Short Run and Long Run Causality in Time Series: Theory

Jean-Marie Dufour; Eric Renault

Econometrica, Volume 66, Issue 5 (Sep., 1998), 1099-1125.

Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28199809%2966%3A5%3C1099%3ASRALRC%3E2.0.CO%3B2-9

Econometrica, Vol. 66, No. 5 (September, 1998), 1099–1125

SHORT RUN AND LONG RUN CAUSALITY IN TIME SERIES: THEORY

BY JEAN-MARIE DUFOUR AND ERIC RENAULT[1]

Causality in the sense of Granger is typically defined in terms of predictibility of a vector of variables one period ahead. Recently, Lütkepohl (1993) proposed to define noncausality between two variables in terms of nonpredictibility at any number of periods ahead. When more than two vectors are considered (i.e., when the information set contains auxiliary variables), these two notions are not equivalent. In this paper, we first generalize the notion of causality by considering causality at a given (arbitrary) horizon h. Then we derive necessary and sufficient conditions for noncausality between vectors of variables (inside a larger vector) up to any given horizon h, where h can be infinite. In particular, for general possibly nonstationary processes with finite second moments, relatively simple *exhaustivity* and *separation* conditions, which are sufficient for noncausality at all horizons, are provided. To deal with cases where such conditions do not apply, we consider a more specific, although still very wide, class of vector autoregressive processes (possibly of infinite order, stationary or nonstationary), which include multivariate ARIMA processes, and we derive general parametric characterizations of noncausality at various horizons for this class (including a *causality chain* characterization). We also observe that the coefficients of lagged variables in forecasts at various horizons $h \geq 1$ can be interpreted as *generalized impulse response coefficients* which yield a complete picture of linear causality properties, in contrast with usual response coefficients which can be quite misleading in this respect.

KEYWORDS: Causality, time series, long run, causality chain, vector autoregression, VAR, VARMA, impulse response, prediction.

1. INTRODUCTION

THE CONCEPT OF CAUSALITY INTRODUCED by Wiener (1956) and Granger (1969) is now a basic notion for studying dynamic relationships between time series. The literature on this topic is considerable; see, for example, the reviews of Pierce and Haugh (1977), Newbold (1982), Geweke (1984), Gouriéroux and Monfort (1990, Chapter X), and Lütkepohl (1991). The original definition of Granger (1969), which is used or adapted by most authors on this topic, refers to the predictibility of a variable $X(t)$, where t is an integer, from its own past, the one of another variable $Y(t)$, and possibly a vector $Z(t)$ of auxiliary variables, *one period ahead*: more precisely, we say that Y causes X in the sense of Granger if the observation of Y up to time t ($Y(\tau)$: $\tau \leq t$) can help one to

[1] This work was supported by the Social Sciences and Humanities Research Council of Canada, the Natural Sciences and Engineering Research Council of Canada, and the Government of Québec (Fonds FCAR and Ministre des relations internationales). The authors thank John Galbraith, Eric Ghysels, Helmut Lütkepohl, Michael Mouchart, Saïd Nsiri, Franz Palm, Pierre Perron, Roch Roy, Peter Schotmann, Lanh Tran, David Tessier, two anonymous referees, and a co-editor for several helpful comments and discussions.

predict $X(t+1)$ when the corresponding observations on X and Z are available $(X(\tau), Z(\tau)\colon \tau \leq t)$; a more formal definition will be given below.

Recently, however, Lütkepohl (1993) has noted that, for multivariate models where a vector of auxiliary variables Z is used in addition to the variables of interest X and Y, it is possible that Y does not cause X in this sense, but can still help to predict X *several periods ahead*; on this, see also Sims (1980) and Renault and Szafarz (1991). For example, the values $Y(\tau)$ up to time t may help to predict $X(t+2)$, even though they are useless to predict $X(t+1)$, because Y may help to predict Z one period ahead, which in turn influences X at a subsequent period. It is clear that studying such indirect effects can have a great interest for analyzing the relationships between time series. In particular, one can distinguish in this way properties of "short-run (non)causality" and "long-run (non)causality."

To the best of our knowledge, these indirect effects and associated noncausality properties have not yet been extensively studied in the literature. On one hand, causality at a given horizon h involves forecasts at horizon h which may depend in a complex way on autoregressive coefficients: for $h \geq 2$, the absence of lagged values of a variable from these forecasts does not generally reduce to zero restrictions on these coefficients. In this respect, so-called *impulse response coefficients* may be better descriptions of lagged causality relationships, but again causality studies "à la Sims" based on innovation accounting are not sufficient to capture all (linear) indirect effects of Y on X. Hsiao (1982) was perhaps the first author to address formally this issue by introducing indirect causality relationships, spurious causality concepts, ... In this respect, the present paper is a continuation of Hsiao's research agenda since we will propose a systematic study and characterization of indirect effects and associated lagged causality relationships. We will observe in particular that Hsiao's definitions do not capture all the effects of interest in the general case where more than one auxiliary variable Z appear in the system.

The paper is organized as follows. In Section 2, we define more general notions of causality that will allow us to study the issues of interest: causality at a given horizon h, where h is a positive integer, and causality up to any given horizon h, where h can be infinite $(1 \leq h \leq \infty)$. These definitions are based on the concept of projection (linear causality), do not require stationarity of the processes considered, and for the horizon one $(h = 1)$ include as a special case the usual definition of causality in the sense of Granger (1969). We can study in this way "short-run causality" (h small) and "long-run causality" (h large) properties. Note "short-run" and "long-run" refer to *forecast horizons* defined with respect to a given point in time, not the role played by past observations which may be more or less close to that point. Then we present several general results on causality up to any given horizon. In particular, we give a *component-wise characterization* of causality properties, which allows a reduction of causality between random vectors to causality between scalar random variables (the components of those vectors), and general sufficient conditions under which noncausality at horizon one is equivalent to noncausality at all horizons. We show this equivalence obtains in two important cases: first when the vectors X and Y contain all the variables considered in the analysis (*exhaustivity condition*), and secondly when all the system variables can be "separated" in two subvectors which do not cause each other at horizon one (*separation condition*). This separation condition is equivalent to a *definition* of noncausality proposed by Hsiao (1982) for systems with more than three variables; Hsiao's condition, however, is *not generally necessary* for noncausality at all horizons (as defined here). All these results are derived for general processes in L^2 (i.e., processes with finite second moments), without any assumption on stationarity or specific parametric forms (such as autoregressive or ARMA models).

One should note that the notion of noncausality at all horizons ($h = \infty$) studied here is not generally equivalent to the one considered by Lütkepohl (1993). The latter, indeed, is not a generalization of the usual concept of noncausality in the sense of Granger (1969), for it is based on whether the *innovations* of a variable have an effect on the other variable (i.e., whether the corresponding coefficients in the moving average (MA) representation are zero), not on whether a given variable can help to predict another one. In multivariate models where auxiliary variables are used to predict, these two notions are not equivalent even for the horizon one; see Dufour and Tessier (1993). Since one of the main characteristics of the Wiener-Granger notion of causality is the emphasis on prediction, we extend it to longer horizons by retaining prediction from *observable variables* as the central concept.

In Section 3, we study the case where the process considered has an autoregressive representation possibly of infinite order. These conditions include as special cases autoregressive processes of finite order (VAR), stationary or nonstationary, a wide class of second-order stationary processes (including long-memory processes, such as fractional processes), and invertible ARMA processes. In particular, we allow for processes with initial conditions (i.e., conditioning on the history of the process up to a given date), so that autoregressive processes with unit or explosive root(s) are included in our setup. It is not required that the covariance matrices of the innovations be constant (i.e., heteroskedastic innovations are allowed). The results presented considerably generalize several results presented by Boudjellaba, Dufour, and Roy (1992, 1994) for the horizon one, and by Renault and Szafarz (1991) for autoregressive processes of order one. We give several characterizations of noncausality at different horizons: for regular processes (i.e., processes with nonsingular innovation covariance matrices), we give necessary and sufficient conditions, while for nonregular processes we show that the same conditions are sufficient. In particular, we give a characterization of noncausality in terms of "causality chains," a formulation which throws considerable light on the relationship between causality at horizons greater than one and the presence of "indirect causal effects." From the causal chain characterization of noncausality, we also derive (as a corollary) necessary conditions for noncausality at all horizons which involve coefficients of the moving average representation of the process, illustrating the link between our definition of noncausality and the one considered by Lütkepohl (1993). When the vector of auxiliary variables $Z(t)$ is univariate (hence in particular for trivariate processes), it is also observed that these conditions are sufficient as well as necessary. We are then able to show that the *separation condition*, which was shown to be sufficient under very general assumptions (Section 2) is also necessary for noncausality at all horizons for the special case of systems which include only one auxiliary variable. In other words, if the auxiliary variable vector Z has only one component, there are only two cases that can make Y not cause X at all horizons: the one where Y does not cause $(X', Z)'$ at horizon 1, and the one where $(Y', Z)'$ does not cause X at horizon 1 (X and Y can be vectors). This separation criterion, when Z is univariate, coincides with the definition of noncausality proposed by Hsiao (1982). Finally, we observe that the coefficients of lagged variables in forecasts at different horizons $h \geq 1$ can be interpreted as *generalized impulse response coefficients* which provide a complete picture of linear causality properties at different horizons. By contrast, usual impulse response coefficients only constitute a small subset of those and can easily give a misleading picture of the causality structure of a vector time series.

In Section 4, we consider the important case of finite order VAR processes, stationary or nonstationary, and show that the characterizations of noncausality obtained for infinite order autoregressive processes reduce in such cases to finite sets of restrictions. These restrictions may then be used for implementing tests. In Section 5 finally, we make a number of concluding remarks and mention briefly the inference problems associated with the causality concepts discussed above.

2. LINEAR CAUSALITY AT DIFFERENT HORIZONS

The concepts of causality studied here are extensions of the original definitions of Wiener (1956) and Granger (1969) in a linear framework similar to the one considered by Hosoya (1977) and Florens and Mouchart (1985). More precisely, noncausality is defined in terms of orthogonality conditions between subspaces of a Hilbert space of random variables with finite second moments. We denote by $L^2 = L^2(\Omega, \mathscr{A}, Q)$ this Hilbert space of real random variables defined on a common probability space $(\Omega, \mathscr{A}, Q)$, with covariance as inner product.

In this context, the "information available at time t" is defined by a closed subspace $I(t)$ of L^2 (Hilbert subspace), where $t \in \mathbb{Z}$ and $\mathbb{Z}$ is the set of the integers. We consider a nondecreasing sequence I of such subspaces, i.e.,

$$I = \{I(t) : t \in \mathbb{Z}, t > \omega\}, \quad \text{and} \quad t < t' \Rightarrow I(t) \subseteq I(t') \quad \text{for all} \quad t > \omega, \tag{2.1}$$

where $\omega \in \mathbb{Z} \cup \{-\infty\}$. We will call $I(t)$ the "reference information set." This means in particular that memory is unbounded and information is not lost as t increases. In addition, we consider an $m_1 \times 1$ vector process of interest X in L^2, i.e., we have

$$(2.2)\qquad X=\{X(t)\colon t\in\mathbb{Z},\, t>\omega\},$$
$$X(t)=\left(x_1(t),\ldots,x_{m_1}(t)\right)',\qquad x_i(t)\in L^2\qquad (i=1,\ldots,m_1),$$

and we suppose that the information sequence I is conformable with X according to the following definition.

Definition 2.1 (Conformable Information Sequence): Under the assumptions (2.1) and (2.2), we say that the information sequence I is *conformable with* X if $X(\omega,t]\subseteq I(t)$ for every integer $t>\omega$, where $X(\omega,t]$ is the Hilbert space spanned by the components $x_i(\tau)$, $i=1,\ldots,m_1$ of $X(\tau)$, $\omega<\tau\le t$.

In other words, the past and present of $X(\tau)$ belong to the information set $I(t)$. At this stage, the "starting point" ω is not specified: in particular, ω may equal $-\infty$ or 0 depending on whether we consider a stationary process on the integers ($t\in\mathbb{Z}$) or a process $\{X(t)\colon t\ge 1\}$ on the positive integers given initial values preceding date 1. The set $I(\omega)=\bigcap_{t>\omega}I(t)$ represents information available at any date $t>\omega$, such as constants, deterministic variables, or initial conditions (on $X(t)$ or other variables).

