begin{vmatrix} \Phi_{11}(z) & \Theta_{11}(z) \\ \Phi_{21}(z) & \Theta_{21}(z) \end{vmatrix} = 0 \iff \begin{vmatrix} 1 - \phi_{11}^{(1)} z & 1 - \theta_{11}^{(1)} z \\ -\phi_{21}^{(1)} z & 0 \end{vmatrix} = 0 \iff \phi_{21}^{(1)} = 0 \]
and
\[ m_{2t} \rightarrow y_t \iff \begin{vmatrix} \Phi_{12}(z) & \Theta_{12}(z) \\ \Phi_{22}(z) & \Theta_{22}(z) \end{vmatrix} = 0 \iff \begin{vmatrix} 0 & 0 \\ 1 - \phi_{22}^{(1)} z & 1 - \theta_{22}^{(4)} z^2 \end{vmatrix} = 0, \]
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 \( m_{2t} \rightarrow y_t \) is satisfied exactly by the model, whereas \( y_t \rightarrow m_{2t} \) 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 \( m_{1t} \rightarrow m_{2t} \iff \phi_{12}^{(1)} = 0 \) and \( m_{2t} \rightarrow m_{1t} \iff \phi_{21}^{(1)} = 0 \). 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 \( y_t \rightarrow y_t, y_t \rightarrow m_{1t}, \) and \( m_{1t} \rightarrow y_t \). The most striking results here is that money stock changes \( (m_{1t} \text{ and } m_{2t}) \) 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 \( (m_{1t} \text{ and } m_{2t}) \) 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
\[ y_t \rightarrow (m_{1t}, m_{2t})' \iff \begin{vmatrix} \Phi_{11}(z) & \Theta_{11}(z) \\ \Phi_{21}(z) & \Theta_{21}(z) \end{vmatrix} = 0 \text{ and } \begin{vmatrix} 1 - \phi_{11}^{(2)} z^2 & 1 - \theta_{11}^{(2)} z^2 \\ 0 & -\theta_{21}^{(5)} z^5 \end{vmatrix} = 0 \iff \theta_{21}^{(5)} = 0. \]
Further, by Corollary 2, \( (m_{1t}, m_{2t})' \rightarrow y_t \iff m_{1t} \rightarrow y_t \) and

Table 1. Causality Tests

<table>
<thead>
<tr>
<th>Null hypothesis (noncausality)</th>
<th>Parametric representation</th>
<th>Wald statistic</th>
<th>Likelihood ratio statistic</th>
<th>Degrees of freedom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bivariate models</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( y_t \rightarrow m_{1t} )</td>
<td>( \phi_{11}^{(1)} = 0 )</td>
<td>13.7*</td>
<td>36.8*</td>
<td>1</td>
</tr>
<tr>
<td>( m_{1t} \rightarrow y_t )</td>
<td>( \phi_{21}^{(1)} = 0 )</td>
<td>25.8*</td>
<td>64.6*</td>
<td>1</td>
</tr>
<tr>
<td>( m_{1t} \rightarrow m_{2t} )</td>
<td>( \phi_{22}^{(1)} = 0 )</td>
<td>16.1*</td>
<td>23.3*</td>
<td>1</td>
</tr>
<tr>
<td>( m_{2t} \rightarrow y_t )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( m_{2t} \rightarrow m_{1t} )</td>
<td>( \phi_{21}^{(2)} = 0 )</td>
<td>10.5*</td>
<td>20.4*</td>
<td>1</td>
</tr>
<tr>
<td>( m_{2t} \rightarrow m_{2t} )</td>
<td>( \phi_{22}^{(2)} = 0 )</td>
<td>96.9*</td>
<td>45.6*</td>
<td>1</td>
</tr>
<tr>
<td>( m_{1t} \rightarrow m_{2t} )</td>
<td>( \phi_{22}^{(3)} = 0 )</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
|\( m_{1t} \rightarrow m_{2t} \) and \( m_{2t} \rightarrow y_t \) | \( \phi_{22}^{(4)} = \phi_{22}^{(5)} = 0 \) | 62.7* | 62.7* | 2

