ELSEVIER Journal of Econometrics 63 (1994) 271–287

JOURNAL OF Econometrics

Simplified conditions for noncausality between vectors in multivariate ARMA models

Hafida Boudjellaba[a], Jean-Marie Dufour*[,b], Roch Roy[c]

[a] *Laurentian University, Sudbury, Ont. P3E 2C6, Canada*
[b] *CRDE and Dépt. Sc. Econ., Université de Montréal, Montréal, Qué. H3C 3J7, Canada*
[c] *CRDE and Dépt. d' Informatique, Université de Montréal, Montréal, Qué. H3C 3J7, Canada*

(Received October 1993)

Abstract

This article derives necessary and sufficient conditions for noncausality between two vectors of variables in stationary invertible ARMA processes. Earlier conditions proposed by Boudjellaba, Dufour, and Roy (1992a) are shown to hold under weaker regularity assumptions and then generalized to cover the important case where the two vectors do not necessarily embody all the variables considered in the analysis. The conditions so obtained can be considerably simpler and easier to implement than earlier ones. Testing of the conditions derived is also discussed and the results are applied to a model of Canadian money, income, and interest rates.

Key words: Granger causality; Multivariate ARMA process; Causality test
JEL classification: C22; C5; E4; E3

1. Introduction

Granger (1969) has proposed a definition of causality between time series which has been applied frequently. Most of the literature on this topic is

* Corresponding author.

This work was supported by the Université Libre de Bruxelles (Institut de Statistique), the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, and the Government of Québec (Fonds FCAR). Earlier versions of this paper were presented at the (EC)^2 Conference in Rotterdam (December 1991), as well as seminars at Queen's University, the University of Southern California, the University of Pennsylvania, and the University of Toronto. We thank Normand Ranger for his programming assistance, and Clive Granger, Cheng Hsiao, Jan Kiviet, Helmut Lütkepohl, James MacKinnon, Angelo Melino, Franz Palm, Dale Poirier, Benedikt Pötscher, Eric Renault, and two anonymous referees for useful comments and discussions.

0304-4076/94/$07.00 © 1994 Elsevier Science B.V. All rights reserved
SSDI 0304407693015687

concerned with bivariate relationships or uses finite-order autoregressive (AR) specifications; see the reviews of Pierce and Haugh (1977), Newbold (1982), Geweke (1984), Gouriéroux and Monfort (1990, Ch. X), and Lütkepohl (1991). In this paper, we study causality in the more general framework of multivariate autoregressive moving average (ARMA) models. While AR models are relatively easy to estimate (hence their popularity), ARMA models have two important advantages. First, ARMA models can be considerably more parsimonious than AR models, and may thus lead to more efficient forecasts as well as more powerful tests. For example, it is easy to see that a large number of parameters may be required for an AR model to approximate even a low-order moving average model (which is autoregressive of infinite order). For earlier work on causality based on finite-order AR specifications, the reader may consult Sims (1980a, b), Hsiao (1979, 1982), Litterman and Weiss (1985), and Lütkepohl (1991). Second, any subvector of a stationary ARMA process still follows an ARMA process (of possibly different order), while a subvector of an AR process does not generally follow an AR process (but an ARMA): the class of ARMA models is invariant to disaggregation, while AR models are not, and thus provides a more coherent modelling framework.[1]

Earlier work on the analysis and testing of causality in the context of multivariate ARMA models can be found in Boudjellaba, Dufour, and Roy (1992a) hereafter referred to as BDR, Eberts and Steece (1984), Kang (1981), Newbold (1982), Newbold and Hotopp (1986), Osborn (1984), and Taylor (1989). 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. Similarly, Newbold (1982) suggested using a likelihood ratio to test a sufficient condition for noncausality in a bivariate ARMA model, while Eberts and Steece (1984) and Taylor (1989) studied Wald, likelihood ratio, and Rao score tests for the Kang's conditions. In the multivariate case with more than two variables, Osborn (1984) examined Granger causality by rewriting the model so that the autoregressive polynomials are the same for all variables. This approach, however, does not take into account all the constraints implied by the ARMA specification and may easily lead to the estimation of an unduly large number of parameters. After giving a formulation of the concept of causality in a multivariate situation which is equivalent but more convenient than the one given by Tjostheim (1981), BDR derived a necessary and sufficient condition for noncausality between two vectors in linear invertible processes and applied it to stationary invertible ARMA processes (see also Boudjellaba, 1988). For ARMA processes, they also obtained simpler conditions in two special cases: (1) the case where the two

[1] For further discussion of the relationship between AR specifications, disaggregation, and causality, see Florens, Mouchart, and Rolin (1993).

vectors reduce to two variables and (2) the case where the two vectors embody all the variables considered in the analysis. These conditions are considerably more tractable from the point of view of implementing tests and take into account all the constraints typically imposed by an ARMA specification.

The main purpose of this paper is to derive simplified conditions, similar to those in BDR, for the general case where the two vectors of variables do not necessarily embody all the variables considered in the analysis. By the same occasion, we also show that two of the theorems already provided in the latter paper (Theorems 2 and 3) are valid under weaker regularity conditions, where an invertibility condition associated with the moving average operator has been suppressed.

The organization and main results of the paper are as follows. In Section 2, we first report slightly less restrictive versions of two theorems obtained in BDR, and then provide generalizations of these results by giving new and relatively simple necessary and sufficient conditions for noncausality between subvectors inside of a system of arbitrary dimension. Section 3 discusses the empirical testing of the conditions derived. In particular, we point out that the multilinear form of the restrictions may lead to nonregular asymptotic distributions (for certain points of the parameter space), a difficulty similar to the one observed in BDR. Such problems can be dealt with by considering subsets of the restrictions (necessary conditions) or by using sequential test strategies. Finally, in Section 4, we present an application of our results to data on two Canadian monetary aggregates (M1 and M2), income, and interest rates. Contrary to the results in BDR, which did not use an interest rate variable and suggested unidirectional causality from the monetary aggregates to income, the analysis presented here strongly suggests the presence of bidirectional causality (feedback) between money and income. We also find feedback between money and interest rates.

2. Causality in multivariate ARMA processes

Let $\{\mathbf{X}_t : t \in \mathbb{Z}\}$ be a $n \times 1$ multivariate stochastic process on the integers $\mathbb{Z}$, with finite second moments, and write $\mathbf{X}_t = (\mathbf{X}'_{1t}, \mathbf{X}'_{2t}, \mathbf{X}'_{3t})'$, where $\mathbf{X}_{it}$ is a vector of dimension $n_i \times 1$ $(n_1 \geqslant 1,\ n_2 \geqslant 1,\ n_3 \geqslant 0)$, $\mathbf{X}_{1t} = (X_{1t}, \dots, X_{n_1 t})'$, and $\mathbf{X}_{2t} = (X_{n_1+1,t}, \dots, X_{n_1+n_2,t})'$; when $n_3 = 0$, we set $\mathbf{X}_t = (\mathbf{X}'_{1t}, \mathbf{X}'_{2t})'$ and $(\mathbf{X}'_{2t}, \mathbf{X}'_{3t})' = \mathbf{X}_{2t}$. The vectors $\mathbf{X}_{1t}$ and $\mathbf{X}_{2t}$ contain variables of interest between which we want to study causality relationships, while $\mathbf{X}_{3t}$ is a (possibly empty) vector of auxiliary variables that are also used as information to obtain forecasts. Further, let $\mathrm{I}_{\mathbf{X}}(t)$ be the Hilbert space generated by the components of $\mathbf{X}_\tau$ for $\tau \leqslant t$ (with covariance as the inner product, and let $\mathrm{I}_{\mathbf{X}_2^\bullet}(t)$ be the closed subspace of $\mathrm{I}_{\mathbf{X}}(t)$ generated by the components of $(\mathbf{X}'_{2\tau}, \mathbf{X}'_{3\tau})$ for $\tau \leqslant t$. The sets $\mathrm{I}_{\mathbf{X}}(t)$ and $\mathrm{I}_{\mathbf{X}_2^\bullet}(t)$ may be described as 'information sets'. For any subspace I_{t-1} of $\mathrm{I}_{\mathbf{X}}(t-1)$ and for $n_1 + 1 \leqslant i \leqslant n_1 + n_2$, we denote $P(X_{it} \mid I_{t-1})$ the affine projection of X_{it} on I_{t-1} (i.e., the best linear prediction of X_{it} based on the variables in I_{t-1} and a constant variable), $\varepsilon_{it}(X_{it} \mid I_{t-1}) = X_{it} - P(X_{it} \mid I_{t-1})$, $\sigma^2(X_{it} \mid I_{t-1}) = E(\varepsilon_{it}(X_{it} \mid I_{t-1})^2)$ the variance of $\varepsilon_{it}(X_{it} \mid I_{t-1})$ and $P(X_{2t} \mid I_{t-1}) = (P(X_{n_1+1,t} \mid I_{t-1}), \ldots, P(X_{n_1+n_2,t} \mid I_{t-1}))'$. Noncausality from $\mathbf{X}_1$ to $\mathbf{X}_2$ is then defined as follows.