In general, knowing $I(t)$ does not allow one to predict perfectly a future value $X(t+h)$, where $h\in\mathbb{N}$ is the *prediction horizon*; $\mathbb{N}=\{1,2,\ldots\}$ represents the positive integers while $\mathbb{N}_0=\{0,1,2,\ldots\}$ is the set of the nonnegative integers. We denote $P[X(t+h)|I(t)]$ the best linear forecast of $X(t+h)$ based on the information $I(t)$: i.e., each component $P[x_i(t+h)|I(t)]$, $1\le i\le m_1$, of $P[X(t+h)|I(t)]$ is the linear projection of $x_i(t+h)$ on the Hilbert subspace $I(t)$ (orthogonal projections with respect to the inner product of L^2). If the minimal information set $I(\omega)$ contains a nonzero constant variable, $P[X(t+h)|I(t)]$ is the *affine regression* of $X(t+h)$ on $I(t)$, which in turn coincides, for Gaussian processes, with the conditional expectation of $X(t+h)$ given $I(t)$. The concept of (linear) causality in the sense of Wiener-Granger from a process Y to a process X is based on studying whether we can improve the forecast of $X(t+1)$ by using, in addition to $I(t)$, information about the past and present values $Y(\tau)$, $\omega<\tau\le t$, of Y. Here we suppose that

$$(2.3)\qquad Y(t)=\left(y_1,(t),\ldots,y_{m_2}(t)\right)',\qquad y_j(t)\in L^2,$$
$$\text{for}\quad 1\le j\le m_2,\qquad t\in\mathbb{Z},$$

and we denote $I(t)+Y(\infty,t]$ the Hilbert space of L^2 generated by $I(t)$ and the components $y_j(\tau)$, $j=1,\ldots,m_2$, $\omega<\tau\le t$. (More generally, for any two subspaces E and F of L^2, we will denote $E+F$ the Hilbert subspace generated by the elements of E and F.) We now give a general definition of noncausality at various horizons, with respect to a "universe" $I(t)$.

DEFINITION 2.2 (Noncausality at Different Horizons): Let the assumptions (2.1) to (2.3) hold and suppose that I is conformable with X. Then, for $h \in \mathbb{N}$, we say that: (i) *Y does not cause X at horizon h given I* (denoted $Y \underset{h}{\nrightarrow} X|I$) if

$$P[X(t+h)|I(t)] = P[X(t+h)|I(t) + Y(\omega, t]], \qquad \forall t > \omega;$$

(ii) *Y does not cause X up to horizon h given I* (denoted $Y \underset{(h)}{\nrightarrow} X|I$) if $Y \underset{k}{\nrightarrow} X|I$ for $k = 1, 2, \ldots, h$; (iii) *Y does not cause X at any horizon given I* (denoted $Y \underset{(\infty)}{\nrightarrow} X|I$) if $Y \underset{k}{\nrightarrow} X|I$ for all $k \in \mathbb{N}$.

We shall make four comments to clarify the latter definition. *First*, it is a natural extension of the usual definition of noncausality proposed by Wiener (1956) and Granger (1969) in terms of predictibility, where the role of the forecast horizon h is emphasized. In the same vein, for cases where no set of auxiliary variables is used (the information set I), related notions ("noncausality of order h" and "global noncausality") cast in terms of independence (vs. the L^2 framework used here) were introduced by Bouissou, Laffont, and Vuong (1986), Florens and Mouchart (1982), and Kohn (1981). However, with auxiliary variables (as noted by Lütkepohl (1993) and Renault and Szafarz (1991)), noncausality at horizon $h = 1$ is neither a necessary nor a sufficient condition for $Y(\omega, t]$ to be useless in predicting X at longer horizons ($h \geq 2$). *Second*, although close in spirit to the notion of "noncausality" (at all horizons) in Lütkepohl (1993), the latter differs from ours because it is based on the absence of effect from the *innovations* of a variable to another variable (zero coefficients in the MA representation of the process) rather than the absence of the past values of a variable in the optimal forecasts of another variable; for further discussion of this point, see Dufour and Tessier (1993). We think Definition 2.2 provides a more natural extension of the usual definition of Granger causality. *Third*, Definition 2.2 allows for nonstationary processes. *Fourth*, as fundamentally a time domain concept, it does not appear to have a simple frequency domain interpretation, e.g., along the lines followed by Hosoya (1991) and Granger and Lin (1995). *Finally*, the "universe" or "reference information set" $I(t)$ considered to study causality properties contains at every date t, beyond the variable of interest X up to date t, an information $I(\omega) = \bigcap_{t > \omega} I(t)$ available at every date $t > \omega$ (e.g., initial conditions, constants, deterministic variables) and information accumulated between dates ω and t (difference between $I(t)$ and $I(\omega)$) about both X and auxiliary variables Z.

It is precisely the presence of auxiliary variables not contained in X or Y that can lead to a situation where Y does not cause X up to horizon h, but causes it at horizon $h + 1$. Without giving further details on the variables contained by the information set I, it is already possible to derive some results on causality at different horizons. The first one generalizes a result given in Boudjellaba, Dufour, and Roy (1992) under more restrictive assumptions and for the horizon one only.

PROPOSITION 2.1 (Componentwise Characterization of Noncausality between Vectors): *Let the assumptions* (2.1) *to* (2.3) *hold, and suppose I is conformable with X. Then the following properties are equivalent for any* $h \in \mathbb{N}$: (i) $Y \underset{h}{\nrightarrow} X \mid I$; (ii) $Y \underset{h}{\nrightarrow} x_i \mid I$, *for* $i = 1, \ldots, m_1$; (iii) $y_j \underset{h}{\nrightarrow} X \mid I$, *for* $j = 1, \ldots, m_2$; (iv) $y_j \underset{h}{\nrightarrow} x_i \mid I$, *for* $i = 1, \ldots, m_1$ *and* $j = 1, \ldots, m_2$.

The proofs of the propositions are given in the Appendix. The above proposition shows that causality between vectors can be studied by considering causality between corresponding components of Y and X. This can lead to important simplifications because real variables are simpler to study than vectors. Note the basic information set I must be the same in the four conditions (i)–(iv) and should not depend on i or j; on this issue (for horizon 1), see Florens and Mouchart (1985, Property 3.5, p. 164). About changing the information set, it is however possible to prove the following proposition.

PROPOSITION 2.2: *Let the assumptions of Proposition* 2.1 *hold, and define* $I_{(j)}(t)$ *as the Hilbert space generated by* $I(t)$ *and the variables* $y_k(\tau)$, $\omega < \tau \leq t$, $k = 1, \ldots, m_2$, $k \neq j$. *Then, for any* $h \in \mathbb{N}$: (*i*) $Y \underset{h}{\nrightarrow} X \mid I \Rightarrow y_j \underset{h}{\nrightarrow} X \mid I_{(j)}$, *for* $j = 1, \ldots, m_2$; (*ii*) *the converse implication is not true in general.*

The latter proposition means that, whenever $Y \underset{h}{\nrightarrow} X \mid I$, and starting from the complete information set $I(t) + Y(\omega, t]$, the forecast accuracy of $X(t+h)$ is not reduced by dropping the information provided by any individual component y_j of Y $(1 \leq j \leq m_2)$. The converse however can hold only if the components of Y satisfy conditions of linear independence (see next section). We now give a proposition which shows clearly that the causality horizon matters only in situations where the universe I involves processes other than X and Y.

PROPOSITION 2.3 (Exhaustivity Condition for Noncausality at all Horizons): *Under the assumptions* (2.1) *to* (2.3), *suppose* $I(t) = H + X(\omega, t]$, *for* $t > \omega$, *where* H *is a* (*possibly empty*) *Hilbert subspace of* L^2. *Then the three following properties are equivalent*:

$$\text{(i) } Y \underset{1}{\nrightarrow} X \mid I; \qquad \text{(ii) } Y \underset{(h)}{\nrightarrow} X \mid I, \quad \forall h \in \mathbb{N}; \qquad \text{(iii) } Y \underset{(\infty)}{\nrightarrow} X \mid I.$$

Proposition 2.3 gives a case where the usual notion of causality in the sense of Granger ($Y \underset{1}{\nrightarrow} X \mid I$) implies that Y cannot help to predict X at every horizon: this case is the one where the only information that gets added to $I(t)$, as t increases, is contained in $X(t)$ and $Y(t)$. For example, any bivariate model satisfies this condition. Note H may contain any variable in L^2 which does not depend on t (e.g., known at every date $t > \omega$). This result may be viewed as an extension of earlier results due to Bouissou, Laffont, and Vuong (1986, Lemma 1) and Florens and Mouchart (1982, p. 590).

Let us now consider a universe $I(t)$ richer than the one of Proposition 2.3, because it contains past and present observations about another vector $Z(t)$ from the process:

$$(2.4) \qquad Z(t) = \big(z_1(t),\dots,z_{m_3}(t)\big)', \qquad z_k(t)\in L^2, \quad \text{for } k=1,\dots,m_3 \quad \text{and} \quad t>\omega,\ t\in\mathbb{Z}.$$

The reference information set at date t is then

$$(2.5) \qquad I(t) = I_{XZ}(t) = H + X(\omega,t] + Z(\omega,t],$$

where H is defined as in Proposition 2.3. In this case, the latter cannot be applied directly to show that $Y \underset{(\infty)}{\not\rightarrow} X \mid I_{XZ}$. However, an important case where the latter property holds is the one where a *separation condition* is satisfied. Suppose indeed we can find two processes $\{Z_1(t)\colon t\in\mathbb{Z},\ t>\omega\}$ and $\{Z_2(t)\colon t\in\mathbb{Z},\ t>\omega\}$ of dimensions m_{31} and m_{32} respectively such that

$$(2.6) \qquad I_{XZ}(t) = I_{XZ_1}(t) + Z_2(\omega,t], \qquad I_{XZ_1}(t) = H + X(\omega,t] + Z_1(\omega,t], \qquad \forall t>\omega.$$

This will hold, in particular, if $(Z_1(t)', Z_2(t)')'$ is an invertible linear transformation of $Z(t)$, e.g., when $Z(t) = (Z_1(t)', Z_2(t)')'$. We shall admit here that m_{31} or m_{32} can be zero, corresponding to cases where either $Z = Z_2$ or $Z = Z_1$. Then we have the following property.

PROPOSITION 2.4 (Separation Condition for Noncausality at all Horizons): *Let H be a Hilbert subspace of L^2, and suppose the assumptions* (2.1) *to* (2.6) *hold. Then the separation condition $(Y', Z_2')' \underset{1}{\not\rightarrow} (Y', Z_1')' \mid I_{XZ_1}$ is sufficient for $Y \underset{(\infty)}{\not\rightarrow} X \mid I_{XZ}$.*

Intuitively, Proposition 2.4 means that whenever the separation condition holds, not only do we have $Y \underset{1}{\not\rightarrow} X \mid I_{XZ}$ but also $Y \underset{(\infty)}{\not\rightarrow} X \mid I_{XZ}$. This comes from the fact that no "causality chain from Y to X" (see Renault and Szafarz (1991)) can operate by going through Z (indirect causality), because the linear transformations of Z that can be "caused by Y" (the components of Z_2) do not cause X.

3. CAUSALITY IN LINEAR INVERTIBLE PROCESSES

We now consider the more specific case of "linear invertible processes," a setup which remains quite general since it includes as special cases both finite order VAR models (stationary or nonstationary) and invertible ARIMA processes. This will allow us to obtain explicit parametric formulations of the noncausality conditions at various horizons. The characterizations so obtained will provide both more insight into the nature of the restrictions (e.g., causality chain characterizations) and a basis for developing tests.

We consider here an $m\times 1$ discrete-time process $\{W(t)\colon t\in\mathbb{Z}\}$ in L^2 with an autoregressive representation (possibly of infinite order):

$$(3.1) \qquad W(t) = \mu(t) + \sum_{j=1}^{\infty} \pi_j W(t-j) + a(t), \qquad \forall t>\omega,$$

where $\mu(t)$ belongs to some Hilbert subspace H of L^2 ($\mu(t) \in H$, $\forall t > \omega$), $\{a(t): t \in \mathbb{Z}\}$ is a sequence of random vectors in L^2 with mean zero, mutually uncorrelated and such that $a(t)$ is orthogonal to the Hilbert space $H + W(-\infty, t]$; we also assume that the series $\sum_{j=1}^{\infty} \pi_j W(t-j)$ converges in quadratic mean (q.m.) for any $t > \omega$. Note the vectors $\mu(t)$ may be nonrandom or equal to zero for all t. It is not required that the covariance matrix of $a(t)$ be constant or nonsingular. Thus $a(t)$ might not be a white noise process. Further, the vector $W(t)$ is partitioned into three subvectors,

$$W(t) = (X(t)', Y(t)', Z(t)')', \qquad t \in \mathbb{Z}, \tag{3.2}$$

where $X(t)$, $Y(t)$, and $Z(t)$ have dimensions m_1, m_2, and m_3 respectively ($m_1 \geq 1$, $m_2 \geq 1$, $m_3 \geq 0$, $m_1 + m_2 + m_3 = m$), and the reference information set is defined by observing at each date t the past and present of $X(t)$ and $Z(t)$, plus the information contained in H:

$$I(t) = I_{XZ}(t) = H + X(-\infty, t] + Z(-\infty, t]. \tag{3.3}$$

In the special case where the vectors $a(t)$ have nonsingular covariance matrices, i.e.,

$$\det(E[a(t)a(t)']) \neq 0, \qquad \forall t > \omega, \tag{3.4}$$

we will say that the process $W(t)$ is *regular*. However, several of the results given below hold without this assumption. Even though the process $W(t)$ is defined for all $t \in \mathbb{Z}$, the representation (3.1) need only hold for $t > \omega$. When $\omega > -\infty$ the values of $W(t)$ for $t \leq \omega$ (i.e., the initial values of the process) may be set at any appropriate values which ensure the convergence of the series in (3.1); for example, this will occur if $W(t) = 0$ for $t < \omega' \leq \omega$. This formalism will allow us to study simultaneously stationary processes on the integers (in which case $\omega = -\infty$) and nonstationary autoregressive processes with initial conditions. In the case of second-order stationary processes, a sufficient condition for the series (3.1) to converge in q.m. is $\sum_{j=1}^{\infty} \|\pi_j\| < \infty$, where $\|\pi_j\|^2 = \text{tr}(\pi_j \pi_j')$. Note also that model (3.1) includes as a special case the model

$$X(t) - \bar{\mu}(t) = \sum_{j=1}^{\infty} \pi_j [X(t-j) - \bar{\mu}(t-j)] + a(t), \qquad t > \omega, \tag{3.5}$$

where each component of $\bar{\mu}(t)$ belongs to H and the series $\sum_{j=1}^{\infty} \pi_j \bar{\mu}(t-j)$ converges in q.m. In (3.5), the function $\bar{\mu}(t)$ may be interpreted as a centering (or detrending) function. Another important case where model (3.1) applies is the one of invertible ARIMA processes.