| Trivariate models             |                           |                |                           |                    |
|\( y_t \rightarrow (m_{1t}, m_{2t})' \) | \( \theta_{21}^{(1)} = 0 \) | 2.1            | 3.5                       | 1                   |
|\( (m_{1t}, m_{2t})' \rightarrow y_t \) | \( \theta_{21}^{(2)} = \phi_{22}^{(2)} = 0 \) | 69.3*          | 62.7*                    | 2                   |
|\( y_t \rightarrow m_{1t} \)  | \( \phi_{21}^{(3)} = 0 \) | 2.1            | 3.5                       | 1                   |
|\( m_{2t} \rightarrow m_{1t} \)| \( \phi_{22}^{(3)} = \phi_{22}^{(5)} = 0 \) | 69.3*          | 62.7*                    | 2                   |
|\( m_{2t} \rightarrow y_t \)  |                           |                |                           |                    |
|\( m_{2t} \rightarrow m_{2t} \)| \( \phi_{22}^{(4)} = \phi_{22}^{(5)} = 0 \) | 11.2*          | 17.2*                    | 1                   |
|\( m_{2t} \rightarrow m_{2t} \)| \( \phi_{22}^{(5)} = 0 \) | 93.0*          | 46.2*                    | 1                   |

* Significant at .06.

a. The condition of noncausality is satisfied exactly by the model.

b. Because the structure of the model implies that \( y_t \rightarrow m_{2t} \) and \( m_{2t} \rightarrow y_t \), the statistics for testing noncausality between \( m_{2t} \) and \( y_t \) are identical to those between \( (m_{1t}, m_{2t}) \) and \( y_t \).
We see that the condition for $m_{21} \not\rightarrow y_i$ holds exactly in this model, so that $(m_{11}, m_{21})' \not\rightarrow y_i \iff m_{11} \not\rightarrow y_i \iff \phi_{12} = 0$ (2) = 0. Using these conditions, we see in Table 1 that $(m_{11}, m_{21})' \not\rightarrow y_i$ is strongly rejected, but $y_i \rightarrow (m_{11}, m_{21})'$ is accepted (at level .05). Causality appears to be unidirectional from $(m_{11}, m_{21})'$ to $y_i$. This result agrees with the one obtained by Osborn (1984) using a different methodology.

By Theorem 1 we also have $m_{11} \not\rightarrow m_{21} \iff \phi_{23} = 0$, $m_{21} \not\rightarrow m_{11} \iff \phi_{32} = 0$, and $m_{11} \not\rightarrow y_i \iff \phi_{12} = 0$, and the conditions for $m_{21} \not\rightarrow y_i$, and $y_i \rightarrow m_{21}$ are satisfied exactly by the model. We thus find that $m_{11} \rightarrow y_i$, whereas $y_i$ and $m_{21}$ do not cause each other. Further, the hypotheses $\phi_{32} = 0$ and $\phi_{23} = 0$ are strongly rejected, suggesting feedback between $m_{11}$ and $m_{21}$.

Thus the causality structure that emerges from the trivariate model of $(y_i, m_{11}, m_{21})'$ is $m_{21} \not\rightarrow m_{11} \rightarrow y_i$. There is no direct link between $m_{21}$ and $y_i$ and no causality running from $m_{11}$ and $m_{21}$ towards $y_i$. Due to the presence of feedback between $m_{11}$ and $m_{21}$, 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 sets 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.
where $\Theta(z)$ denotes the adjoint matrix of $\Theta(z)$. The matrix $\Theta^*(z)$ can be written as $\Theta^*(z) = (-1)^k A_k(z)\Phi(z)$, where $A_k$ is the minor of $\Theta(z)$ associated with the $(i,j)$ element. Writing
\[ \Pi^*(z) = \Theta^*(z) \Phi(z) = (\Pi^*_{ij}(z)), \] (A.9)
it follows from Corollary 1 and (A.8) that $X_i$ does not cause $X_j \iff \Pi^*_{ij}(z) = 0$, because det $\Theta(z) \neq 0$ for all $z \in C$ such that $|z| \leq 1$. We see from (A.9) that
\[ \Pi^*_{ij}(z) = \sum_{k=1}^n (-1)^{k+i,j} \Phi_k(z) A_k(z) = \pm \left( \sum_{k=1}^n (-1)^{k+1} \Phi_k(z) A_k(z) \right) \]
= $\pm$det$(\Phi(z), \Theta_{ij}(z))$,
and the result is proved.