Definition 1. The vector $\mathbf{X}_1$ *does not cause* $\mathbf{X}_2$ *(given* $\mathbf{X}_3$*), denoted* $\mathbf{X}_1 \nrightarrow \mathbf{X}_2 \mid \mathbf{X}_3$, *if each component of the error vector* $\mathbf{X}_{2t} - P(\mathbf{X}_{2t} \mid \mathrm{I}_{\mathbf{X}_2 \cdot}(t-1))$ *is orthogonal to* $\mathrm{I}_{\mathbf{X}}(t-1)$, *for all t.*

Note that this definition is equivalent to stating (as in BDR) that $\sigma^2(X_{it} \mid \mathrm{I}_{\mathbf{X}}(t-1)) = \sigma^2(X_{it} \mid \mathrm{I}_{\mathbf{X}_2 \cdot}(t-1))$, $i = n_1 + 1, \ldots, n_1 + n_2$, for all t, i.e., the mean squared linear prediction error of $\mathbf{X}_{2t}$ based on the past of $\mathbf{X}_2$ and $\mathbf{X}_3$ is not improved by taking into account the past of $\mathbf{X}_1$; for further discussion of causality in terms of orthogonality conditions, see Florens and Mouchart (1985). Of course, the qualification 'given $\mathbf{X}_3$' is irrelevant when $\mathbf{X}_{3t}$ is empty; for such cases, we will write $\mathbf{X}_1 \nrightarrow \mathbf{X}_2$. In this paper, '$\mathbf{X}_1$ does not cause $\mathbf{X}_2$' will always mean that $\mathbf{X}_1$ does not cause $\mathbf{X}_2$ given the other variables in $\mathbf{X}_3$ (if any).

Suppose now that $\{\mathbf{X}_t\}$ is a n-dimensional stationary and invertible ARMA (p, q) process with zero mean and regular innovations, i.e.,

$$\mathbf{\Phi}(B)\mathbf{X}_t = \mathbf{\Theta}(B)\mathbf{a}_t, \tag{2.1}$$

where B is the usual backshift operator, $\mathbf{\Phi}(z) = \mathrm{I}_n - \mathbf{\Phi}_1 z - \cdots - \mathbf{\Phi}_p z^p$, $\mathbf{\Theta}(z) = \mathrm{I}_n - \mathbf{\Theta}_1 z - \cdots - \mathbf{\Theta}_q z^q$ are matrix polynomials, with $\mathbf{I}_n$ denoting the identity matrix of order n, and $\{\mathbf{a}_t : t \in \mathbb{Z}\}$ is a white noise process with nonsingular convariance matrix $\mathbf{V}$. 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. There is no loss of generality in the assumption that $E(\mathbf{X}_t) = \mathbf{0}$.

Let us consider first the case where $\mathbf{X}_t$ is partitioned into two subvectors: $\mathbf{X}_t = (\mathbf{X}'_{1t}, \mathbf{X}'_{2t})'$ where $\mathbf{X}_{it}$ has dimension $n_i \times 1$, $i = 1, 2$, and $n_1 + n_2 = n$, so that the model (2.1) can be rewritten as

$$\begin{bmatrix} \mathbf{\Phi}_{11}(B) & \mathbf{\Phi}_{12}(B) \\ \mathbf{\Phi}_{21}(B) & \mathbf{\Phi}_{22}(B) \end{bmatrix} \begin{bmatrix} \mathbf{X}_{1t} \\ \mathbf{X}_{2t} \end{bmatrix} = \begin{bmatrix} \mathbf{\Theta}_{11}(B) & \mathbf{\Theta}_{12}(B) \\ \mathbf{\Theta}_{21}(B) & \mathbf{\Theta}_{22}(B) \end{bmatrix} \begin{bmatrix} \mathbf{a}_{1t} \\ \mathbf{a}_{2t} \end{bmatrix}, \tag{2.2}$$

where $\mathbf{\Phi}_{ij}(z)$ and $\mathbf{\Theta}_{ij}(z)$ are $n_i \times n_j$ matrix polynomials, $i, j = 1, 2$. In other words, $\mathbf{X}_{1t}$ and $\mathbf{X}_{2t}$ include all the variables considered in the information set $\mathrm{I}_{\mathbf{X}}(t)$. In this case, necessary and sufficient conditions for noncausality between $\mathbf{X}_1$ and $\mathbf{X}_2$ are given by the two following theorems, which improve two earlier results given in BDR (Theorems 2 and 3). The symbol det(.) refers to the determinant of a matrix.

Theorem 1. Suppose that the stationary ARMA process (2.2) is invertible. Then $\mathbf{X}_1$ *does not cause* $\mathbf{X}_2$ *if and only if*

$$\mathbf{\Phi}_{21}(z) - \mathbf{\Theta}_{21}(z)\mathbf{\Theta}_{11}(z)^{-1}\mathbf{\Phi}_{11}(z) = \mathbf{0} \quad \textit{for} \quad |z| < \delta, \tag{2.3}$$

where δ *is some positive constant.*

Theorem 2. Suppose that the stationary ARMA process (2.2) is invertible. Then $\mathbf{X}_1$ *does not cause* $\mathbf{X}_2$ *if and only if*

$$\det\begin{bmatrix} \mathbf{\Phi}_{11}^{\cdot j}(z) & \mathbf{\Theta}_{11}(z) \\ \phi_{n_1+i,j}(z) & \mathbf{\Theta}_{21}^{i\cdot}(z) \end{bmatrix} = 0, \qquad \forall z \in \mathbb{C}, \tag{2.4}$$

for $i = 1, \ldots, n_2$ *and* $j = 1, \ldots, n_1$, *where* $\mathbf{\Phi}_{11}^{\cdot j}(z)$ *is the* j*th column of* $\mathbf{\Phi}_{11}(z)$, $\mathbf{\Theta}_{21}^{i\cdot}(z)$ *is the* i*th row of* $\mathbf{\Theta}_{21}(z)$, *and* $\phi_{n_1+i,j}(z)$ *is the* (i,j)*-element of* $\mathbf{\Phi}_{21}(z)$.

The proofs of the theorems appear in the Appendix. Theorems 1 and 2 provide extensions of Theorems 2 and 3 in BDR because the assumption that $\det(\mathbf{\Theta}_{11}(z)) \neq 0$ for $|z| \leqslant 1$ has now been dropped. Note that $|z| < \delta$ is added to (2.3) only to ensure that the power series on the right-hand side converges. The latter converges provided $\det(\mathbf{\Theta}_{11}(z)) \neq 0$ for $|z| < \delta$ and some $\delta > 0$. Since $\mathbf{\Theta}_{11}(0) = \mathbf{I}_{n_1}$, it is clear that such a δ always exists. To see that the generalization provided by Theorem 1 above is substantial, consider the simple case where $n_1 = n_2 = 1$ with $\mathbf{\Theta}_{11}(z) = 1 - 1.1z$, $\mathbf{\Theta}_{12}(z) = -0.5z$, $\mathbf{\Theta}_{21}(z) = 1.1z$, and $\mathbf{\Theta}_{22}(z) = 1 + 0.5z$. Then we see easily that $\det(\mathbf{\Theta}(z)) = 1 - 0.6z$ so that the equation $\det(\mathbf{\Theta}(z)) = 0$ has all its roots outside the unit circle; further $\det(\mathbf{\Theta}(z)) = 1 - 1.1z \neq 0$ for $|z| < \delta = (1.1)^{-1} < 1$, but $\det(\mathbf{\Theta}_{11}(z)) = 0$ for $z = (1.1)^{-1}$. This example shows that $\det(\mathbf{\Theta}_{11}(z)) = 0$ may have a root inside the unit circle, while $\det(\mathbf{\Theta}(z)) = 0$ has all its roots outside, so that dropping the assumption $\det(\mathbf{\Theta}_{11}(z)) = 0$ for $|z| \leqslant 1$ provides a substantial extension of Theorem 3 in BDR. The appropriate noncausality restrictions on the parameters of $\mathbf{\Phi}(B)$ and $\mathbf{\Theta}(B)$ are obtained by setting the coefficients of the power series (or polynomials) defined in (2.3) or (2.4) equal to zero. Theorem 1 gives a matrix generalization of the condition given by Kang (1981) for noncausality between $\mathbf{X}_1$ and $\mathbf{X}_2$ when $n_1 = n_2 = 1$: $\mathbf{\Theta}_{11}(z)\mathbf{\Phi}_{21}(z) - \mathbf{\Theta}_{21}(z)\mathbf{\Phi}_{11}(z) = 0$. Theorem 2, on the other hand, yields a more convenient necessary and sufficient condition to be used in applications, because it only involves a finite number of polynomials.