The autoregressive form (3.1) naturally yields forecasts at any horizon h, given the information $H + W(-\infty, t]$ available at time t. The latter may be computed easily from the formula

$$\begin{aligned} (3.6) \qquad & P[W(t+h) | H + W(-\infty, t]] \\ &= \sum_{k=0}^{h-1} \pi_1^{(k)} \mu(t+h-k) + \sum_{j=1}^{\infty} \pi_j^{(h)} W(t+1-j), \quad \forall t > \omega, \forall h \in \mathbb{N}, \end{aligned}$$

where we set $\pi_1^{(0)} = I_m$ and, for each $j \in \mathbb{N}$, the sequence of matrices $\pi_j^{(h)}$, $h \in \mathbb{N}$, is defined recursively by

$$(3.7) \qquad \pi_j^{(1)} = \pi_j, \qquad \pi_j^{(h+1)} = \pi_{j+h} + \sum_{l=1}^{h} \pi_{h-l+1}\pi_j^{(l)} \qquad (h = 1,2,\dots).$$

Furthermore, it will be useful to observe that any sequence of matrices $\pi_j^{(h)}$, where $j \in \mathbb{N}$ and $h \in \mathbb{N}$, that satisfies (3.7) also satisfies the recursion

$$(3.8) \qquad \pi_j^{(1)} = \pi_j, \qquad \pi_j^{(h+1)} = \pi_{j+1}^{(h)} + \pi_1^{(h)}\pi_j \qquad (h = 1,2,\dots).$$

For an explicit derivation of (3.6)–(3.8) in the context of model (3.1), the reader may consult Dufour and Renault (1994). We shall now characterize the noncausality $Y \underset{h}{\nrightarrow} X | I_{XZ}$ from natural partitions of the matrices $\pi_j^{(h)}$ conformable with X, Y, and Z:

$$(3.9) \qquad \pi_j^{(h)} = \begin{bmatrix} \pi_{XXj}^{(h)} & \pi_{XYj}^{(h)} & \pi_{XZj}^{(h)} \\ \pi_{YXj}^{(h)} & \pi_{YYj}^{(h)} & \pi_{YZj}^{(h)} \\ \pi_{ZXj}^{(h)} & \pi_{ZYj}^{(h)} & \pi_{ZZj}^{(h)} \end{bmatrix}.$$

Our basic result on this issue is given in the following theorem, which is a generalization of Proposition 1 in Boudjellaba, Dufour, and Roy (1992).

THEOREM 3.1 (Projection Coefficient Characterization of Noncausality at Horizon h): *Under the assumptions* (3.1) *to* (3.3), *the condition*

$$(3.10) \qquad \pi_{XYj}^{(h)} = 0, \qquad \forall j \in \mathbb{N},$$

is sufficient for $Y \underset{h}{\nrightarrow} X | I_{XZ}$, *where* $h \in \mathbb{N}$. *If, furthermore, the process* $W(t)$ *is regular* (*assumption* (3.4)) *then* $Y \underset{h}{\nrightarrow} X | I_{XZ} \Leftrightarrow \pi_{XYj}^{(h)} = 0$, $\forall j \in \mathbb{N}$.

The latter theorem jointly with (3.8) allows one to understand why in the presence of an auxiliary variable vector $Z(t)$, Y can cause X at horizon $h+1$ even though it does not at horizon h. By the recursion (3.8), we have

$$(3.11) \qquad \pi_{XYj}^{(h+1)} = \pi_{XY,j+1}^{(h)} + \pi_{XX1}^{(h)}\pi_{XYj} + \pi_{XY1}^{(h)}\pi_{YYj} + \pi_{XZ1}^{(h)}\pi_{ZYj}$$

which upon using Theorem 3.1 entails the following result.

COROLLARY 3.1: *Under the assumptions* (3.1) *to* (3.4),

$$Y \underset{(h)}{\nrightarrow} X | I_{XZ} \Rightarrow \pi_{XYj}^{(h+1)} = \pi_{XZ1}^{(h)}\pi_{ZYj}, \qquad \forall j \in \mathbb{N}.$$

In other words, when there is no causality from Y to X up to horizon h, causality can still appear at horizon $h+1$ if the auxiliary variable(s) Z can cause X at horizon h ($\pi_{XZ1}^{(h)} \neq 0$) and Y can cause Z at horizon 1 ($\pi_{ZY_j} \neq 0$). Thus the presence of Z can introduce indirect causality from Y to X going through Z. This leads to *sufficient conditions* for noncausality from Y to X which gives a more explicit form to Proposition 2.4. For example, if $Z(t) = (Z_1(t)', Z_2(t)')'$, we see that $Y \underset{(h)}{\nrightarrow} X | I_{XZ} \Rightarrow \pi_{XYj}^{(h+1)} = \pi_{XZ_1 1}^{(h)} \pi_{Z_1 Yj} + \pi_{XZ_2 1}^{(h)} \pi_{Z_2 Yj}$, hence sufficient conditions like $Y \underset{1}{\nrightarrow} (X', Z_1')' | I_{XZ}$ and $Z_2 \underset{(h)}{\nrightarrow} X | I_{XYZ_1} \Rightarrow Y \underset{(h+1)}{\nrightarrow} X | I_{XZ}$, which can be verified easily. Theorem 3.2 below gives a complete characterization of noncausality from Y to X in terms of "causality chains." To prove it, we will need two lemmas (of separate interest) on the properties of matrix sequence $\pi_j^{(h)}$ that satisfy recursion (3.8).

LEMMA 3.1: *Let $\pi_j^{(h)}$, $j \in \mathbb{N}$, $h \in \mathbb{N}$, be any sequence of $m \times m$ matrices, which are partitioned as in* (3.9) *and satisfy the recursion* (3.8). *If $\pi_{XYj}^{(k)} = 0$, $\forall j \in \mathbb{N}$, $k = 1, \ldots, h$, then for any integer p such that $2 \le p \le h$, we have*

$$\pi_{XZ1}^{(h)} \pi_{ZYj} = \sum_{l=1}^{p} \pi_{XZl}^{(h-p+1)} \left\{ \sum_{J(p-l)} \left[\prod_{i=1}^{p-l} \pi_{ZZi}^{n_i} \right] \right\} \pi_{ZYj}, \qquad \forall j \in \mathbb{N}, \tag{3.12}$$

where $J(l) = \{(n_1, n_2, \ldots, n_l)$: $\sum_{i=1}^{l} i n_i = l$ and $n_i \in \mathbb{N}_0$, $i = 1, \ldots, l\}$, $\prod_{i=1}^{p-l} \pi_{ZZi}^{n_i} = \pi_{ZZ1}^{n_1} \pi_{ZZ2}^{n_2} \cdots \pi_{ZZ,k-l}^{n_{p-l}}$, $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$, with the convention $\sum_{J(0)} [\prod \pi_{ZZi}^{n_i}] = I_{m_3}$.

LEMMA 3.2: *Under the assumptions of Lemma 3.1, the three following conditions are equivalent for $h \ge 2$:*

$$\pi_{XYj}^{(k)} = 0, \qquad \forall j \in \mathbb{N}, \qquad (k = 1, \ldots, h); \tag{3.13}$$

$$\pi_{XYj} = 0, \qquad \forall j \in \mathbb{N}, \qquad \textit{and} \tag{3.14}$$

$$\pi_{XZ1}^{(k)} \pi_{ZYj} = 0, \qquad \forall j \in \mathbb{N} \qquad (k = 1, \ldots, h-1);$$

$$\pi_{XYj} = 0, \qquad \forall j \in \mathbb{N}, \qquad \textit{and} \tag{3.15}$$

$$R_{XZ}^{(k)} \pi_{ZYj} = 0, \qquad \forall j \in \mathbb{N} \qquad (k = 1, \ldots, h-1),$$

where $R_{XZ}^{(k)} = \sum_{l=1}^{k} \pi_{XZl} \{\sum_{J(k-l)} [\prod_{i=1}^{k-l} \pi_{ZZi}]\}$.

THEOREM 3.2 (Causality Chain Characterization of Noncausality at Horizon h): *Under the assumptions* (3.1) *to* (3.3), *each one of the three equivalent conditions* (3.13), (3.14), *and* (3.15) *is sufficient for $Y \underset{(h)}{\nrightarrow} X | I_{XZ}$. If furthermore the process $W(t)$ is regular* (*assumption* (3.4)), *then each one of the conditions* (3.13), (3.14), *and* (3.15) *is necessary and sufficient for $Y \underset{(h)}{\nrightarrow} X | I_{XZ}$.*

The interest of the criteria (3.14) and (3.15) of noncausality $Y \underset{(h)}{\nrightarrow} X | I_{XZ}$, as opposed to (3.13) (derived in Theorem 3.1), comes from the fact that they are more clearly linked to the fundamental autoregressive coefficients π_j of the representation (3.1). Criterion (3.14) shows that noncausality at horizon h occurs when two conditions hold: (i) there is noncausality at horizon 1 ($\pi_{XYj}=0$, $\forall j$), and (ii) the composed effects $\pi_{XZ1}^{(k)}\pi_{ZYj}$ that run first from Y to Z at horizon 1 (π_{ZYj}) and then from Z to X at horizons less than h ($\pi_{XZ1}^{(k)}$, $k=1,\ldots,h-1$) are zero. Criterion (3.14) gives an explicit expression for these composed effects by relating them to all possible "causality chains" that run from Y to Z, from components of Z to other components of Z, and then from Z to X (see the expression (3.15)).

Criterion (3.14) also provides a link between our concept of "noncausality at all horizons" and the coefficients of the moving average (MA) representation of the process ("impulse response coefficients;" see Sims (1980)). To do this, consider the formal series:

$$(3.16) \qquad \pi(z)=I_m-\sum_{j=1}^{\infty}\pi_j z^j, \qquad \psi(z)=\pi(z)^{-1}=I_m+\sum_{j=1}^{\infty}\psi_j z^j.$$

These formal series, when applied to lag operators, characterize the autoregressive representation (3.1), $\pi(B)W(t)=\mu(t)+a(t)$, and eventually the moving average representation, $W(t)=\psi(B)\mu(t)+\psi(B)a(t)$, provided the series involved converge in q.m. It will be useful to notice here another algebraic property of the matrices $\pi_j^{(h)}$. Since $\pi(z)\psi(z)=I_m$ and using (3.7), it is easy to see that

$$(3.17) \qquad \pi_1^{(h)}=\psi_h, \qquad \forall h\geq 0.$$

In other words, the coefficient matrix ψ_h of the MA representation of $W(t)$ is simply the coefficient matrix of $W(t)$ in the best forecast of $W(t+h)$ as defined in (3.6). By the definition of $\pi_1^{(h)}$, the impulse response coefficient ψ_{jkh} can thus be interpreted (and indeed could be defined) as giving the corrections to be made on $P[W_j(t+h)|I(t)]$ when the component number k of $W(t)$ is modified by one unit, while the other variables in $I(t)$ are kept unchanged ($1\leq j\leq m$, $1\leq k\leq m$). Let us now partition each matrix ψ_h conformably with X, Y, and Z:

$$(3.18) \qquad \psi_h=\begin{bmatrix}\psi_{XXh} & \psi_{XYh} & \psi_{XZh}\\ \psi_{YXh} & \psi_{YYh} & \psi_{YZh}\\ \psi_{ZXh} & \psi_{ZYh} & \psi_{ZZh}\end{bmatrix}, \qquad h\geq 0,$$

and similarly for $\psi(z)$ in (3.16). By combining (3.17) with Theorem 3.1, it is then easy to see that noncausality at horizon h entails zero restrictions on the impulse response coefficients of a regular process.

COROLLARY 3.2 (Necessary Conditions for Noncausality at Horizon h): *Under the assumptions* (3.1) *to* (3.4), *the condition*

$$(3.19) \qquad \pi_{XYj}=0, \qquad \forall j\in\mathbb{N}, \qquad \textit{and} \qquad \psi_{XYh}=0$$

is necessary for $Y \underset{h}{\nrightarrow} X|I_{XZ}$ (*where* $1\leq h\leq\infty$).