**Theorem 2.** The process (3.3) being invertible, it can be expressed as an (infinite) autoregressive process:
\[ \Pi(B) \begin{bmatrix} X_i \\ X_{z_i} \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}, \]
where
\[ \Pi(z) = \begin{bmatrix} \Pi_{11}(z) & \Pi_{12}(z) \\ \Pi_{21}(z) & \Pi_{22}(z) \end{bmatrix} = \Theta(z)^{-1} \Phi(z), \quad |z| \leq 1. \]

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 $|z| \leq 1$,
\[ \Theta^{-1} = \begin{pmatrix} \Theta^{-1} & -\Theta^{-1} \Theta D^{-1} \\ -\Theta D^{-1} & \Theta^{-1} \Theta D^{-1} \end{pmatrix}, \]
where $I$ denotes the identity matrix and $D = \Theta_{22} - \Theta_{21}\Theta_{11}^{-1}\Theta_{12}$. The existence of $D^{-1}$ follows from the assumption that $\Theta_{11}$ exists and from the invertibility of model (3.3), because det $\Theta = \det(\Theta_{11}) \times \det(\Theta_{22} - \Theta_{21}\Theta_{11}^{-1}\Theta_{12})$; see, for example, Searle and Hausman (1970, p. 111). Then, from Proposition 1, $X_1$ does not cause $X_2 \iff \Pi_{21} = -D^{-1}\Theta_{21}\Theta_{11}^{-1}\Phi_{11} + D^{-1}\Phi_{21} = 0$ 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
\[ A = \Theta_{11} = \begin{bmatrix} \Theta_{11} & \cdots & \Theta_{1s} \\ \cdots & \cdots & \cdots \\ \Theta_{11} & \cdots & \Theta_{1s} \end{bmatrix}, \]
\[ B = \Phi_{11} = \begin{bmatrix} \Phi_{11} \\ \cdots \\ \Phi_{11} \end{bmatrix}, \]
\[ C = \Phi_{21} = \begin{bmatrix} \Phi_{11} \\ \cdots \\ \Phi_{11} \end{bmatrix}, \] and
\[ D = \Theta_{21} = (\Theta_{s+1,1} \cdots \Theta_{s+1,s}). \]

From Theorem 2, $X_i$ does not cause $X_j \iff C - DA^{-1}B = 0 \iff DC - DA^{-1}B = 0$, where $^*$ denotes the adjoint matrix and the determinant of $A: A^* = ([(-1)^{j+i} A_{ij}]^*)$. The $j$th component of the line vector $DA^*B = \alpha_j = \sum_{k=1}^n (-1)^{k+j} \Phi_k A_k$, and the $j$th component of the vector $DC = \beta_j = \Phi_{s+1,j} D$. Hence $\Delta C - DA^{-1}B = 0 \iff E_j = \beta_j - \alpha_j = 0$, $j = 1, \ldots, r$. But we can write $\alpha_j = -\sum_{k=1}^n (-1)^{k+j} \Phi_k A_k$ and, consequently, $E_j$ is the determinant of the matrix
\[ \Phi_{ij} \quad \Theta_{11} \quad \cdots \quad \Theta_{1r} \\
\vdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\
\Phi_{ij} \quad \Theta_{r1} \quad \cdots \quad \Theta_{rr} \\
\Phi_{s+1,j} \quad \Theta_{s+1,1} \quad \cdots \quad \Theta_{s+1,s} \\
\Phi_{s+2,j} \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\
\Phi_{s+r,j} \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\
\Phi_{s+r,s} \\
\] where $j = 1, \ldots, r$. When $n_2 = s > 1$,
\[ C = \begin{bmatrix} \Phi_{s+1,1} & \cdots & \Phi_{s+1,r} \\ \cdots & \cdots & \cdots \\ \Phi_{s+r,1} & \cdots & \Phi_{s+r,r} \end{bmatrix}, \]
\[ D = \begin{bmatrix} \Theta_{s+1,1} & \cdots & \Theta_{s+1,r} \\ \cdots & \cdots & \cdots \\ \Theta_{s+r,1} & \cdots & \Theta_{s+r,r} \end{bmatrix}, \]
and
\[ C - DA^{-1}B = 0 \iff \Phi_{s+1,1} \cdots \Phi_{s+r,s} \]
\[ - [\Theta_{s+1,1} \cdots \Theta_{s+r,s}] A^{-1} B = 0, \quad i = 1, \ldots, s \iff \]
(3.5) is satisfied for $i = 1, \ldots, s$ and $j = 1, \ldots, r$.

[Received August 1990. Revised November 1991.]

**REFERENCES**