Consider now the case where $\mathbf{X}_{1t}$ and $\mathbf{X}_{2t}$ do not include all the variables considered in the analysis, but a third vector of 'auxiliary variables' $\mathbf{X}_{3t}$ is also used to forecast $\mathbf{X}_{2t}$. In this case, we suppose that $\mathbf{X}_t$ is partitioned into three subvectors $\mathbf{X}_t = (\mathbf{X}'_{1t}, \mathbf{X}'_{2t}, \mathbf{X}'_{3t})'$ where $\mathbf{X}_{it}$ has dimension $n_i \times 1$, $i = 1, 2, 3$, with $n_1 + n_2 + n_3 = n$ and $n_i \geqslant 1$, $i = 1, 2, 3$. In this case, (2.1) can be written as

$$\begin{bmatrix} \mathbf{\Phi}_{11}(B) & \mathbf{\Phi}_{12}(B) & \mathbf{\Phi}_{13}(B) \\ \mathbf{\Phi}_{21}(B) & \mathbf{\Phi}_{22}(B) & \mathbf{\Phi}_{23}(B) \\ \mathbf{\Phi}_{31}(B) & \mathbf{\Phi}_{32}(B) & \mathbf{\Phi}_{33}(B) \end{bmatrix} \begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{bmatrix} = \begin{bmatrix} \mathbf{\Theta}_{11}(B) & \mathbf{\Theta}_{12}(B) & \mathbf{\Theta}_{13}(B) \\ \mathbf{\Theta}_{21}(B) & \mathbf{\Theta}_{22}(B) & \mathbf{\Theta}_{23}(B) \\ \mathbf{\Theta}_{31}(B) & \mathbf{\Theta}_{32}(B) & \mathbf{\Theta}_{33}(B) \end{bmatrix} \begin{bmatrix} a_{1t} \\ a_{2t} \\ a_{3t} \end{bmatrix}, \tag{2.5}$$

where $\mathbf{\Phi}_{ij}(z)$ and $\mathbf{\Theta}_{ij}(z)$ are $n_i \times n_j$ matrices, $i, j = 1, 2, 3$. Let also

$$\bar{\mathbf{\Theta}}_{ij}(z) = \mathbf{\Theta}_{ij}(z) - \mathbf{\Theta}_{i3}(z)\mathbf{\Theta}_{33}(z)^{-1}\mathbf{\Theta}_{3j}(z), \qquad i, j = 1, 2, \tag{2.6}$$

$$\bar{\mathbf{\Phi}}_{ij}(z) = \mathbf{\Phi}_{ij}(z) - \mathbf{\Theta}_{i3}(z)\mathbf{\Theta}_{33}(z)^{-1}\mathbf{\Phi}_{3j}(z), \qquad i, j = 1, 2, \tag{2.7}$$

$$\Delta_2(z) = \bar{\mathbf{\Theta}}_{22}(z) - \bar{\mathbf{\Theta}}_{21}(z)\bar{\mathbf{\Theta}}_{11}(z)^{-1}\bar{\mathbf{\Theta}}_{12}(z). \tag{2.8}$$

Clearly the matrices $\bar{\mathbf{\Theta}}_{ij}(z)$ and $\bar{\mathbf{\Phi}}_{ij}(z)$ are $n_i \times n_j$ matrices, while $\Delta_2(z)$ has dimension $n_2 \times n_2$. By the argument used in the proofs of Theorems 1 and 2, it is also easy to see that $\mathbf{\Theta}_{33}(z)$ and $\bar{\mathbf{\Theta}}_{11}(z)$ are invertible when $|z| < \delta$, for some $\delta > 0$. We now give a generalization of Theorem 1 for the case where a third set of variables $\mathbf{X}_3$ is present in the analysis and is used to compute forecasts.

Theorem 3. Suppose that the stationary ARMA process (2.5) is invertible. Then $\mathbf{X}_1$ *does not cause* $\mathbf{X}_2$ *(given* $\mathbf{X}_3$*) if and only if*

$$\bar{\mathbf{\Phi}}_{21}(z) - \bar{\mathbf{\Theta}}_{21}(z)\bar{\mathbf{\Theta}}_{11}(z)^{-1}\bar{\mathbf{\Phi}}_{11}(z) = \mathbf{0} \quad \textit{for} \quad |z| < \delta, \tag{2.9}$$

where δ *is some positive constant.*

By using the convention $\bar{\mathbf{\Phi}}_{ij}(z) = \mathbf{\Phi}_{ij}(z)$ and $\bar{\mathbf{\Theta}}_{ij}(z) = \mathbf{\Theta}_{ij}(z)$ when $n_3 = 0$ (no vector of auxiliary variables), we see clearly that the condition (2.9) includes (2.3) as a special case. Theorem 3 however may lead one to consider an infinity of (possibly redundant) restrictions. In the following theorem, we generalize Theorem 2 and provide a more convenient characterization of noncausality between $\mathbf{X}_1$ and $\mathbf{X}_2$.

Theorem 4. Suppose that the stationary ARMA process (2.5) is invertible. Then $\mathbf{X}_1$ *does not cause* $\mathbf{X}_2$ *(given* $\mathbf{X}_3$*) if and only if*

$$\Gamma_{ij}(z) = \det \begin{bmatrix} \mathbf{\Phi}_{11}^{\bullet j}(z) & \mathbf{\Theta}_{11}(z) & \mathbf{\Theta}_{13}(z) \\ \phi_{n_1+i,j}(z) & \mathbf{\Theta}_{21}^{i\bullet}(z) & \mathbf{\Theta}_{23}^{i\bullet}(z) \\ \mathbf{\Phi}_{31}^{\bullet j}(z) & \mathbf{\Theta}_{31}(z) & \mathbf{\Theta}_{33}(z) \end{bmatrix} = 0, \qquad \forall z \in \mathbb{C},$$

for $i = 1, \ldots, n_2$ *and* $j = 1, \ldots, n_1$, *where* $\mathbf{\Phi}_{lk}^{\bullet j}(z)$ *is the* j*th column of the matrix* $\mathbf{\Phi}_{lk}(z)$, $\mathbf{\Theta}_{lk}^{i\bullet}(z)$ *is the* i*th row of the matrix* $\mathbf{\Theta}_{lk}(z)$, *and* $\phi_{n_1+i,j}(z)$ *is the* (i,j)*-element of* $\mathbf{\Phi}_{21}(z)$.