It is important to note that condition (3.19) is *only necessary* for $Y \underset{h}{\nrightarrow} X | I_{XZ}$: the necessary and sufficient condition given by Theorem 3.1 requires $\pi_{XYj}^{(h)} = 0$ *for all* $j \geq 1$. If we now combine (3.17) with Theorem 3.2, we can get conditions for $Y \underset{(h)}{\nrightarrow} X | I_{XZ}$ that involve both impulse responses and autoregressive parameters.

COROLLARY 3.3 (Impulse Response Characterization of Noncausality up to Horizon h): *Under the assumptions* (3.1) *to* (3.3), *the condition*

$$\text{(3.20)} \qquad \pi_{XYj} = 0, \quad \forall j \in \mathbb{N}, \quad \textit{and}$$
$$\psi_{XZk}\pi_{ZYj} = 0, \quad \forall j \in \mathbb{N} \qquad (k = 1, \ldots, h-1),$$

is sufficient for $Y \underset{(h)}{\nrightarrow} X | I_{XZ}$. *If furthermore the process* $W(t)$ *is regular* (*assumption* 3.4)), *condition* (3.20) *is necessary and sufficient for* $Y \underset{(h)}{\nrightarrow} X | I_{XZ}$.

For noncausality at all horizons, it is possible to derive more compact characterizations given in the following corollary (where $\mathbb{C}$ is the set of complex numbers).

COROLLARY 3.4 (Necessary Conditions for Noncausality at All Horizons): *Let the assumptions* (3.1) *to* (3.4) *hold, and suppose the power series* $\pi(z)$ *converges when* $z \in \mathbb{C}$ *and* $|z| < \rho$, *for some* $\rho > 0$. *Then the three following conditions are equivalent and each one of them yields a necessary condition for* $Y \underset{(\infty)}{\nrightarrow} X | I_{XZ}$:

$$\text{(3.21)} \qquad \pi_{XY}(z) \equiv 0 \quad \textit{and} \quad \pi_{XZ}(z)\pi_{ZZ}(z)^{-1}\pi_{ZY}(z) \equiv 0,$$

$$\text{(3.22)} \qquad \pi_{XY}(z) \equiv 0 \quad \textit{and} \quad \psi_{XY}(z) \equiv 0,$$

$$\text{(3.23)} \qquad \psi_{XY}(z) \equiv 0 \quad \textit{and} \quad \psi_{XZ}(z)\psi_{ZZ}(z)^{-1}\psi_{ZY}(z) \equiv 0,$$

where $\psi(z)$ *is defined by* (3.16) *and* (3.18), *and the symbol* $\equiv$ *means that the two formal series* (*in* z) *considered are identical* (*i.e.*, *the coefficients of the corresponding powers of* z *are equal on both sides of* $\equiv$).

The assumption that $\pi(z) = I_m - \sum_{j=1}^{\infty} \pi_j z^j$ converges for $|z| < \rho$ will be met in almost all cases of practical interest, since it is satisfied whenever the sequence $\|\pi_j\|$, $j \geq 1$, is bounded and even if $\|\pi_j\|$ grows at an exponential rate as j increases (e.g., if $\|\pi_j\| = C\rho_1^j$, with $\rho_1 > 1$ and $\rho = \rho_1^{-1}$). Note condition (3.23) is formally identical with (3.21) on permuting π and ψ, which is possible because π and ψ play symmetric roles in (3.22). Note also condition (3.22) was stated by Bruneau and Nicolaï (1992) for the special case of stationary finite-order VAR processes. These (necessary) conditions eliminate both the *direct* effect of Y on X (by canceling the autoregressive operator $\pi_{XY}(L)$) and various *indirect* effects of Y on X (by cancelling the coefficients of the innovations of Y in the MA representation of X). Although it might appear at first sight that these conditions should also be sufficient for Y not to cause X at all horizons, we can see from Theorems 3.2 that they are not *when Z is multivariate*. Indeed, since $R_{XZ}(z) \equiv -\pi_{XZ}(z)\pi_{ZZ}(z)^{-1}$ (see Lemma A.3 in the Appendix), condition (3.21) can be written:

$$(3.24) \qquad \pi_{XYj}=0, \qquad \forall j \geq 1, \qquad \text{and} \qquad \sum_{j=1}^{k-1} R_{ZYj}^{(k-j)}\pi_{ZYj}=0, \qquad \forall k>1.$$

On the other hand, the necessary and sufficient condition (3.15) from Theorem 3.2 is

$$(3.25) \qquad \pi_{XYj}=0, \qquad \forall j \geq 1, \qquad \text{and} \qquad R_{XZ}^{(k)}\pi_{ZYj}=0, \qquad \forall k \geq 1, \qquad \forall j \geq 1,$$

which is not equivalent to (3.24). The comparison of (3.24) and (3.25) shows in particular that the condition $\pi_{XY}(z) \equiv \psi_{XY}(z) \equiv 0$ is generally insufficient for $Y \underset{(\infty)}{\nrightarrow} X | I_{XZ}$.

There is, however, an interesting special case where the conditions of Corollary 3.4 are also sufficient for Y not to cause X at all horizons, namely when Z is a univariate process. This result is reported in the following corollary.

COROLLARY 3.5 (Characterizations of Noncausality at All Horizons for Z Univariate): *Under the assumptions of Corollary* 3.4, *suppose the process* $Z(t)$ *is univariate* ($m_3=1$). *Then the property* $Y \underset{(\infty)}{\nrightarrow} X | I_{XZ}$ *is equivalent to each one of the three conditions* (3.21), (3.22), *and* (3.23).

The case where Z is multivariate is more complicated because causality relations internal to Z must be taken into account. This is the reason why a separation criterion is available only when Z is univariate, as given in the following corollary.

COROLLARY 3.6 (Separation Criterion when Z is Univariate): *Under the assumptions* (3.1) *to* (3.4) *with* $Z(t)$ *univariate* ($m_3=1$), $Y \underset{(\infty)}{\nrightarrow} X | I_{XZ}$ *if and only if at least one of the two following conditions is satisfied*:

$(A) \qquad Y \underset{1}{\nrightarrow} (X', Z)' | I_{XZ};$

$(B) \qquad (Y', Z)' \underset{1}{\nrightarrow} X | I_X.$

From Proposition 2.4, we know conditions (A) and (B) above are each sufficient for $Y \underset{(\infty)}{\nrightarrow} X | I_{XZ}$, irrespective of the dimension of $Z(t)$. When $Z(t)$ is univariate, at least one of these two conditions must be satisfied to have $Y \underset{(\infty)}{\nrightarrow} X | I_{XZ}$, because in this case causality chains from Y to X via Z cannot be compensated by causality chains in opposite directions, internal to Z. More precisely, the necessary condition (3.21), for the special case of systems which include only one auxiliary variable Z, can be split into two (nonexclusive) alternatives: $\pi_{XZ}(z)\pi_{ZZ}(z)^{-1}\pi_{ZY}(z) \equiv 0 \Leftrightarrow \pi_{ZY}(z) \equiv 0$ or $\pi_{XZ}(z) \equiv 0$, the first one (respectively, the second one) corresponding to (A) (respectively, to (B)). Note also that verifying (A) *or* (B) corresponds to the definition of "noncausality" proposed by Hsiao (1982, p. 247, Definition 3) for the case where X, Y, and Z are univariate. A general study of possible compensation schemes shows that Hsiao's definition is not generally equivalent to noncausality at all horizons (except precisely for a trivariate process); see Dufour and Renault (1994).

To summarize, our definition of noncausality at all horizons $(Y \underset{(\infty)}{\nrightarrow} X | I_{XZ})$ appears to be in the general case a *linear predictability* property which is strictly included between two better known properties: (i) a definition "à la Sims" in terms of innovations which can be characterized as in Lütkepohl (1993) by various restrictions on impulse response coefficients (see (3.23)), but appears to be too weak since it does not capture some indirect effects of Y on X (see the comparison between (3.24) and (3.25)); (ii) a definition "à la Hsiao" in terms of "separation" ((A) *or* (B) in Corollary 3.6) which is too restrictive since it precludes some causality relationships from Y to Z and from Z to X which are not necessarily responsible for some causal chain from Y to X. These three definitions are not generally equivalent.

In general, the informational content of the double array $(\pi_{XYj}^{(h)})$, $h = 1, 2, \ldots, j = 1, 2, \ldots,$ is much richer than the one of the two sequences (π_{XYj}), $j = 1, 2, \ldots,$ and (ψ_{XYh}), $h = 1, 2 \ldots$. In particular, the array $\pi_{XYj}^{(h)}$, $h = 1, 2, \ldots,$ $j \geq 1$, yields *generalized impulse response coefficients* (corresponding to different delays) which provide a complete picture of linear causality properties at different horizons, while usual impulse response coefficients (which correspond to the array $\pi_{XY1}^{(h)}$, $h = 1, 2, \ldots$) can be very misleading in this respect. To see this better, consider the case where X and Y are scalar processes and the (infinite dimensional) matrix $[\pi_{XYj}^{(h)}]$ has rows indexed by $h \geq 1$ and columns by $j \geq 1$. On one hand, the first row characterizes causality links from Y to X at horizon 1. Even if there is no causality at horizon 1, causal effects from Y to X may occur at greater horizons due to indirect causality chains. For instance, by Corollary 3.1, if the first row is zero, the second one provides the values of the products $\pi_{XZ1}\pi_{ZYj}$, $j = 1, 2 \ldots,$ which can be different from zero. On the other hand, the first column gives impulse response coefficients $\psi_{XYh} = \pi_{XY1}^{(h)}$ (before orthogonalization). Following a practice popularized by Sims (1980), it is usual to characterize causality links from Y to X at various horizons by the pattern of this column, i.e., the graph of ψ_{XYh} as a function of h. A current belief in applied macroeconomics is to consider that this graph summarizes causal links that may appear, directly or indirectly, at various lags $h = 1, 2, \ldots$. But we wish to stress here that, irrespective of whether the first impulse response coefficient $\psi_{XY1} = \pi_{XY1}$ is zero or not, the rest of the first column $(\psi_{XYh}, h \geq 2)$ may be zero until a large horizon H, leading one to believe that there are no lagged direct or indirect causal links, while the corresponding part of the second column $[\pi_{XYh}^{(2)}, h = 2, 3, \ldots, H]$ is nonzero. This is shown clearly by the following theorem.

THEOREM 3.3: *Let the assumptions* (3.1) *and* (3.2) *hold with the notations*:

$$\pi_j = \begin{bmatrix} \pi_{XXj} & \pi_{XYj} & \pi_{XZj} \\ \pi_{YXj} & \pi_{YYj} & \pi_{YZj} \\ \pi_{ZXj} & \pi_{ZYj} & \pi_{ZZj} \end{bmatrix} = \begin{bmatrix} \pi_{X.j} \\ \pi_{Y.j} \\ \pi_{Z.j} \end{bmatrix} = [\pi_{.Xj}, \pi_{.Yj}, \pi_{.Zj}].$$

Then, for any given values of $\pi_1, \pi_{Y.j}, \pi_{Z.j}, \pi_{.Xj}, \pi_{.Zj}$, $j \geq 2$, *and for any given integer* H *greater than* 1, *the matrices* π_{XYj}, $j = 2, 3, \ldots, H$, *can be chosen so that* $\pi_{XY1}^{(h)} = 0$, *for* $h = 2, 3, \ldots, H$. *Moreover, we have in that case*

$$\pi_{XY2}^{(h)} = -\pi_{X.1}^{(h)} \pi_{.Y1} = -\sum_{q=1}^{h} \pi_{X.q} \pi_1^{(h-q)} \pi_{.Y1}, \qquad \text{for} \quad h = 1, 2, \ldots, H-1. \tag{3.26}$$

In other words, each column $j \geq 1$, i.e., each impulse response function of order j ($\pi_{XYj}^{(h)}$, $h = 1, 2, \ldots$), is important for characterizing causality properties at different horizons, while the current practice, which considers only the first column ($j = 1$), may be very misleading (it is easy to build numerical examples illustrating this fact). In particular, the second identity in (3.26) shows why $\pi_{XY2}^{(h)}$ will be nonzero in general, even if $\pi_{XY1}^{(h)} = \psi_{XYh} = 0$ for $h = 1, 2, \ldots, H$, because the impulse response matrices $\pi_1, \pi_1^{(2)}, \ldots, \pi_1^{(h-1)}$ remain largely unconstrained (except for the matrices $\pi_{XY1}^{(l)}$, $l = 1, 2, \ldots, H$). This shows that standard (order one) impulse response analysis should be completed by considering higher-order impulse response functions: $\pi_{XYj}^{(h)}$, $h = 1, 2, 3, \ldots$ where $j > 1$. It is only by looking at these that one can get a complete picture of long horizon (indirect) linear causality properties. Furthermore, in the same way that the coefficients $\psi_{XYj} = \pi_{XYj}^{(1)}$ tell one how the forecast $P[X(t+h)|I(t)]$ should be modified when $Y(t)$ is modified by one unit (keeping the other variables in $I(t)$ unchanged), higher order impulse responses provide answers to more complex experiments such as how $P[X(t+h)|I(t)]$ should be corrected when $Y(t-1)$ is changed by one unit. Similar interpretations and generalizations can be given to the impulse responses associated with "orthogonalized" innovations, but this would go beyond the scope of the present paper.