Compared with Theorem 1 in BDR, Theorem 4 above reduces the dimension of the determinants to be evaluated. For example, consider the case where $\mathbf{X}_t = (X_{1t}, X_{2t}, X_{3t}, X_{4t})'$ is a four-dimensional stationary and invertible ARMA process as defined in (2.1). Then, from Theorem 1 of BDR, X_1 does not cause $(X_2, X_3)'$ if and only if

$$\det \begin{bmatrix} \phi_{11}(z) & \theta_{11}(z) & \theta_{1i}(z) & \theta_{14}(z) \\ \phi_{21}(z) & \theta_{21}(z) & \theta_{2i}(z) & \theta_{24}(z) \\ \phi_{31}(z) & \theta_{31}(z) & \theta_{3i}(z) & \theta_{34}(z) \\ \phi_{41}(z) & \theta_{41}(z) & \theta_{4i}(z) & \theta_{44}(z) \end{bmatrix} = 0, \qquad i = 2, 3,$$

where $\phi_{ij}(z)$ and $\theta_{ij}(z)$ are the (i,j)-elements of $\mathbf{\Phi}(z)$ and $\mathbf{\Theta}(z)$ respectively. On the other hand, from Theorem 4 above, we see that X_1 does not cause $(X_2, X_3)'$ if and only if

$$\Gamma_{1i}(z) = \det \begin{bmatrix} \phi_{11}(z) & \theta_{11}(z) & \theta_{14}(z) \\ \phi_{1+i,1}(z) & \theta_{1+i,1}(z) & \theta_{1+i,4}(z) \\ \phi_{41}(z) & \theta_{41}(z) & \theta_{44}(z) \end{bmatrix} = 0, \qquad i = 1, 2.$$

The latter condition leads to the evaluation of 3×3 determinants, while the former requires the evaluation of 4×4 determinants. More generally, the determinants involved in Theorem 4 have dimension $n_1 + n_3 + 1$ instead of $n_1 + n_2 + n_3$. So to take a more extreme example, if $n_1 = n_3 = 1$ and $n_2 = 8$, the dimension is 3×3 instead of 10×10, a very sizable reduction in complexity.

In Theorems 1 to 4, we have assumed that the stationary ARMA process $\mathbf{X}_t$ is invertible, i.e., the polynomial $\det(\mathbf{\Theta}(z))$ has no roots on the unit circle ($|z| = 1$). Without loss of generality, we can assume that $\det(\mathbf{\Theta}(z)) \neq 0$ for $|z| \leqslant 1$ since for any stationary and invertible ARMA process, it is always possible to choose a representation such that $\det(\mathbf{\Theta}(z))$ has no roots inside the unit circle ($|z| < 1$); see Nsiri and Roy (1993, Sect. 3.1). However, for a noninvertible ARMA process, roots on the unit circle cannot be eliminated; see Hannan and Deistler (1988, Ch. 1). Thus the invertibility condition (which implies that the process admits an autoregressive representation) entails restrictions on the covariance structure of the process. Note that the above results can also be applied to an invertible

ARIMA(p, d, q) process, provided $\mathbf{X}_t$ is replaced by $(1 - B)^d \mathbf{X}_t$. The noncausality conditions given then apply to $(1 - B)^d \mathbf{X}_t$ instead of $\mathbf{X}_t$. It is important to note here that the invertibility assumption precludes the presence of cointegrating relationships among the components of $\mathbf{X}_t$; see Gouriéroux and Monfort (1990, Ch. XI). There is no incompatibility between an ARIMA specification and the presence of cointegrating relationships, but the latter imply that the polynomial $\det(\mathbf{\Theta}(z))$ associated with the representation $\mathbf{\Phi}(B)(1 - B)^d \mathbf{X}_t = \mathbf{\Theta}(B)\mathbf{a}_t$ has roots on the unit circle and so the process $\mathbf{X}_t$ is not invertible. Though it appears quite plausible that the characterizations in Theorems 1 to 4 remain valid when the polynomial $\det(\mathbf{\Theta}(z))$ has roots on the unit circle, Proposition 1 in BDR relies heavily on the invertibility assumption and important modifications appear required to extend the results to noninvertible processes. Furthermore, the distributional theory of the test statistics is likely to be more complicated in such cases. Both these problems go beyond the scope of the present paper.

3. Testing

The conditions given in Section 2 can be used in two different ways. First, given a theoretical (or estimated) ARMA(p, q) model that involves a number of parameter restrictions (such as zero restrictions), it is possible that the model satisfies exactly certain noncausality properties. Since it is not typically easy to see these properties when $q \geqslant 1$, conditions such as those in Theorems 2 and 4 provide a simple way of checking whether noncausality restrictions hold exactly in a given model. For an example of this type of application, the reader may consult the empirical section of BDR. Second, when the restrictions for a given noncausality property do not hold exactly, one is led to the problem of testing these restrictions. The general problem of testing conditions of the type considered in Theorems 2 and 4 has already been discussed in BDR (Section 5), so there is no need to describe it in detail here. Nevertheless, it will be useful to outline succinctly the approach followed. There are three main steps for assessing empirically causality properties in the context of multivariate ARMA modelling.

(1) Build a multivariate ARMA model for the series, for example by following the procedure of Tiao and Box (1981).
(2) Using the results of Section 2, derive the noncausality conditions and express them in terms of the autoregressive and moving average parameters of the estimated model. Denoting $\boldsymbol{\beta}$ the vector of all AR and MA parameters, the noncausality conditions lead to (possibly nonlinear) constraints on an $l \times 1$

subvector, $\boldsymbol{\beta}_1$ of $\boldsymbol{\beta}$. We will denote these restrictions by H_0: $R_j(\boldsymbol{\beta}_1) = 0$, $j = 1, \dots, k$, where $k \leqslant l$.

(3) If the restrictions do not hold exactly, choose a test criterion. The most convenient ones are typically the likelihood ratio (LR) and the Wald test statistics. The LR statistic for testing H_0 is $LR = 2[L_N(\hat{\boldsymbol{\beta}}) - L_N(\hat{\boldsymbol{\beta}}^{\circ})]$, where $L_N(\boldsymbol{\beta})$ is the log-likelihood function, N is the sample size, $\hat{\boldsymbol{\beta}}$ is the unrestricted maximum likelihood (ML) estimator of $\boldsymbol{\beta}$, and $\hat{\boldsymbol{\beta}}^{\circ}$ is the restricted estimator of $\boldsymbol{\beta}$ (under H_0). On the other hand, the Wald statistic for testing H_0 is $W = N\mathbf{R}(\hat{\boldsymbol{\beta}}_1)'[\mathbf{T}(\hat{\boldsymbol{\beta}}_1)\hat{\mathbf{V}}_1(\hat{\boldsymbol{\beta}})\mathbf{T}(\hat{\boldsymbol{\beta}}_1)']^{-1}\mathbf{R}(\hat{\boldsymbol{\beta}}_1)$, where $\mathbf{R}(\boldsymbol{\beta}_1) = [R_1(\boldsymbol{\beta}_1), \dots, R_k(\boldsymbol{\beta}_1)]'$, $\mathbf{T}(\boldsymbol{\beta}_1) = \partial\mathbf{R}(\boldsymbol{\beta}_1)/\partial\boldsymbol{\beta}_1'$, and $\hat{\mathbf{V}}_1(\hat{\boldsymbol{\beta}})$ is a consistent estimate of the asymptotic covariance matrix of $\sqrt{N}\,(\hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1)$. We assume that $\hat{\mathbf{V}}_1(\hat{\boldsymbol{\beta}})$ is invertible. Under usual regularity conditions, the asymptotic distribution of LR or W is chi-square with k degrees of freedom.

Note that specification procedures may impose causality restrictions by setting various coefficients to zero so that some causality testing is implicitly incorporated at the modelling stage. The formal causality tests performed after the specification stage of the ARMA modelling process consider only the restrictions which are not exactly imposed at the specification stage.

It is easy to see from Theorems 2 and 4 that the restrictions to be tested are either linear or multilinear, i.e., the functions $R_j(\boldsymbol{\beta}_1)$ which define H_0 are either linear in $\boldsymbol{\beta}_1$ or sums of products of the components of $\boldsymbol{\beta}_1$. Clearly, for the test criteria to follow $\chi^2(k)$ distributions asymptotically, it is important that $\mathbf{R}(\boldsymbol{\beta}_1) = \mathbf{0}$ does not include redundant restrictions; if it did, the matrix of the first derivatives of $\mathbf{R}(\boldsymbol{\beta}_1)$ would not have full row rank and the asymptotic covariance matrix of $\mathbf{R}(\hat{\boldsymbol{\beta}}_1)$ would be singular.

As already observed in BDR, it is important to note also that the restrictions (3.1) may lead to situations where these regularity conditions do not hold (at least for certain points of the parameter space). The nature of these problems and the possible approaches to deal with them [e.g., separate tests of simpler necessary (or sufficient) conditions, sequential testing] are completely analogous to those discussed in BDR. The reader may see the latter article for further discussion.