4. NONCAUSALITY AT ALL HORIZONS IN VAR PROCESSES

The main problem associated with the results of Theorems 3.1 and 3.2 is that they generally yield an infinite number of restrictions, and so they may not be easy to test from a finite sample. This is due, of course, to the fact that model (3.1) involves an infinite number of parameters. To get empirically testable restrictions, we need to consider a finitely parameterized model. In this section, we consider the case of a vector autoregressive process of order p. For this case, we have the following proposition.

PROPOSITION 4.5 (Truncation Rule for Noncausality at All Horizons in VAR Processes): *If the process* $\{W(t): t \in \mathbb{Z}\}$ *satisfies the assumptions* (3.1) *to* (3.3) *and* $\pi_k = 0$ *for* $k > p$, *then* $\pi_{XYj}^{(h)} = 0$, $\forall j \in \mathbb{N}$, *for* $h = 1, \ldots, m_3 p + 1 \Rightarrow Y \underset{(\infty)}{\nrightarrow} X | I_{XZ}$. *If, furthermore, the process* $W(t)$ *is regular* (*assumption* (3.4)), *then* $Y \underset{(\infty)}{\nrightarrow} X | I_{YZ} \Leftrightarrow Y \underset{(m_3 p + 1)}{\nrightarrow} X | I_{XZ}$.

In other words, for autoregressive processes of order p, it is sufficient to have noncausality up to horizon $m_3 p + 1$ for noncausality at all horizons to hold. It is interesting to note that a similar result holds almost trivially for a moving average process of order $q[\psi_j = 0$ for $j > q$, where $\psi(z) = \pi(z)^{-1}]$: for $h > q$, $W(t+h)$ is orthogonal to $I_W(t)$ and thus noncausality up to horizon q is sufficient to have noncausality at all horizons. A truncation result similar to the one of Proposition 3.4 also holds for the necessary condition of Corollary 3.4.

PROPOSITION 4.6 (Necessary Condition for Noncausality at All Horizons in VAR Processes): *Under the assumptions* (3.1) *to* (3.4) *with* $\pi_k = 0$ *for* $k > p$, *the necessary condition* (3.21) *for* $Y \underset{(\infty)}{\nrightarrow} X | I_{XZ}$ *is equivalent to the following finite set of conditions*: (*i*) $\pi_{XYj} = 0$, *for* $j = 1, \ldots, p$; (*ii*) *the coefficients of* z^k *in the formal series* $\pi_{XZ}(z)\pi_{ZZ}(z)^{-1}\pi_{ZY}(z)$ *are equal to zero for* $k = 1, 2, \ldots, p(m_3 + 1)$.

This truncation result is similar to the one of Lütkepohl (1993) for the coefficients of the MA representation of a VAR process. Thus, to check empirically the necessary condition of Corollary 3.4, one needs to only consider $pm_1m_2 + p(m_3+1)m_1m_2 = pm_1m_2(m_3+2)$ restrictions on the coefficients of the autoregressive representation. On the other hand, Proposition 4.5 entails that testing $Y \underset{(\infty)}{\nrightarrow} X | I_{XZ}$ requires one to test $pm_1m_2 + (m_3 p)(pm_1m_2) = pm_1m_2(m_3 p + 1)$ restrictions, because testing causality up to horizon $m_3 p + 1$ leads one to check that the $m_3 pp$ matrices $R_{XZ}^{(k)}\pi_{ZYj}$, $k = 1, \ldots, m_3 p$, $j = 1, \ldots, p$ are all zero.

As expected, the necessary and sufficient conditions for noncausality at all horizons involve strictly more restrictions than the necessary conditions implied by Corollary 3.4. The only exceptions are the special cases $p = 1$ or $(p = 2, m_3 = 1)$, where the two sets of conditions are the same. For $p = 1$, the latter equivalence is not surprising because we then have $\pi_1^{(h)} = \pi_1^h = \psi_h$ for all $h \geq 1$, so that the constraints $\pi_{XY}(z) \equiv 0$ and $\psi_{XY}(z) \equiv 0$ given by Corollary 3.4 are indeed sufficient to characterize noncausality at all horizons. Further, it is important to note that $pm_1m_2(m_3 p + 1)$ is an *upper bound* on the effective number of restrictions which characterize noncausality at all horizons. For example, when $m_3 = 1$, we get from Corollary 3.5 and Proposition 4.6 that $Y \underset{(\infty)}{\nrightarrow} X | I_{XZ}$ is ensured by $3pm_1m_2$ restrictions, instead of $(p+1)pm_1m_2$ restrictions.

Concerning the practical applications of the above results, we wish to stress the following point. While the need of a multivariate analysis ($m_3 \neq 0$) to characterize causality between two variables ($m_1 = m_2 = 1$) is now well documented, the consequences of a multivariate environment ($m_3 > 1$), and in particular the difference between the conditions of Propositions 4.5 and 4.6 in VAR systems, is often forgotten. Besides the aforementioned misinterpretation of standard impulse response coefficients in macroeconometrics, it is of interest to look at two other fields of econometrics where our warning should be relevant: financial econometrics and the econometrics of marketing.

In debates about market efficiency in financial econometrics, several authors (e.g., Fama and French (1988), Campbell and Shiller (1988), Poterba and Summers (1988), Bekaert and Hodrick (1992)) have stressed the importance of long-horizon predictability (or causality) for asset returns. In particular, Bekaert and Hodrick (1992; BH) note that "vector autogressions (...) facilitate calculations of implied long-horizon statistics, such as variance ratios" and help to answer questions like "Does a forward premium in the foreign currency predict appreciation of the domestic currency at all horizons?" They consider a six-variable VAR system involving for a given pair of countries (U.S. and either Japan, U.K., or Germany) the U.S. equity market excess return, the companion country equity excess return, the relevant foreign exchange excess return, dividend yields, and the forward premium, estimated over the period 1981:1–1989:12 (monthly data). Since this VAR system incorporates not only the variables of interest for causality but also four "environmental" variables ($m_3 = 4$), the tools developed in this paper are needed in general to analyze the long-run causality relationship from the forward premium to foreign exchange market excess return. In this respect, BH are lucky to find that the minimized value of a Schwartz criterion is associated with a one-lag VAR system ($p = 1$), which allows them to perform their long-run causality analysis through variance ratios which aggregate only standard impulse response coefficients.

This is not the case for other data sets such as, in marketing research, the one considered by Cordier and Indjehagopian (1986; CI) to analyze the French hog market in Brittany. They fit a VAR(7) model on five price series relevant to this market (carcass, loin, ham, belly, and European carcass index), which "fluctuate simultaneously with complex links of causality and feedbacks," using 349 weekly observations (week 8 in 1975 to week 44 in 1981). Once the parameters of a VAR model have been estimated within the sample period ($t = 1, \ldots, 349$), CI compute a sequence of forecasts at time $n = 349$ for horizons $h = 1, 2, \ldots, 7$. The post-sample prediction performance is measured by the root mean square error $RMSE = (\sum_{c=1}^{h} (A_{n+c} - \hat{x}_n(c))^2 / h)^{1/2}$, where A_t is the actual value at time t and $\hat{x}_n(c)$ the c-step ahead forecast made at time n. CI conclude that a multivariate analysis provides a better forecast performance than a univariate ARMA model and that "the improvement in the forecast is smaller for leading variables such as carcass than for caused variables such as index." In view of our results, the index *RMSE* is a rather confusing causality measure which mixes different horizons and the analysis should be made clearer by looking at the generalized impulse responses in the 5×5 matrices $\pi_j^{(h)}$, $j \geq 1$, $h \geq 1$.

5. CONCLUDING REMARKS

An obvious application of the above results is the development of tests for hypotheses of the form $Y \underset{(h)}{\nrightarrow} X$ or $Y \underset{(\infty)}{\nrightarrow} X$. Provided the number of parameters in the model is finite (as for example in finite order VAR models) and standard regularity conditions hold, it is clear that Wald-type or likelihood-ratio-type tests may be applied here. Note however that the conditions given by Theorems 3.1 and 3.2 are generally nonlinear. In some cases, like the one where a separation condition holds (e.g., when Z is univariate), it is possible to reduce these nonlinear conditions to combinations of linear conditions which can be tested by testing separately causality hypotheses at the horizon one (with appropriate level adjustments to control the overall level of the procedure). But more generally we need to test zero restrictions on multilinear functions of the coefficients of the matrices π_j in (3.1). Such restrictions can lead to Jacobian matrices of the restrictions having less than full rank under the null hypothesis (for some illustrations, see Boudjellaba, Dufour, and Roy (1992, 1994)) and thus produce test statistics with nonstandard asymptotic distributions (see Andrews (1987)). Special methods are required to deal with such problems. Since those require lengthy developments, the appropriate statistical methodology and various applications are described in a separate paper (Dufour and Renault (1995)).

C.R.D.E. and Dépt. de Sciences Économiques, Université de Montréal, C.P. 6128, Succursale Centre-ville, Montréal (Québec) H3C 3J7, Canada; jean.marie.dufour@umontreal.ca; http://www.crde.umontreal.ca/personnel/dufour.html

and

CREST-INSEE, Bâtiment Malakoff 2—Timbre J301, 15 Bd. Gabriel Péri, 92245 Malakoff Cedex, France; renault@pareto.ensae.fr

Manuscript received August, 1994; final revision received August, 1997.

APPENDIX: PROOFS

PROOF OF PROPOSITION 2.1: The equivalence between (i) and (ii) is obvious from the definitions of $P[X(t+h)|I(t)]$ and $P[X(t+h)|I(t)+Y(\omega,t]]$ as the vectors of the forecasts $P[x_i(t+h)|I(t)]$ and $P[x_i(t+h)|I(t)+Y(\omega,t]]$ respectively, $i=1,\dots,m_1$. Consider now the equivalence between (i) and (iii). If $Y \underset{h}{\nrightarrow} X|I$, we have by definition

$$\text{(A.1)} \qquad P[X(t+h)|I(t)+Y(\omega,t]] = P[X(t+h)|I(t)], \qquad \forall t > \omega.$$

Thus each component of $P[X(t+h)|I(t)+Y(\omega,t]]$ is an element of $I(t) \subseteq I(t)+y_j(\omega,t] \subseteq I(t)+Y(\omega,t]$, where $y_j(\omega,t]$ is the Hilbert space generated by the variables $y_j(\tau)$, $\omega < \tau \leq t$. Using the properties of iterated projections, we then see that

$$\begin{aligned} P[X(t+h)|I(t)+y_j(\omega,t]] &= P[P[X(t+h)|I(t)+Y(\omega,t]]|I(t)+y_j(\omega,t]] \\ &= P[P[X(t+h)|I(t)]|I(t)+y_j(\omega,t]] \\ &= P[X(t+h)|I(t)], \quad \text{for} \quad t>\omega \quad \text{and} \quad j=1,\ldots,m_2, \end{aligned}$$

which means that $y_j \underset{h}{\nrightarrow} X|I$ for $j=1,\ldots,m_2$. Thus (i) $\Rightarrow$ (iii). Conversely, (iii) entails

$$P[X(t+h)|I(t)+y_j(\omega,t]] = P[X(t+h)|I(t)], \quad \forall t>\omega, \quad \text{for} \quad j=1,\ldots,m_2,$$

so that each component of $X(t+h)-P[X(t+h)|I(t)]$ must be orthogonal to the Hilbert subspace $I(t)+y_j(\omega,t]$, $j=1,\ldots,m_2$, hence also to the Hilbert subspace $I(t)+Y(\omega,t]$ which is generated by the latter subspaces. Thus we have $P[X(t+h)|I(t)] = P[X(t+h)|I(t)+Y(\omega,t]]$, i.e., $Y \underset{h}{\nrightarrow} X|I$. Thus (iii) $\Leftrightarrow$ (i). Finally, the equivalence between (iii) and (iv) follows from the definitions of $P[X(t+h)|I(t)]$ and $P[X(t+h)|I(t)+y_j(\omega,t]]$ as the vectors of the forecasts $P[x_i(t+h)|I(t)]$ and $P[x_i(t+h)|I(t)+y_j(\omega,t]]$ respectively, $i=1,\ldots,m_1$. *Q.E.D.*