4. Causality tests between money, income, and interest rates

Causality relations between money and income, and money and interest rates have been much debated in the economic literature; see Sims (1972, 1980b), Feige and Pearce (1979), Hsiao (1979), Osborn (1984), and BDR. To illustrate the causality conditions and tests given above, we will now study causality relations between money and income, as well as between money and interest rates in Canada. The money and income data are those of Hsiao (1979). They consist of quarterly seasonally adjusted nominal GNP, M1, and M2 over the period 1955 to 1977 (92 observations). A listing of these data is available in the Appendix of Hsiao (1979). We also considered quarterly interest rates (INT) data on Canada Treasury bonds for the corresponding period. The latter series comes from a monthly series available in CANSIM (series B14007); monthly observations were aggregated into quarterly ones by taking the arithmetic mean over the corresponding months.

The natural logarithm of GNP, M1, and M2 was taken in order to stabilize variances, and all series were differenced to have stationary processes. Furthermore, the first differences of ln(GNP), ln(M1), and ln(M2) were multiplied by 100 to ensure that their sample variances be of the same order of magnitude as the one of INT. In the following, we will denote $y_t = 100(1-B)\ln(\mathrm{GNP}_t)$, $m_{1t} = 100(1-B)\ln(\mathrm{M1}_t)$, $m_{2t} = 100(1-B)\ln(\mathrm{M2}_t)$, $r_t = (1-B)\ \mathrm{INT}_t$. Using the approach of Tiao and Box (1981) and the SCA statistical package (see Liu and Hudak, 1986), a four-variable ARMA model was estimated. At the estimation stage, the full model was first estimated by a Gaussian maximum likelihood method (the 'exact' method available in SCA); then each parameter estimate smaller than one standard error in absolute value was set at zero. The modified model was reestimated until all the parameter estimates were greater than one standard error in absolute value. The final model is described in Table 1. It satisfies the diagnostic checks suggested by Tiao and Box (1981) to ensure model adequacy. We also checked that all the roots of $\det(\hat{\boldsymbol{\Theta}}(z))$ are outside the unit circle.

Writing $\mathbf{X}_t = (y_t, m_{1t}, m_{2t}, r_t)'$, $\mathbf{a}_t = (a_{yt}, a_{1t}, a_{2t}, a_{rt})'$, and $\boldsymbol{\Theta}_0 = (\theta_{0y}, \theta_{01}, \theta_{02}, \theta_{0r})'$, the model described in Table 1 can be represented by the equation

$$\boldsymbol{\Phi}(B)X_t = \boldsymbol{\Theta}_0 + \boldsymbol{\Theta}(B)a_t, \tag{4.1}$$

where $\boldsymbol{\Phi}(B) = [\phi_{ij}(B)]$ and $\boldsymbol{\Theta}(B) = [\theta_{ij}(B)]$ are 4×4 matrices; $\phi_{ij}^{(k)}$ and $\theta_{ij}^{(k)}$ will denote the coefficients of B^k in $\phi_{ij}(B)$ and $\theta_{ij}(B)$ respectively.

Let us first analyze causality between money and income, with the interest rate being the auxilliary variable. Using Theorem 4, we find that $y_t \nrightarrow (m_{1t}, m_{2t})' \mid r_t \Leftrightarrow \Gamma_{i1}(z) \equiv 0$, $i = 1, 2$, where

$$\Gamma_{i1}(z) = \det \begin{bmatrix} \phi_{11}(z) & \theta_{11}(z) & \theta_{14}(z) \\ \phi_{1+i,1}(z) & \theta_{1+i,1}(z) & \theta_{1+i,4}(z) \\ \phi_{41}(z) & \theta_{41}(z) & \theta_{44}(z) \end{bmatrix}.$$

Table 1

Estimated multivariate ARMA model (estimated standard errors are given in parentheses)

$$
\begin{bmatrix}
1 - \underset{(0.079)}{0.193B} - \underset{(0.078)}{0.168B^2} - \underset{(0.076)}{0.216B^3} & -\underset{(0.073)}{0.184B} - \underset{(0.078)}{0.293B^2} + \underset{(0.079)}{0.333B^2} & \underset{(0.112)}{0.133B^2} - \underset{(0.110)}{0.196B^3} & \underset{(0.156)}{0.231B} - \underset{(0.170)}{0.526B^2} \\
-\underset{(0.056)}{0.250B} & 1 - \underset{(0.091)}{0.115B^2} & -\underset{(0.103)}{0.294B} - \underset{(0.124)}{1.64B^2} & \underset{(0.144)}{0.713B} \\
-\underset{(0.029)}{0.197B} - \underset{(0.032)}{0.117B^3} & \underset{(0.062)}{0.137B} - \underset{(0.056)}{0.265B^2} + \underset{(0.060)}{0.141B^3} & 1 - \underset{(0.088)}{0.497B} - \underset{(0.100)}{0.189B^2} - \underset{(0.084)}{0.98B^3} & \underset{(0.095)}{0.429B} - \underset{(0.096)}{0.316B^2} - \underset{(0.087)}{0.168B^3} \\
-\underset{(0.043)}{0.075B} - \underset{(0.040)}{0.082B^2} & -\underset{(0.040)}{0.071B} - \underset{(0.040)}{0.090B^3} & \underset{(0.047)}{0.113B} & 1 - \underset{(0.092)}{0.239B} + \underset{(0.095)}{0.268B^2} + \underset{(0.095)}{0.153B^3}
\end{bmatrix}
\begin{bmatrix} y_t \\ m_{1t} \\ m_{2t} \\ r_t \end{bmatrix}
$$

$$
= \begin{bmatrix} \underset{(0.238)}{0.538} \\ -\underset{(0.080)}{0.063} \\ -\underset{(0.054)}{0.116} \\ -\underset{(0.128)}{0.297} \end{bmatrix}
+ \begin{bmatrix}
1 - \underset{(0.102)}{0.730B^4} & 0 & \underset{(0.116)}{0.796B^4} & -\underset{(0.193)}{0.628B^4} \\
0 & 1 - \underset{(0.047)}{0.877B^4} & 0 & 0 \\
\underset{(0.065)}{0.077B^4} & -\underset{(0.034)}{0.057B^4} & 1 - \underset{(0.052)}{0.992B^4} & -\underset{(0.127)}{0.274B^4} \\
-\underset{(0.064)}{0.329B^4} & \underset{(0.048)}{0.210B^4} & \underset{(0.074)}{0.210B^4} & 1
\end{bmatrix}
\begin{bmatrix} a_{yt} \\ a_{1t} \\ a_{2t} \\ a_{rt} \end{bmatrix}
$$

Error covariance matrix

$$
\begin{bmatrix}
0.786 & - & - & - \\
0.407 & 1.513 & - & - \\
0.428 & 0.607 & 0.794 & \\
-0.132 & -0.138 & -0.114 & 0.290
\end{bmatrix}
$$

From Table 1, we see that

$$\Gamma_{11}(z) = \det \begin{bmatrix} 1 - \phi_{11}^{(1)} z - \phi_{11}^{(2)} z^2 - \phi_{11}^{(3)} z^3 & 1 - \theta_{11}^{(4)} z^4 & -\theta_{14}^{(4)} z^4 \\ -\phi_{21}^{(1)} z & 0 & 0 \\ -\phi_{41}^{(1)} z - \phi_{41}^{(2)} z^2 & -\theta_{41}^{(4)} z^4 & 1 \end{bmatrix}$$

$$= \phi_{21}^{(1)} (z - \theta_{11}^{(4)} z^5 - \theta_{14}^{(4)} \theta_{41}^{(4)} z^9) \equiv 0 \Leftrightarrow \phi_{21}^{(1)} = 0. \tag{4.2}$$