Proof of Proposition 2.2: If $Y \underset{h}{\nrightarrow} X|I$, the identity (A.1) holds, and each component of $P[X(t+h)|I(t)+Y(\omega,t]]$ is an element of $I(t) \subseteq I_{(j)}(t) \subseteq I(t)+Y(\omega,t]$. Thus, by using the properties of iterated projections, we have for each $j=1,\ldots,m_2$:

$$\begin{aligned} P[X(t+h)|I_{(j)}(t)] &= P[P[X(t+h)|I(t)+Y(\omega,t]]|I_{(j)}(t)] \\ &= P[P[X(t+h)|I(t)]|I_{(j)}(t)] \\ &= P[X(t+h)|I(t)], \quad \text{for} \quad t>\omega, \end{aligned}$$

which means that $y_j \underset{h}{\nrightarrow} X|I_{(j)}$. On the other hand, we can have

$$\text{(A.2)} \qquad P[X(t+h)|I_{(j)}(t)] = P[X(t+h)|I(t)], \quad \forall t>\omega, \quad \text{for} \quad j=1,\ldots,m_2,$$

without (A.1) holding, if $m_2>1$ and the m_2 components of $Y(t)$ are identical, a situation where it is clear that $Y \underset{h}{\nrightarrow} X|I$ may not hold. *Q.E.D.*

Proof of Proposition 2.3: Since (ii) $\Rightarrow$ (i) by definition, we need to show first the converse implication, i.e., $Y \underset{1}{\nrightarrow} X|I \Rightarrow Y \underset{(h)}{\nrightarrow} X|I$, for any $h \geq 1$. The proof is done by induction. Suppose $Y \underset{(h)}{\nrightarrow} X|I$. Then, by the properties of iterated projections and since $Y \underset{1}{\nrightarrow} X|I$,

$$\begin{aligned} &P[X(t+h+1)|I(t)+Y(\omega,t]] \\ &\quad = P[P[X(t+h+1)|I(t+h)+Y(\omega,t+h]]|I(t)+Y(\omega,t]] \\ &\quad = P[P[X(t+h+1)|I(t+h)]|I(t)+Y(\omega,t]], \quad \forall t>\omega. \end{aligned}$$

$P[X(t+h+1)|I(t+h)]$ is an $m \times 1$ vector whose elements belong to

$$\begin{aligned} I(t+h) &= H+X(\omega,t+h] = H+X(\omega,t]+X[t+1,t+h] \\ &= I(t)+X[t+1,t+h] \end{aligned}$$

where $X[t+1,t+h]$ is the Hilbert subspace generated by the components $x_i(\tau)$, $i=1,\ldots,m_1$, $t+1 \leq \tau \leq t+h$, and thus we can write $P[X(t+h+1)|I(t+h)] = a_h(t)+b_h(t)$, where $a_h(t) \in I(t)$ and $b_h(t) \in X[t+1,t+h]$. Further, since $Y \underset{(h)}{\nrightarrow} X|I$, each component $x_i(\tau)$, $i=1,\ldots,m_1$, $t+1 \leq \tau \leq$

$t+h$, satisfies $P[x_i(\tau)|I(t)+Y(\omega,t]]=P[x_i(\tau)|I(t)]$, $\forall t>\omega$, which implies $P[b_h(t)|I(t)+Y(\omega,t]]=P[b_h(t)|I(t)]$, $\forall t>\omega$. We thus have:

$$\begin{aligned} P[X(t+h+1)|I(t)+Y(\omega,t]] &= P[a_h(t)|I(t)+Y(\omega,t]]+P[b_h(t)|I(t)+Y(\omega,t]] \\ &= a_h(t)+P[b_h(t)|I(t)], \qquad \forall t>\omega, \end{aligned}$$

so that each component of $P[X(t+h+1)|I(t)+Y(\omega,t]]$ belongs to $I(t)$. Consequently, $P[X(t+h+1)|I(t)+Y(\omega,t]]=P[X(t+h+1)|I(t)]$, $\forall t>\omega$, which means $Y \underset{h+1}{\nrightarrow} X|I$. Thus (i) $\Leftrightarrow$ (ii). The equivalence between (ii) and (iii) follows trivially from Definition 2.2. *Q.E.D.*

Proof of Proposition 2.4: Using Proposition 2.3, we can infer from the separation condition that $(Y',Z_2')' \underset{(\infty)}{\nrightarrow} (X',Z_1')'|I_{XZ_1}$, hence by Proposition 2.1, $(Y',Z_2')' \underset{(\infty)}{\nrightarrow} X|I_{XZ_1}$. Then, since $I_{XZ_1}(t)\subseteq I_{XZ}(t)$, we have $Y \underset{(\infty)}{\nrightarrow} X|I_{XZ}$. *Q.E.D.*

Proof of Theorem 3.1: Let $t\in\mathbb{Z}$, $t>\omega$, and $I(t)=I_{XZ}(t)$. We deduce from (3.6)–(3.9):

$$\begin{aligned} P[X(t+h)|I(t)+Y(-\infty,t]] &= \mu_{Xh}(t)+\sum_{j=1}^{\infty}[\pi_{XXj}^{(h)}X(t+1-j)+\pi_{XYj}^{(h)}Y(t+1-j)+\pi_{XZj}^{(h)}Z(t+1-j)] \end{aligned}$$

where $\mu_h(t)=\sum_{k=0}^{h-1}\pi_1^{(k)}\mu(t+h-k)=[\mu_{Xh}(t)',\mu_{Yh}(t)',\mu_{Zh}(t)']'$ and $\mu_{Xh}(t)$ has dimension $m_1\times 1$. Condition (3.10) then implies that

$$\begin{aligned} P[X(t+h)|I(t)+Y(-\infty,t]] &= \sum_{j=1}^{\infty}\left[\pi_{XXj}^{(h)}X(t+1-j)+\pi_{XZj}^{(h)}Z(t+1-j)\right]+\mu_h(t), \end{aligned}$$

and the components of $P[X(t+h)|I(t)+Y(-\infty,t]]$ thus all belong to the Hilbert space $I(t)=H+X(-\infty,t]+Z(-\infty,t]$. In other words, $P[X(t+h)|I(t)+Y(-\infty,t]]=P[X(t+h)|I(t)]$. Consequently, condition (3.10) is sufficient for $Y \underset{h}{\nrightarrow} X|I$.

Suppose now the matrices $E[a(t)a(t)']$ are nonsingular for $t>\omega$. If $Y \underset{h}{\nrightarrow} X|I$, all the components of $P[X(t+h)|I(t)+Y(-\infty,t]]$, belong to the Hilbert space $I(t)$, for $t>\omega$. Thus $P[X(t+h)|I(t)+Y(-\infty,t]]$, which can be written $P[X(t+h)|I(t)+Y(-\infty,t]]=\sum_{j=1}^{\infty}\pi_{X.j}^{(h)}\,W(t+1-j)$ where $\pi_{X.j}^{(h)}=[\pi_{XXj}^{(h)},\pi^{(h)}{}_{XYj},\pi_{XZj}^{(h)}]$, can also be expressed as the limit in quadratic mean (q.m.) of a sequence $U_T=\sum_{j=1}^{T}\phi^{(T)}{}_jW(t+1-j)$, $T\in\mathbb{N}$, where the components of U_T all belong to $I(t)$: $U_T=\sum_{j=1}^{T}[\phi_{Xj}^{(T)}X(t+1-j)+\phi_{Zj}^{(T)}Z(t+1-j)]$. Consequently, defining $\phi_j^{(T)}=0$ for $j>T$, and

$$\bar{U}_T(t)=P[X(t+h)|I(t)+Y(-\infty,t]]-U_T=\sum_{j=1}^{\infty}[\pi_{X.j}^{(h)}-\phi_j^{(T)}]W(t+1-j),$$

we see $\bar{U}_T(t)$ converges in q.m. to 0, hence $E[\bar{U}_T(t)a(t)']=[\pi_{X.1}^{(h)}-\phi_1^{(T)}]E[a(t)a(t)']\underset{T\to\infty}{\to}0$, because $E[W(t)a(t)']=E[a(t)a(t)']$ and $E[W(t+1-j)a(t)']=0$ for $j\geq 2$. Since the matrix $E[a(t)a(t)']$ is nonsingular, we must have $\pi_{X.1}^{(h)}-\phi_1^{(T)}\underset{T\to\infty}{\to}0$. And, since $\phi_1^{(T)}=[\phi_{X1}^{(T)},0,\phi_{Z1}^{(T)}]$, this implies that $\pi_{XY1}^{(h)}=0$. We thus see that $\sum_{j=2}^{\infty}[\pi_{X.j}^{(h)}-\phi_j^{(T)}]W(t+1-j)$ converges in q.m. to zero and a similar argument (with $t-1>\omega$) yields $\pi_{XY2}^{(h)}=0$. Proceeding analogously for increasing j, we get: $\pi_{XYj}^{(h)}=0$ for $j=1,2,\dots$. *Q.E.D.*

PROOF OF LEMMA 3.1: We shall prove (3.12) in two steps: first, for $p = 2$ (for any $h \geq 2$), and then by recurrence on h for $2 < p \leq h$. Let $p = 2 \leq h$. From (3.8), it follows that

$$\pi_{XYj}^{(2)} = \pi_{XX1}\pi_{XYj} + \pi_{XY1}\pi_{YYj} + \pi_{XZ1}\pi_{ZYj} + \pi_{XY,j+1}^{(2)}, \qquad \forall j \geq 1,$$

and

$$\pi_{XZ1}^{(h)} = \pi_{XX1}^{(h-1)}\pi_{XZ1} + \pi_{XY1}^{(h-1)}\pi_{YZ1} + \pi_{XZ1}^{(h-1)}\pi_{ZZ1} + \pi_{XZ2}^{(h-1)}.$$

From the assumption that $\pi_{XYj}^{(k)} = 0$, $\forall j \in \mathbb{N}$, for $k = 1, \ldots, h$, we then have $\pi_{XYj} = \pi_{XYj}^{(2)} = \pi_{XYj}^{(h-1)} = 0$ for all j, so that $\pi_{XYj}^{(2)} = \pi_{XZ1}\pi_{ZYj} = 0$, and

$$\begin{aligned}\pi_{XZ1}^{(h)}\pi_{ZYj} &= [\pi_{XX1}^{(h-1)}\pi_{XZ1} + \pi_{XZ1}^{(h-1)}\pi_{ZZ1} + \pi_{XZ2}^{(h-1)}]\pi_{ZYj}\\ &= [\pi_{XZ1}^{(h-1)}\pi_{ZZ1} + \pi_{XZ2}^{(h-1)}]\pi_{ZYj}, \qquad \forall j \in \mathbb{N},\end{aligned}$$

which is identical with (3.12) with $p = 2$.

Let us now call $P(h)$ the property obtained when (3.12) holds for all integers p such that $2 \leq p \leq h$. From the first step above, it is clear that $P(h)$ holds for $h = 2$ (since then we must have $p = 2$). Take now $h \geq 3$ (otherwise, the proof is complete), and suppose that $P(k)$ holds for $k = 2, \ldots, h-1$. We need to show that $P(h)$ then also holds, i.e., we need to prove (3.12) for all integers p such that $2 \leq p \leq h$. From the first step, we know that (3.12) holds for $p = 2$. By the mathematical induction principle, it will suffice to show that (3.12) must hold for $p = \bar{p} + 1$ whenever it does for $p = 1, \ldots, \bar{p}$ (where $\bar{p} < h$). If we assume (3.12) for $p = 2, \ldots, \bar{p} < h$ and take p to be any integer such that $2 \leq p \leq \bar{p}$, we have

$$\pi_{XZ1}^{(h)}\pi_{ZYj} = \left\{\sum_{l=1}^{p} \pi_{XZl}^{(h-p+1)}\left[\sum_{J(p-l)} \prod_i \pi_{ZZi}^{n_i}\right]\right\}\pi_{ZYj}, \qquad \forall j \in \mathbb{N},$$

where we write (to simplify the notation) $\Sigma_{J(l)}\Pi_i \pi_{ZZi}^{n_i} = \Sigma_{J(l)}[\Pi_{i=1}^{l}\pi_{ZZi}^{n_i}]$. But, by (3.8), we have $\pi_l^{(h-p+1)} = \pi_1^{(h-p)}\pi_l + \pi_{l+1}^{(h-p)}$, hence

$$\pi_{XZl}^{(h-p+1)} = \pi_{XX1}^{(h-p)}\pi_{XZl}\pi_{XY1}^{(h-p)}\pi_{YZl} + \pi_{XZ1}^{(h-p)}\pi_{ZZl} + \pi_{XZ,l+1}^{(h-p)}.$$

Since $\pi_{XYj}^{(k)} = 0$, $\forall j \in \mathbb{N}$, for $k = 1, \ldots, h$, we have $\pi_{XY1}^{(h-p)} = 0$, so that

$$\text{(A.3)} \qquad \begin{aligned}&\left\{\sum_{l=1}^{p} \pi_{XZl}^{(h-p+1)}\left[\sum_{J(p-l)} \prod_i \pi_{ZZi}^{n_i}\right]\right\}\pi_{ZYj}\\ &= \pi_{XX1}^{(h-p)}\left\{\sum_{l=1}^{p} \pi_{XZl}\left[\sum_{J(p-l)} \prod_i \pi_{ZZi}^{n_i}\right]\right\}\pi_{ZYj}\\ &\quad + \left\{\sum_{l=1}^{p} [\pi_{XZ,l+1}^{(h-p)} + \pi_{XZ1}^{(h-p)}\pi_{ZZl}] \sum_{J(p-l)} \prod_i \pi_{ZZi}^{n_i}\right\}\pi_{ZYj}.\end{aligned}$$