Similarly, we have

$$\Gamma_{21}(z) = \det \begin{bmatrix} 1 - \phi_{11}^{(1)} z - \phi_{11}^{(2)} z^2 - \phi_{11}^{(3)} z^3 & 1 - \theta_{11}^{(4)} z^4 & -\theta_{14}^{(4)} z^4 \\ -\phi_{31}^{(1)} z \qquad -\phi_{31}^{(3)} z^3 & -\theta_{31}^{(4)} z^4 & -\theta_{34}^{(4)} z^4 \\ -\phi_{41}^{(1)} z - \phi_{41}^{(2)} z^2 & -\theta_{41}^{(4)} z^4 & 1 \end{bmatrix},$$

and using a symbolic manipulation program like MATHEMATICA (see Wolfram, 1991), the polynomial representation can be easily obtained. Here, we get

$$\Gamma_{21}(z) = \phi_{31}^{(1)} z + \phi_{31}^{(3)} z^3 - \theta_{31}^{(4)} z^4 + \sum_{k=5}^{11} c_k z^k,$$

where the c_k's are nonlinear functions of the parameters appearing in $\Gamma_{21}(z)$. It is rather tedious algebra to write the necessary and sufficient conditions for $\Gamma_{21}(z) \equiv 0$. However, it is immediately seen that

$$\Gamma_{21}(z) \equiv 0 \Rightarrow \phi_{31}^{(1)} = \phi_{31}^{(3)} = \theta_{31}^{(4)} = 0. \tag{4.3}$$

Putting together (4.2) and (4.3), we find that

$$\phi_{21}^{(1)} = \phi_{31}^{(1)} = \phi_{31}^{(3)} = \theta_{31}^{(4)} = 0 \tag{4.4}$$

are necessary conditions for y_t not to cause $(m_{1t}, m_{2t})'$. The Wald and LR statistics for testing (4.4) take the values 62.3 and 37.7 respectively (causality tests are summarized in Table 2). Since the asymptotic null distribution of these test statistics is $\chi^2(4)$, the hypothesis (4.4) is strongly rejected at about any conventional significance level and we conclude that y_t causes $(m_{1t}, m_{2t})'$. Without taking into account interest rates, BDR found that y_t does not cause $(m_{1t}, m_{2t})'$ from a trivariate ARMA model.

Table 2
Causality tests

Null hypothesis	Number of constrained parameters	First-order linear constraints	Wald stat.[a]	L.R. stat.[a]	D.f.
$y_t \nrightarrow (m_{1t}, m_{2t})' \mid r_t$	13	$\phi_{21}^{(1)} = \phi_{31}^{(1)} =$	62.3	37.7	4
		$\phi_{31}^{(3)} = \theta_{31}^{(4)} = 0$			
$(m_{1t}, m_{2t})' \nrightarrow y_t \mid r_t$	25	$\phi_{12}^{(1)} = \phi_{12}^{(2)} =$	77.2	102.1	6
		$\phi_{12}^{(3)} = \phi_{13}^{(2)} =$			
		$\phi_{13}^{(3)} = \theta_{13}^{(4)} = 0$			
$r_t \nrightarrow (m_{1t}, m_{2t})' \mid y_t$	14	$\phi_{24}^{(1)} = \phi_{34}^{(1)} =$	73.6	47.5	5
		$\phi_{34}^{(2)} = \phi_{34}^{(3)} =$			
		$\theta_{34}^{(4)} = 0$			
$(m_{1t}, m_{2t})' \nrightarrow r_t \mid y_t$	26	$\phi_{42}^{(1)} = \phi_{42}^{(3)} =$	56.5	41.3	5
		$\phi_{43}^{(1)} = \theta_{42}^{(4)} =$			
		$\theta_{43}^{(4)} = 0$			

[a] All p-values are smaller than 0.0005.

Conversely, for causality running from money to income, we have: $(m_{1t}, m_{2t})' \nrightarrow y_t \mid r_t \Leftrightarrow \Gamma_{1j}(z) \equiv 0$, $j = 1, 2$, where

$$\Gamma_{1j}(z) = \det\begin{bmatrix} \phi_{2,1+j}(z) & \theta_{22}(z) & \theta_{23}(z) & \theta_{24}(z) \\ \phi_{3,1+j}(z) & \theta_{32}(z) & \theta_{33}(z) & \theta_{34}(z) \\ \phi_{1,1+j}(z) & \theta_{12}(z) & \theta_{13}(z) & \theta_{14}(z) \\ \phi_{4,1+j}(z) & \theta_{42}(z) & \theta_{43}(z) & \theta_{44}(z) \end{bmatrix}$$

$$= \det\begin{bmatrix} \phi_{1,1+j}(z) & \theta_{12}(z) & \theta_{13}(z) & \theta_{14}(z) \\ \phi_{2,1+j}(z) & \theta_{22}(z) & \theta_{23}(z) & \theta_{24}(z) \\ \phi_{3,1+j}(z) & \theta_{32}(z) & \theta_{33}(z) & \theta_{34}(z) \\ \phi_{4,1+j}(z) & \theta_{42}(z) & \theta_{43}(z) & \theta_{44}(z) \end{bmatrix}$$

With the help of MATHEMATICA, we find

$$\Gamma_{11}(z) = -\phi_{12}^{(1)}z - \phi_{12}^{(2)}z^2 - \phi_{12}^{(3)}z^3 + \sum_{k=5}^{15} c_k z^k,$$

$$\Gamma_{12}(z) = -\phi_{13}^{(2)}z^2 - \phi_{13}^{(3)}z^3 - \theta_{13}^{(4)}z^4 + \sum_{k=5}^{15} d_k z^k,$$

where the c_k's and the d_k's are again nonlinear functions of the parameters appearing in $\Gamma_{11}(z)$ and $\Gamma_{12}(z)$ respectively. Therefore, a set of necessary restrictions for $(m_{1t}, m_{2t})'$ not to cause y_t is given by

$$\phi_{12}^{(1)} = \phi_{12}^{(2)} = \phi_{12}^{(3)} = \phi_{13}^{(2)} = \phi_{13}^{(3)} = \phi_{13}^{(4)} = 0. \tag{4.5}$$

The Wald and LR statistics for testing (4.5) are 77.2 and 102.1 respectively. Since the asymptotic null distribution of the test statistics is $\chi^2(6)$, the hypothesis (4.5) is strongly rejected and we conclude that $(m_{1t}, m_{2t})'$ causes y_t. A similar conclusion was reached in BDR with a trivariate model.

A causality analysis between money and interest rates was also performed when income is an auxilliary variable. The results are summarized in Table 2 and the details of the analysis are given in Boudjellaba, Dufour, and Roy (1992b).

The above results thus strongly suggest the presence of bidirectional causality (feedback) between money and income, when interest rates are also used to forecast, as well as between money and interests rates when income is also used to forecast. The first of these two conclusions appears to support the earlier finding of Sims (1980b) with American data.

Appendix: Proofs

Theorem 1. Since $\mathbf{\Theta}(z) = \mathbf{I}_n - \mathbf{\Theta}_1 z - \cdots - \mathbf{\Theta}_q z^q$, we have $\det(\mathbf{\Theta}_{11}(0)) = 1$ so that, by the continuity of the function $\det(\mathbf{\Theta}_{11}(z))$, $\det(\mathbf{\Theta}_{11}(z)) \neq 0$ for z sufficiently close to zero, say $|z| < \delta$ where $\delta > 0$. The rest of the proof is identical to the proof of Theorem 2 in BDR, with $|z| < \delta$ instead of $|z| \leqslant 1$.

Theorem 2. By the same argument as in Theorem 1, $\det(\mathbf{\Theta}_{11}(z)) \neq 0$ for $|z| < \delta$, where $\delta > 0$. The rest of the proof is identical to the proof of Theorem 3 in BDR, with $|z| < \delta$ instead of $|z| \leqslant 1$.