We will now show that the first term of the latter sum is zero. Since $p \leq h-1$, we know that $P(p)$ must hold, i.e., $\pi_{XZ1}^{(p)}\pi_{ZYj} = [\Sigma_{l=1}^{q}\pi_{XZl}^{(p-q+1)}\Sigma_{J(q-l)}\Pi_i \pi_{ZZi}^{n_i}]\pi_{ZYj}$, for $2 \leq q \leq p$ (by the recurrence assumption). In particular, by taking $q=p$, we get $\pi_{XZ1}^{(p)}\pi_{ZYj} = [\Sigma_{l=1}^{p}\pi_{XZl}\Sigma_{J(p-l)}\Pi_i \pi_{ZZi}^{n_i}]\pi_{ZYj}$. Since $\pi_{XYj}^{(k)} = 0$, $\forall j \in \mathbb{N}$, for $k=1,\dots,p+1$ (for $p+1 \leq h$), we must have $\pi_{XZ1}^{(p)}\pi_{ZYj} = 0$ (see (3.11)), which implies that the first term of (A.3) is zero. Consequently,

$$\pi_{XZ1}^{(h)}\pi_{ZYj} = \left\{ \sum_{l=1}^{p} [\pi_{XZ,l+1}^{(h-p)} + \pi_{XZ1}^{(h-p)}\pi_{ZZl}] \sum_{J(p-l)} \prod_i \pi_{ZZi}^{n_i} \right\} \pi_{ZYj}$$

$$= \left\{ \sum_{l=2}^{p+1} \pi_{XZl}^{(h-p)} \left[\sum_{J(p+1-l)} \prod_i \pi_{ZZi}^{n_i} \right] \right\} \pi_{ZYj}$$

$$+ \pi_{XZ1}^{(h-p)} \left\{ \sum_{l=1}^{p} \pi_{ZZl} \left[\sum_{J(p-l)} \prod_i \pi_{ZZi}^{n_i} \right] \right\} \pi_{ZYj}.$$

But it is clear from the definition of $J(p)$ that $\Sigma_{l=1}^{p}\pi_{ZZl}[\Sigma_{J(p-l)}\Pi_i \pi_{ZZi}^{n_i}] = \Sigma_{J(p)}\Pi_i \pi_{ZZi}^{n_i}$, hence $\pi_{XZ1}^{(h)}\pi_{ZYj} = \{\Sigma_{l=1}^{p+1}\pi_{XZl}^{(h-p)}[\Sigma_{J(p+1-l)}\Pi_i \pi_{ZZi}^{n_i}]\}\pi_{ZYj}$, which means that (3.12) holds with p replaced by $p+1$. We can take $p=\bar{p}$, so that (3.12) holds for $p=2,3,\dots,\bar{p}+1$, hence (by recurrence) for $p=2,\dots,h$. Property $P(h)$ is thus established. *Q.E.D*

PROOF OF LEMMA 3.2: We need to show that the equivalence between (3.13), (3.14), and (3.15) holds for any $h \geq 2$. Again we shall proceed by recurrence considering first the equivalence between (3.13) and (3.14), and then the one between (3.13) and (3.15).

The equivalence between (3.13) and (3.14) follows by applying the recursion (3.11), which is implied by (3.8). For $h=2$, the result clearly holds since $\pi_{XYj}^{(1)} = \pi_{XYj}$, and $\pi_{XYj} = 0$ for all j entails (on applying (3.11)): $\pi_{XYj}^{(2)} = \pi_{XZ1}^{(1)}\pi_{ZYj}$, $\forall j \in \mathbb{N}$. Suppose now that the equivalence holds for some $h \geq 2$. Then, given that $\pi_{XYj}^{(k)} = 0$, $\forall j \in \mathbb{N}$, for $k=1,\dots,h$, it follows from (3.11) that $\pi_{XYj}^{(h+1)} = \pi_{XZ1}^{(h)}\pi_{ZYj}$, $\forall j \in \mathbb{N}$, and the equivalence holds for $h+1$. The equivalence between (3.13) and (3.14) for any $h \geq 2$ follows by recurrence.

The equivalence between (3.13) and (3.15) holds for $h=2$ because the criteria (3.14) and (3.15) are then identical. Suppose now that the equivalence between (3.13) and (3.15) holds for some $h \geq 2$. Then given that $\pi_{XYj}^{(k)} = 0$, $\forall j \in \mathbb{N}$, $k=1,\dots,h$, we see from (3.11) and Lemma 3.1 that $\pi_{XYj}^{(h+1)} = \pi_{XZ1}^{(h)}\pi_{ZYj} = R_{XZ}^{(h)}\pi_{ZYj}$, $\forall j \in \mathbb{N}$, and the equivalence between (3.13) and (3.15) also holds for $h+1$. Conditions (3.13) and (3.15) are thus equivalent, $\forall h \geq 2$. *Q.E.D.*

PROOF OF THEOREM 3.2: Under the assumptions (3.1) to (3.3), it follows from Theorem 3.1 that condition (3.13) is sufficient for $Y \underset{(h)}{\nrightarrow} X|I_{XZ}$. Further, under these conditions, the recursion (3.8) applies so that (3.14) and (3.15) are each equivalent to (3.13) by Lemma 3.1. When $W(t)$ is a regular process, Theorem 3.1 also entails that (3.13), hence also (3.14) and (3.15), is necessary and sufficient for $Y \underset{(h)}{\nrightarrow} X|I_{XZ}$. *Q.E.D.*

PROOF OF COROLLARY 3.3: To prove this corollary we shall need the following lemma relating the matrix $R_{XZ}^{(k)}$ to the formal series $\pi_{XZ}(z)$ and $\pi_{ZZ}(z)$.

LEMMA A.3: *If we denote by $R_{XZ}(z) = \Sigma_{k=1}^{\infty}R_{XZ}^{(k)}z^k$ the formal series associated with the coefficients $R_{XZ}^{(k)}$ defined in Lemma 3.2, then $R_{XZ}(z) \equiv -\pi_{XZ}(z)\pi_{ZZ}(z)^{-1}$.*

PROOF OF LEMMA A.3: From the definitions (3.16), we see that $\pi_{ZZ}(z) = I_{m_3} - \sum_{k=1}^{\infty} \pi_{ZZk} z^k$ and $\pi_{XZ}(z) = -\sum_{j=1}^{\infty} \pi_{XZj} z^j$, hence

$$[\pi_{ZZ}(z)]^{-1} = I_{m_3} + \sum_{l=1}^{\infty} \left[\sum_{k=1}^{\infty} \pi_{ZZk} z^k \right]^l = \sum_{k=0}^{\infty} \left\{ \sum_{J(k)} \left[\prod_{i=1}^{k} \pi_{ZZi}^{n_i} \right] \right\} z^k,$$

where $J(k)$, $k \in \mathbb{N}_0$, is defined as in Lemma 3.1, and

$$-\pi_{XZ}(z)[\pi_{ZZ}(z)]^{-1} = \sum_{k=1}^{\infty} \left\{ \sum_{l=1}^{k} \pi_{XZl} \left[\sum_{J(k-l)} \left(\prod_{i=1}^{k-l} \pi_{ZZi}^{n_i} \right) \right] \right\} z^k$$

$$= \sum_{k=1}^{\infty} R_{XZ}^{(k)} z^k = R_{XZ}(z). \qquad Q.E.D.$$

PROOF OF COROLLARY 3.4: By Theorem 3.2, the condition $\pi_{XYj} = 0$, $\forall j \geq 1$, jointly with $R_{XZ}^{(k)} \pi_{ZYj} = 0$, $\forall j \geq 1$, $\forall k \geq 1$, is necessary and sufficient for $Y \underset{(\infty)}{\nrightarrow} X | I_{XZ}$. This condition can be written in terms of generating functions, as follows: $\pi_{XY}(z) \equiv 0$ and $R_{XZ}(z)\pi_{ZYj} \equiv 0$, $\forall j \geq 1$. If we multiply $R_{XZ}(z)\pi_{ZYj}$ by z^j and sum over $j \in \mathbb{N}$, we get the following necessary condition for $Y \underset{(\infty)}{\nrightarrow} X | I_{XZ}$: $\pi_{XY}(z) \equiv 0$ and $R_{XZ}(z)\pi_{ZY}(z) \equiv 0$. Using Lemma A.3, this can be written: $\pi_{XY}(z) \equiv 0$ and $\pi_{XZ}(z)\pi_{ZZ}(z)^{-1}\pi_{XY}(z) \equiv 0$, hence (3.21).

Let us now assume $z \in \mathbb{C}$ and $|z| < \delta$, so that $\pi(z)$ converges in $\mathbb{C}$ for $|z| < \delta$. To show (3.21) and (3.22) are equivalent, it is then sufficient to use standard formulae for inverting partitioned matrices (where we omit the symbol z to simplify the notation): $\psi_{XY} = -\pi_{XX.Z}^{-1}[\pi_{XY} - \pi_{XZ}\pi_{ZZ}^{-1}\pi_{ZY}]\tilde{\pi}_{YY}^{-1}$, where $\pi_{XX.Z} = \pi_{XX} - \pi_{XZ}\pi_{ZZ}^{-1}\pi_{ZX}$,

$$\tilde{\pi}_{YY} = \pi_{YY} - [\pi_{YX}, \pi_{YZ}]A^{-1} \begin{bmatrix} \pi_{XY} \\ \pi_{ZY} \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} \pi_{XX} & \pi_{XZ} \\ \pi_{ZX} & \pi_{ZZ} \end{bmatrix}.$$

Since $\pi(0) = I_m$ (a nonsingular matrix), the inverses of the matrices $\pi_{ZZ}, A, \pi_{XX.Z}$, and $\tilde{\pi}_{YY}$ all exist in a sufficiently small disk centered at zero, say for $|z| < \bar{\delta}$ (where $0 < \bar{\delta} < 1$); for a similar argument, see Dufour and Tessier (1993, Proposition 1). For $|z| < \bar{\delta}$, we see that: $\pi_{XZ}(z)\pi_{ZZ}(z)^{-1}\pi_{ZY}(z) \equiv 0$ and $\pi_{XY}(z) \equiv 0 \Leftrightarrow \psi_{XY}(z) \equiv 0$ and $\pi_{XY}(z) \equiv 0$, which shows (3.21) and (3.22) are equivalent. The equivalence between (3.22) and (3.23) is deduced from the one between (3.21) and (3.22) on permuting π and ψ, as these play symmetric roles in (3.22) *Q.E.D.*

PROOF OF COROLLARY 3.5: When $Z(t)$ is univariate, $R_{XZ}(z) = -\pi_{XZ}(z)\pi_{ZZ}(z)^{-1} = [R_{x_1 Z}(z), \ldots, R_{x_{m_1} Z}(z)]'$ is an $m_1 \times 1$ vector of scalar formal series $R_{x_i Z}(z)$ in z, while $\pi_{ZY}(z) = [\pi_{Zy_1}(z), \ldots, \pi_{Zy_{m_2}}(z)]$ is a $1 \times m_2$ vector of scalar formal series $\pi_{Zy_j}(z)$ in z. Then the condition $\pi_{XZ}(z)\pi_{ZZ}(z)^{-1}\pi_{ZY}(z) \equiv 0$ in Corollary 3.4 means that $R_{x_i Z}(z)\pi_{Zy_j}(z) \equiv 0$, for $i = 1, \ldots, m_1$ and $j = 1, \ldots, m_2$. If we now interpret z as a complex number ($z \in \mathbb{C}$), the assumption that $\pi(z)$ converges for $|z| < \delta$ implies that the series $\pi(z)^{-1}$ must converge in a circle $|z| < \bar{\delta}$, where $\bar{\delta} > 0$ (because $\pi(0) = I_m$ is nonsingular). Then the series $R_{x_i Z}(z)$ and $\pi_{Zy_j}(z)$ represent functions which are analytic in the circle $|z| < \bar{\delta}$, and $R_{x_i Z}(z)\pi_{Zy_j}(z) \equiv 0$ can hold only if either $R_{x_i Z}(z) \equiv 0$ or $\pi_{Zy_j}(z) \equiv 0$. Consequently, we must have: $R_{x_i Z}^{(k)} \pi_{Zy_{lj}} = 0$, $\forall k \in \mathbb{N}$, $\forall j \in \mathbb{N}$, for $i = 1, \ldots, m_1$ and $l = 1, \ldots, m_2$, or equivalently, $R_{XZ}^{(k)} \pi_{ZYj} \equiv 0$, $\forall k \in \mathbb{N}$, $\forall j \in \mathbb{N}$. Thus the condition (3.21) implies (3.13) when $Z(t)$ is univariate, and Corollary 3.5 follows from Theorem 3.2 and Corollary 3.4. *Q.E.D.*