Theorem 3. The process (2.5) being invertible can be expressed as an (infinite) autoregressive process: $\mathbf{\Pi}(B)\mathbf{X}_t = \mathbf{a}_t$ where $\mathbf{\Pi}(B) = \mathbf{\Theta}(B)^{-1}\mathbf{\Phi}(B) = [\mathbf{\Pi}_{ij}(B)]_{i,j=1,2,3}$. From Proposition 1 of BDR, we know that $\mathbf{X}_1$ does not cause

$\mathbf{X}_2$ if and only if $\mathbf{\Pi}_{21}(z) = \mathbf{0}$, $\forall z \in \mathbb{C}$. We thus need to evaluate $\mathbf{\Pi}_{21}(z)$. Let us set (the argument z will be omitted for simplicity)

$$\mathbf{A}_{00} = \begin{bmatrix} \mathbf{\Theta}_{11} & \mathbf{\Theta}_{22} \\ \mathbf{\Theta}_{21} & \mathbf{\Theta}_{22} \end{bmatrix}, \quad \mathbf{A}_{03} = \begin{bmatrix} \mathbf{\Theta}_{13} \\ \mathbf{\Theta}_{23} \end{bmatrix}, \quad \mathbf{A}_{30} = [\mathbf{\Theta}_{31}, \mathbf{\Theta}_{32}], \quad \mathbf{A}_{33} = \mathbf{\Theta}_{33}.$$

Since $\mathbf{\Theta}(0) = \mathbf{I}_n \neq 0$ and using procedure (a) of Searle and Hausman (1970, p. 113) for partitioned inversion, we can write

$$\mathbf{\Theta}^{-1} = \begin{bmatrix} \mathbf{\Theta}^{11} & \mathbf{\Theta}^{12} & \mathbf{\Theta}^{13} \\ \mathbf{\Theta}^{21} & \mathbf{\Theta}^{22} & \mathbf{\Theta}^{23} \\ \mathbf{\Theta}^{31} & \mathbf{\Theta}^{32} & \mathbf{\Theta}^{33} \end{bmatrix} = \begin{bmatrix} A^{00} & A^{03} \\ A^{30} & A^{33} \end{bmatrix} \tag{A.1}$$

for $|z| < \delta$, where $\delta > 0$ is some positive constant, $\mathbf{A}^{00} = (A_{00} - A_{03}A_{33}^{-1}A_{30})^{-1}$, $\mathbf{A}^{03} = -A^{00}A_{03}A_{33}^{-1}$, $\mathbf{A}^{30} = -A_{33}^{-1}A_{30}A^{00}$, $\mathbf{A}^{33} = A_{33}^{-1} - A^{30}A_{03}A_{33}^{-1}$, and $\mathbf{\Theta}^{ij}$ is a $n_i \times n_j$ matrix for $i, j = 1, 2, 3$. Further, using the inverse of a partitioned matrix from procedure (b) of Searle and Hausman (1970, p. 114), $\mathbf{A}^{00}$ can be written

$$\mathbf{A}^{00} = \begin{bmatrix} \mathbf{\Theta}^{11} & \mathbf{\Theta}^{12} \\ \mathbf{\Theta}^{21} & \mathbf{\Theta}^{22} \end{bmatrix}, \tag{A.2}$$

where $\mathbf{\Theta}^{11} = \bar{\Theta}_{11}^{-1}(\mathbf{I}_{n_1} + \bar{\Theta}_{12}\Delta_2^{-1}\bar{\Theta}_{21}\bar{\Theta}_{11}^{-1})$, $\Theta^{12} = -\bar{\Theta}_{11}^{-1}\bar{\Theta}_{12}\Delta_2^{-1}$, $\Theta^{21} = -\Delta_2^{-1}\bar{\Theta}_{21}\bar{\Theta}_{11}^{-1}$ and $\mathbf{\Theta}^{22} = \Delta_2^{-1}$. Using (A.1) and (A.2), we can then evaluate easily the other components of $\mathbf{\Theta}^{-1}$:

$$\mathbf{A}^{03} = \begin{bmatrix} -\mathbf{\Theta}^{11}\mathbf{\Theta}_{13}\mathbf{\Theta}_{33}^{-1} - \mathbf{\Theta}^{12}\mathbf{\Theta}_{23}\mathbf{\Theta}_{33}^{-1} \\ -\mathbf{\Theta}^{21}\mathbf{\Theta}_{13}\mathbf{\Theta}_{33}^{-1} - \mathbf{\Theta}^{22}\mathbf{\Theta}_{23}\mathbf{\Theta}_{33}^{-1} \end{bmatrix} = \begin{bmatrix} \mathbf{\Theta}^{13} \\ \mathbf{\Theta}^{23} \end{bmatrix},$$

$$\begin{aligned} \mathbf{A}^{30} &= [-\mathbf{\Theta}_{33}^{-1}\mathbf{\Theta}_{31}\mathbf{\Theta}^{11} - \mathbf{\Theta}_{33}^{-1}\mathbf{\Theta}_{32}\mathbf{\Theta}^{21}, -\mathbf{\Theta}_{33}^{-1}\mathbf{\Theta}_{31}\mathbf{\Theta}^{12} - \mathbf{\Theta}_{33}^{-1}\mathbf{\Theta}_{32}\mathbf{\Theta}^{22}], \\ &= [\mathbf{\Theta}^{31}, \mathbf{\Theta}^{32}], \end{aligned}$$

$$\mathbf{A}^{33} = \mathbf{\Theta}_{33}^{-1} - A^{30}A_{03}\mathbf{\Theta}_{33}^{-1}.$$

Consequently, we can also compute $\mathbf{\Pi}_{21}$:

$$\begin{aligned} \mathbf{\Pi}_{21} &= \mathbf{\Theta}^{21}\mathbf{\Phi}_{11} + \mathbf{\Theta}^{22}\mathbf{\Phi}_{21} + \mathbf{\Theta}^{23}\mathbf{\Phi}_{31} \\ &= -\Delta_2^{-1}\bar{\mathbf{\Theta}}_{21}\bar{\mathbf{\Theta}}_{11}^{-1}\mathbf{\Phi}_{11} + \Delta_2^{-1}\mathbf{\Phi}_{21} \\ &\quad + \Delta_2^{-1}(\bar{\mathbf{\Theta}}_{21}\bar{\mathbf{\Theta}}_{11}^{-1}\mathbf{\Theta}_{13} - \mathbf{\Theta}_{23})\mathbf{\Theta}_{33}^{-1}\mathbf{\Phi}_{31} \\ &= \Delta_2^{-1}[\bar{\mathbf{\Phi}}_{21} - \bar{\mathbf{\Theta}}_{21}\bar{\mathbf{\Theta}}_{11}^{-1}\bar{\mathbf{\Phi}}_{11}]. \end{aligned}$$

Thus, since $\Delta_2(0) = \mathbf{I}_{n_2} \neq \mathbf{0}$, we have $\mathbf{\Pi}_{21}(z) = \mathbf{0}$ for $|z| < \delta \Leftrightarrow \bar{\mathbf{\Phi}}_{21} - \bar{\mathbf{\Theta}}_{21}\bar{\mathbf{\Theta}}_{11}^{-1}\bar{\mathbf{\Phi}}_{11} = \mathbf{0}$ for $|z| < \delta$. Q.E.D.

Theorem 4. The argument z will be omitted for simplicity. By Theorem 3, $\mathbf{X}_1$ does not cause $\mathbf{X}_2$ if and only if condition (2.9) holds, where δ is some positive constant and $\det(\mathbf{\Theta}_{33}) \neq 0$ for $|z| < \delta$. From now on, we shall assume that $|z| < \delta$. By an argument similar to the one used in the proof of Theorem 2 in BDR, we have:

$$\bar{\mathbf{\Phi}}_{21} - \bar{\mathbf{\Theta}}_{21}\bar{\mathbf{\Theta}}_{11}^{-1}\bar{\mathbf{\Phi}}_{11} \equiv \mathbf{0} \Leftrightarrow \det(\mathbf{D}_{ij}) \equiv 0, \qquad 1 \leqslant i \leqslant n_2, \quad 1 \leqslant j \leqslant n_1, \tag{A.3}$$

where $\mathbf{D}_{ij}$ is the $(n_1 + 1)\times(n_1 + 1)$ matrix defined by

$$D_{ij} = \begin{bmatrix} \bar{\mathbf{\Phi}}_{11}^{\cdot j} & \bar{\mathbf{\Theta}}_{11} \\ \bar{\phi}_{n_1+i,j} & \bar{\mathbf{\Theta}}_{21}^{i\cdot} \end{bmatrix},$$