PROOF OF COROLLARY 3.6: The result follows from condition (3.21) as in the proof of Corollary 3.5. When $Z(t)$ is univariate, $\pi_{XZ}(z)\pi_{ZZ}(z)^{-1}\pi_{ZY}(z) \equiv 0 \Leftrightarrow \pi_{XZ}(z)\pi_{ZY}(z) \equiv 0 \Leftrightarrow \pi_{XZ}(z) \equiv 0$ or $\pi_{ZY}(z) \equiv 0$, where the last equivalence follows from the observation that the product of two analytic functions is zero only if one of the functions is zero. *Q.E.D.*

PROOF OF THEOREM 3.3: Let us call $Q(\overline{H})$ the property obtained when $\pi_{XY1}^{(h)} = 0$, for $h = 2,3,\ldots,H$, with $H = \overline{H}$. By (3.7), $Q(H)$ is equivalent to $\tilde{Q}(H)$: $\pi_{XYh} = -\sum_{q=1}^{h-1}\pi_{X.q}\pi_{.Y1}^{(h-q)}$, for $h = 2,3,\ldots,H$. Thus $Q(2)$ [or $\tilde{Q}(2)$] holds by choosing $\pi_{XY2} = -\pi_{X.1}\pi_{.Y1}$. Then to prove that $Q(H)$ may hold (for a convenient choice of π_{XYj}, $j = 2,3,\ldots,H$) for all integers $H \geq 2$, it will be sufficient, by mathematical induction, to show that $\tilde{Q}(H)$ may hold for $H = \overline{H}+1$ (with an appropriate choice of $\pi_{XY,\overline{H}+1}$), whenever it does for $H = 2,3,\ldots,\overline{H}$ (where $\overline{H} > 2$). But, assuming $\tilde{Q}(H)$ for $H = 2,3,\ldots,\overline{H}$ is tantamount to assuming $\tilde{Q}(\overline{H})$, i.e., a set of $\overline{H}$ equalities which involve only the matrices π_j, $j = 1,2,\ldots,\overline{H}$. To see this, it is sufficient to show that $\pi_1^{(h)}$ is a well-defined function of $\pi_1, \pi_2, \ldots, \pi_h$, for all positive integers h. But this last statement is a straightforward corollary of the recursion (3.8): $\pi_1^{(h)} = \pi_2^{(h-1)} + \pi_1^{(h-1)}\pi_1$. Therefore we are able to ensure $\tilde{Q}(\overline{H}+1)$ by choosing $\pi_{XY,\overline{H}+1} = -\sum_{q=1}^{\overline{H}}\pi_{X.q}\pi_{.Y1}^{(h-q)}$, since $\pi_{\overline{H}+1}$ is not constrained by the $\overline{H}$ equalities $\tilde{Q}(\overline{H})$. The first part of Theorem 3.3 is thus established and the proof will be complete if we prove that $Q(H)$ (or $\tilde{Q}(H)$) entails $\pi_{XY2}^{(h)} = -\sum_{q=1}^{h}\pi_{X.q}\pi_1^{(h-q)}\pi_{.Y1}$, for $h = 1,2,\ldots,H-1$. But, by the recursion (3.8), $\pi_{XY2}^{(h)} = \pi_{XY1}^{(h+1)} - \pi_{X.1}^{(h)}\pi_{.Y1} = -\pi_{X.1}^{(h)}\pi_{.Y1}$, for $h = 1,2,\ldots,H-1$, if $Q(H)$ holds. The proof is then complete on noting that, by (3.7), $\pi_{X.1}^{(h)} = \pi_{X.h} + \sum_{q=1}^{h-1}\pi_{X.q}\pi_1^{(h-q)}$. *Q.E.D.*

PROOF OF PROPOSITION 4.5: To prove this proposition we shall use the following lemma on power series, which generalizes a property used by Lütkepohl (1993) for a similar problem; for a proof of this lemma, see Dufour and Tessier (1996).

LEMMA A.4: *Let $f(z) = \sum_{j=0}^{\infty}a_j z^j$ be a complex-valued power series in $z \in \mathbb{C}$ (with $a_j \in \mathbb{C}$ for all j), convergent for $|z| < \delta$, and such that $\sum_{j=0}^{\infty}a_j z^j = (\sum_{k=0}^{p}b_k z^k)(\sum_{l=0}^{\infty}c_l z^l)$ for $|z| < \delta$, where $\delta > 0$, $0 \leq p < \infty$, $c_0 = 1$ and the power series $\sum_{l=0}^{\infty}c_l z^l$ converges for $|z| < \delta$. Then: $a_j = 0, \forall j \geq 0 \Leftrightarrow a_j = 0$ for $j = 0,1,\ldots,p$.*

Given Theorem 3.2, we now prove Proposition 4.5 by showing that $R_{XZ}^{(k)}\pi_{ZYj} = 0$, $\forall j \geq 1$, for $k = 1,2,\ldots,m_3 p$, implies $R_{XZ}^{(k)}\pi_{ZYj} = 0$, $\forall j \geq 1$, $\forall k \geq 1$, where the latter identity is equivalent to $R_{XZ}(z)\pi_{ZYj} \equiv 0$, $\forall j \geq 1$. From Lemma A.3, we have

$$R_{XZ}(z)\pi_{ZYj} = -\pi_{XZ}(z)\pi_{ZZ}(z)^{-1}\pi_{ZYj}$$

$$= -\{\det[\pi_{ZZ}(z)]\}^{-1}\pi_{XZ}(z)\pi_{ZZ}^{*}(z)\pi_{ZYj},$$

where $\pi_{ZZ}^{*}(z)$ is the transposed of the matrix of cofactors of $\pi_{ZZ}(z)$. Since $\pi_{ZZ}(z)$ is an $m_3 \times m_3$ matrix whose elements are polynomials of degree not greater than p (by assumption), $\pi_{ZZ}^{*}(z)$ is a matrix of polynomials of degree not greater than $(m_3 - 1)p$. Let us now consider a given pair $(x_i(t), y_l(t))'$ of components of $X(t) = [x_1(t),\ldots,x_{m_1}(t)]'$ and $Y(t) = [y_1(t),\ldots,Y_{m_2}(t)]'$, where $1 \leq i \leq m_1$ and $1 \leq l \leq m_2$. We have $R_{x_iZ}(z)\pi_{Zy_lj} = -\{\det[\pi_{ZZ}(z)]\}^{-1}\pi_{x_iZ}(z)\pi_{ZZ}^{*}(z)\pi_{Zy_lj}$, where $\{\det[\pi_{ZZ}(z)]\}^{-1} = \sum_{l=0}^{\infty}c_l z^l$ is a formal series with $c_0 = 1$ and $\pi_{x_iZ}(z)\pi_{ZZ}^{*}(z)\pi_{Zy_lj}$ is a polynomial of degree not greater than $m_3 p$. Using Lemma A.4, we can then state that $R_{x_iZ}(z)\pi_{Zy_lj} = 0$, $\forall j \geq 1$, if and only if, for $j \geq 1$, the coefficients associated with powers of z not greater than $m_3 p$ in $R_{x_iZ}(z)\pi_{Zy_lj}$ are equal to zero: $R_{x_iZ}^{(k)}\pi_{Zy_lj} = 0$, for $k = 1,2,\ldots,m_3 p$. The result then follows from

Theorem 3.2. *Q.E.D.*

PROOF OF PROPOSITION 4.6: For each pair $(x_i(t), y_l(t))'$, $1 \le i \le m_1$, $1 \le l \le m_2$, we have

$$\pi_{x_iZ}(z)\pi_{ZZ}(z)^{-1}\pi_{Zy_l}(z) \equiv 0 \Leftrightarrow \{\det[\pi_{ZZ}(z)]\}^{-1}\pi_{x_iZ}(z)\pi_{ZZ}^*(z)\pi_{Zy_l}(z) \equiv 0,$$

where $\pi_{x_iZ}(z)\pi_{ZZ}^*(z)\pi_{Zy_l}(z)$ is a polynomial of degree not greater than $(m_3+1)p$. The result then follows on applying Lemma A.4. *Q.E.D.*

REFERENCES

ANDREWS, D. W. K. (1987): "Asymptotic Results for Generalized Wald Tests," *Econometric Theory*, 3, 348–358.

BEKAERT, G., AND R. J. HODRICK (1992): "Characterizing Predictable Components in Excess Returns on Equity and Foreign Exchange Markets," *Journal of Finance*, 47, 467–509.

BOUDJELLABA, H., J.-M. DUFOUR, AND R. ROY (1992): "Testing Causality Between Two Vectors in Multivariate ARMA Models," *Journal of the American Statistical Association*, 87, 1082–1090.

——— (1994): "Simplified Conditions for Non-Causality between Two Vectors in Multivariate ARMA Models," *Journal of Econometrics*, 63, 271–287.

BOUISSOU, M. B., J. J. LAFFONT, AND Q. H. VUONG (1986): "Tests of Noncausality Under Markov Assumptions for Qualitative Panel Data," *Econometrica*, 54, 395–414.

BRUNEAU, C., AND J.-P. NICOLAÏ (1992): "Probabilistic Foundations of Causal Analysis in a Stationary Vectorial Autoregressive Model," Technical Report, CREST-ENSAE, Paris.

CAMPBELL, J. Y., AND R. J. SHILLER (1988): "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Journal of Financial Studies*, 1, 195–228.

CORDIER, J., AND J. P. INDJEHAGOPIAN (1986): "Multidimensional Analysis of a Commodity Price System," *International Journal of Forecasting*, 2, 153–189.

DUFOUR, J.-M., AND E. RENAULT (1994): "Short-Run and Long-Run Causality in Time Series: Theory," Technical Report, Centre de Recherche et Développement en Économique, Université de Montréal, and Institut d'Economie Industrielle, Université des Sciences Sociales de Toulouse.

——— (1995): "Short-Run and Long-Run Causality in Time Series: Inference," Technical Report, Centre de Recherche et Développement en Économique, Université de Montréal, and Institut d'Economie Industrielle, Université des Sciences Sociales de Toulouse.

DUFOUR, J.-M., AND D. TESSIER (1993): "On the Relationship Between Impulse Reponse Analysis: Innovation Accounting and Granger Causality," *Economics Letters*, 42, 327–333.

——— (1996): "Parsimonious Autoregressive Conditions for Noncausality in Multivariate ARMA Models," Technical Report, C.R.D.E., Université de Montréal.

FAMA, E. F., AND K. R. FRENCH (1988): "Permanent and Temporary Components of Stock Prices," *Journal of Political Economy*, 96, 246–273.

FLORENS, J.-P., AND M. MOUCHART (1982): "A Note on Noncausality," *Econometrica*, 50, 583–606.

——— (1985): "A Linear Theory for Non-Causality," *Econometrica*, 53, 157–175.

GEWEKE, J. (1984): "Inference and Causality in Economic Time Series," in *Handbook of Econometrics, Volume 2*, ed. by Z. Griliches and M. D. Intrilligator. Amsterdam: North-Holland, pp. 1102–1144.

GOURIÉROUX, C., AND A. MONFORT (1990): *Séries Temporelles et Modèles Dynamiques*. Paris: Economica.

GRANGER, C. W. J. (1969): "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," *Econometrica*, 37, 424–459.

GRANGER, C. W. J., AND J.-L. LIN (1995): "Causality in the Long Run," *Econometric Theory*, 11, 530–536.

HOSOYA, Y. (1977): "On the Granger Condition for Non-Causality," *Econometrica*, 45, 1735–1736.

——— (1991): "The Decomposition and Measurement of the Interdependence Between Second-Order Stationary Processes," *Probability Theory and Related Fields*, 88, 429–444.

HSIAO, C. (1982): "Autoregressive Modeling and Causal Ordering of Economic Variables," *Journal of Economic Dynamics and Control*, 4, 243–259.

KOHN, R. (1981): "A Characterization of Granger-Sims Exogeneity," *Economics Letters*, 8, 129–133.

LÜTKEPOHL, H. (1991): *Introduction to Multiple Time Series Analysis*. Berlin: Springer-Verlag.

——— (1993): "Testing for Causation Between Two Variables in Higher Dimensional VAR Models," in *Studies in Applied Econometrics*, ed. by H. Schneeweiss and K. Zimmerman. Heidelberg: Springer-Verlag.

NEWBOLD, P. (1982): "Causality Testing in Economics," in *Time Series Analysis: Theory and Practice* 1, ed. by O. D. Anderson. Amsterdam: North-Holland.

PIERCE, D. A., AND L. D. HAUGH (1977): "Causality in Temporal Systems: Characterizations and Survey," *Journal of Econometrics*, 5, 265–293.

POTERBA, J. M., AND L. H. SUMMERS (1988): "Mean Reversion in Stock Prices: Evidence and Implications," *Journal of Financial Economics*, 22, 27–59.

RENAULT, E., AND A. SZAFARZ (1991): "True Versus Spurious Instantaneous Causality," Technical Report 9103, Centre d'Economie Mathématique, Université Libre de Bruxelles.

SIMS, C. (1980): "Macroeconomics and Reality," *Econometrica*, 48, 1–48.

WIENER, N. (1956): "The Theory of Prediction," in *The Theory of Prediction*, ed. by E. F. Beckenback. New York: McGraw-Hill, Chapter 8.