$\bar{\phi}_{n_1+i,j}$ is the (i,j)-element of $\bar{\mathbf{\Phi}}_{21}$, and '$\equiv$' refers to an identity that holds for all $|z| < \delta$. Now, let

$$D_{ij}^{*} = \begin{bmatrix} \bar{\mathbf{\Phi}}_{11}^{\cdot j} & \bar{\mathbf{\Theta}}_{11} & \mathbf{0} \\ \bar{\phi}_{n_1+i,j} & \bar{\mathbf{\Theta}}_{21}^{i\cdot} & \mathbf{0} \\ \mathbf{\Phi}_{31}^{\cdot j} & \mathbf{\Theta}_{31} & \mathbf{\Theta}_{33} \end{bmatrix}, \qquad \bar{D}_{ij} = \begin{bmatrix} \mathbf{\Phi}_{11}^{\cdot j} & \mathbf{\Theta}_{11} & \mathbf{\Theta}_{13} \\ \phi_{n_1+i,j} & \mathbf{\Theta}_{21}^{i\cdot} & \mathbf{\Theta}_{23}^{i\cdot} \\ \mathbf{\Phi}_{31}^{\cdot j} & \mathbf{\Theta}_{31} & \mathbf{\Theta}_{33} \end{bmatrix},$$

and let $\alpha_i' = (\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{in_2})$ be the ith row of $\mathbf{\Theta}_{23}\mathbf{\Theta}_{33}^{-1}$, $1 \leqslant i \leqslant n_2$. Clearly $\det(\mathbf{D}_{ij}^{*}) = \det(\mathbf{D}_{ij})\det(\mathbf{\Theta}_{33})$. Further, since $\bar{\mathbf{\Phi}}_{11}^{\cdot j} = \mathbf{\Phi}_{11}^{\cdot j} - \mathbf{\Theta}_{13}\mathbf{\Theta}_{33}^{-1}\mathbf{\Phi}_{31}^{\cdot j}$, $\bar{\mathbf{\Theta}}_{11} = \mathbf{\Theta}_{11} - \mathbf{\Theta}_{13}\mathbf{\Theta}_{33}^{-1}\mathbf{\Theta}_{31}$, $\bar{\mathbf{\Phi}}_{21}^{\cdot j} = \mathbf{\Phi}_{21}^{\cdot j} - \mathbf{\Theta}_{23}\mathbf{\Theta}_{33}^{-1}\mathbf{\Phi}_{31}^{\cdot j}$, $\bar{\mathbf{\Theta}}_{21} = \mathbf{\Theta}_{21} - \mathbf{\Theta}_{13}\mathbf{\Theta}_{33}^{-1}\mathbf{\Theta}_{31}$, $\mathbf{\Theta}_{13} = \mathbf{\Theta}_{13}\mathbf{\Theta}_{33}^{-1}\mathbf{\Theta}_{33}$, and $\mathbf{\Theta}_{23} = \mathbf{\Theta}_{23}\mathbf{\Theta}_{33}^{-1}\mathbf{\Theta}_{33}$, the matrix $\bar{\mathbf{D}}_{ij}$, can be obtained from $\mathbf{D}_{ij}^{*}$ by adding $\mathbf{\Theta}_{13}\mathbf{\Theta}_{33}^{-1}[\mathbf{\Phi}_{31}^{\cdot j}, \mathbf{\Theta}_{31}, \mathbf{\Theta}_{33}]$ to the first n_1 rows of $\mathbf{D}_{ij}^{*}$, and $\alpha_i'[\mathbf{\Phi}_{31}^{\cdot j}, \mathbf{\Theta}_{31}, \mathbf{\Theta}_{33}]$ to the $(n_1 + 1)$th row $[\bar{\phi}_{n_1+i,j}, \bar{\mathbf{\Theta}}_{21}^{i\cdot}, \mathbf{0}]$. Thus the rows of $\bar{\mathbf{D}}_{ij}$ are obtained by adding to the first $n_1 + 1$ rows of $\mathbf{D}_{ij}^{*}$ linear combinations of other rows of $\mathbf{D}_{ij}^{*}$, so that the determinants of $\bar{\mathbf{D}}_{ij}$ and $\mathbf{D}_{ij}^{*}$ are identical: $\Gamma_{ij} = \det(\bar{\mathbf{D}}_{ij}) = \det(\mathbf{D}_{ij}^{*})$. Since $\det(\mathbf{\Theta}_{33}) \neq 0$, we can conclude from condition (2.9) and (A.3) that $\mathbf{X}_1$ does not cause $\mathbf{X}_2$ if and only if $\Gamma_{ij} = 0$, for $|z| < \delta$, $1 \leqslant i \leqslant n_2$, $1 \leqslant j \leqslant n_1$. Further, since $\Gamma_{ij} = \Gamma_{ij}(z)$ is a polynomial in z ($1 \leqslant i \leqslant n_2$, $1 \leqslant j \leqslant n_1$), i.e., $\Gamma_{ij}(z)$ is a linear combination of a **finite** number of powers of z, the latter condition is equivalent to $\Gamma_{ij} = 0, \forall z \in \mathbb{C}$, for $1 \leqslant i \leqslant n_2$, $1 \leqslant j \leqslant n_1$. Q.E.D.

References

Boudjellaba, H., 1988, Tests de causalité dans les modèles autoregressifs et moyennes mobiles multivariés, Unpublished Ph.D. thesis (Université de Montréal, Département d'Informatique et de Recherche Opèrationnelle, Montréal).

Boudjellaba, H., J.-M. Dufour, and R. Roy, 1992a, Testing causality between two vectors in multivariate ARMA models, Journal of the American Statistical Association 87, 1082–1090.

Boudjellaba, H., J.-M. Dufour, and R. Roy, 1992b, Simplified conditions for non-causality between vectors in multivariate ARMA models, Technical report 4092 (Université de Montréal, CRDE, Montréal).

Eberts, R.W. and B.M. Steece, 1984, A test for Granger-causality in a multivariate ARMA model, Empirical Economics 9, 51–58.

Feige, E.L. and D.K. Pearce, 1979, The causal relationship between money and income: Some caveats for time series analysis, Review of Economics and Statistics 61, 521–533.

Florens, J.-P. and M. Mouchart, 1985, A linear theory for noncausality, Econometrica 53, 157–175.

Florens, J.-P., M. Mouchart, and J.-M. Rolin, 1993, Noncausality and marginal Markov processes, Econometric Theory 9, 222–240.

Geweke, J., 1984, Inference and causality in economic time series models, in: Z. Griliches and M. Intriligator, eds., Handbook of econometrics, Vol. II (North-Holland, Amsterdam) 1102–1144.

Gouriéroux, C. and A. Monfort, 1990, Séries temporelles et modèles dynamiques (Economica, Paris).

Granger, C.W.J., 1969, Investigating causal relations by econometric models and cross-spectral methods, Econometrica 37, 424–438.

Hannan, E.J. and M. Deistler, 1988, The statistical theory of linear systems (Wiley, New York, NY).

Hsiao, C., 1979, Autoregressive modelling of Canadian money and income data, Journal of the American Statistical Association 74, 553–560.

Hsiao, C., 1982, Autoregressive modelling and causal orderings 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 L. Weiss, 1985, Money, real interest rates and output: A reinterpretation of postwar U.S. data, Econometrica 53, 129–156.

Liu, L. and G.B. Hudak, 1986, The SCA statistical system: Reference manual for forecasting and time series analysis (Scientific Computing Associates, DeKalb, IL).

Lütkepohl, H., 1991, Introduction to multiple time series analysis (Springer-Verlag, Berlin).

Newbold, P., 1982, Causality testing in economics, in: O.D. Anderson, ed., Time series analysis: Theory and practice 1 (North-Holland, Amsterdam) 701–716.

Newbold, P. and M. Hotopp, 1986, Testing causality using efficiently parametrized vector ARMA models, Applied Mathematics and Computation 20, 329–348.

Nsiri, S. and R. Roy, 1993, On the invertibility of multivariate linear processes, Journal of Time Series Analysis 14, 305–316.

Osborn, D.R., 1984, Causality testing and its implications for dynamic econometric models, The Economic Journal 94, 82–95.

Pierce, D.A. and L.D. Haugh, 1977, Causality in temrporal systems: Characterizations and survey, Journal of Econometrics 5, 265–293.

Searle, S.R. and W.H. Hausman, 1970, Matrix algebra for business and economics (Wiley, New York, NY).

Sims, C., 1972, Money, income and causality, American Economic Review 62, 540–552.

Sims, C., 1980a, Macroeconomics and reality, Econometrica 48, 1–48.

Sims, C., 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 model, Empirical Economics 14, 257–271.

Tiao, G.C. and G.E.P. Box, 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.

Wolfram, S., 1991, Mathematica, a system for doing mathematics by computer, 2nd ed. (Addison-Wesley, Redwood City, CA).