Université de Montréal
Faculté des arts et des sciences
Département de sciences économiques

CAHIER 9538

SHORT-RUN AND LONG-RUN CAUSALITY IN TIME SERIES : THEORY

Jean-Marie DUFOUR[1] and Eric RENAULT[2]

1 Département de sciences économiques and Centre de recherche et développement en économique (C.R.D.E.), Université de Montréal.

2 GREMAQ and Institut d'économie industrielle, Université des Sciences Sociales, France.

September 1995

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 ministère des Relations internationales). The authors thank John Galbraith, Eric Ghysels, Helmut Lütkepohl, Michel Mouchart, Saïd Nsiri, Franz Palm, Pierre Perron, Roch Roy, Peter Schotmann, Lanh Tran and David Tessier for several helpful comments and discussions.

C.P. 6128, succursale Centre-ville
Montréal (Québec) H3C 3J7

Télécopieur (FAX): (514) 343-5831
Courrier électronique (E-Mail): econo@tornade.ERE.Umontreal.CA

Ce cahier a également été publié au Centre de recherche et développement en économique (C.R.D.E.) (publication no 3595).

Dépôt légal - 1995
Bibliothèque nationale du Québec
Bibliothèque nationale du Canada

ISSN 0709-9231

RÉSUMÉ

La causalité au sens de Wiener-Granger est un concept qui réfère à la prévisibilité d'un vecteur de variables aléatoires par un autre vecteur, une période à l'avance. Récemment, Lütkepohl (1993) a proposé de définir la non-causalité entre deux variables aléatoires par la non-prévisibilité à tous les horizons. Si plus de deux vecteurs de variables sont utilisés pour calculer les prévisions (i.e., lorsque l'ensemble d'information comprend des variables auxiliaires), ces deux notions ne sont pas équivalentes. Dans ce texte, nous généralisons d'abord la notion de causalité en considérant la causalité à n'importe quel horizon *h*. Nous donnons des conditions nécessaires et suffisantes pour la non-causalité entre deux vecteurs aléatoires (à l'intérieur d'un vecteur plus grand) jusqu'à n'importe quel horizon *h*, où *h* peut être infini. En particulier, pour des processus possiblement non stationnaires dont les seconds moments sont finis, nous dérivons des conditions relativement simples d'**exhaustivité** et de **séparabilité** qui sont suffisantes pour la non-causalité à tous les horizons. Pour traiter les cas où ces conditions ne s'appliquent pas, nous considérons une classe plus particulière, bien qu'encore très générale, de processus vectoriels autorégressifs (possiblement d'ordre infini, stationnaire ou non stationnaire), incluant comme cas spéciaux les processus ARIMA multivariés. Nous donnons des caractérisations paramétriques générales de la non-causalité à différents horizons pour cette classe, incluant une caractérisation en termes de **chaînes de causalité**. Nous introduisons aussi la notion de **décomposition causale séparante** et utilisons ce concept pour obtenir des **conditions de séparabilité généralisées** qui sont à la fois nécessaires et suffisantes pour la non-causalité à différents horizons.

Mots clés : causalité, long terme, chaîne de causalité, autorégression vectorielle, processus ARMA multivarié, test.

ABSTRACT

Causality in the sense of Granger is typically defined in terms of predictability of a vector of variables one period ahead. Recently, Lütkepohl (1993) proposed to define non-causality between two variables in terms of non-predictability 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 non-causality between vectors of variables (inside a larger vector) up to any given horizon *h*, where *h* can be infinite. In particular, for general possibly non-stationary processes with finite second moments, relatively simple **exhaustivity** and **separation** conditions, which are sufficient for non-causality 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 non-stationary), which include multivariate ARIMA processes, and we derive general parametric characterizations of non-causality at various horizons for this class (including a **causality chain** characterization). The notion of **separating causal decomposition** is also introduced and used to obtain **generalized separation conditions** which are both necessary and sufficient for non-causality at various horizons in the same setup.

Key words : causality, long run, causality chain, vector autoregression, multivariate ARMA model, testing.

1 Introduction

The concept of causality introduced by Wiener (1956) and Granger (1969) is now a basic notion for studying the dynamic relationships between time series. The literature on this topic is considerable ; see, for example, the reviews of Pierce and Haugh (1979), 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 predict $X(t+1)$ when the corresponding observations on X and Z are available $(X(\tau), Y(\tau) : \tau \leq t)$; 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 issue, 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 through they are useless to predict $X(t+1)$. This is due to the fact that Y may help to predict Z one period ahead, which in turn has an effect on 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".

The purpose of the present paper is to study and characterize more precisely such effects. First, in Section 2, we define more general notions of causality that will allow us to study these issues : 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. Then we present several general results on causality up to any given horizon. In particular, we give a **componentwise 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 non-causality at horizon one is equivalent to non-causality at all horizons. We show that 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 the one where 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 non-causality proposed by Hsiao (1982) for systems with more than three variables ; Hsiao's condition, however, is **not generally necessary** for non-causality 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 non-causality 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 non-causality 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 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, of course, include as special cases autoregressive processes of finite order (VAR), stationary or non-stationary, a wide class of second-order stationary processes (including long-memory processes, such as fractional processes), and invertible ARMA processes. It is not required that the covariance matrices of the innovations be constant (i.e., heteroskedastic innovations are allowed). The results presented generalize considerably 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. After deriving a convenient formula for computing forecasts at different horizons, we give several characterizations of non-causality at these horizons : for regular processes (i.e., processes with non-singular innovation covariance matrices), we give necessary and sufficient conditions, while for non-regular processes we show that the same conditions are sufficient. In particular, we give a characterization of non-causality 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 non-causality, we also derive (as a corollary) necessary conditions for non-causality at all horizons which involve coefficients of the moving average representation of the process, illustrating the link between our definition of non-causality 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.

In Section 4, we consider the important case of finite order VAR processes, stationary or non-stationary, and show that the characterizations of non-causality obtained for infinite order autoregressive processes reduce in such cases to finite sets of restrictions. These restrictions may then be used for implementing tests. To obtain simpler conditions for testing as well to throw more light on the causality notions considered, we also derive in Section 5 **generalized separation conditions** which are sufficient (for non-causality) and intuitively appealing. These involve the introduction of the notion of **separating causal decomposition** by which the auxiliary variables Z of the system are decomposed into factors which do not cause X at horizon one and factors which do cause X at horizon one. Causality at horizons greater than one may then be characterized relatively simply in terms of these underlying causal factors. In particular, for the special case of systems which include only one auxiliary variable, we show that the **separation condition**, which was shown to be sufficient under very general assumptions (Section 2) is also necessary for non-causality at all horizons. In other words, if the auxiliary variable vector Z has only one component, there are only two cases that can make Y not to cause X at

all horizons : the one where Y causes $(X', Z)'$ at horizon 1 and the one where $(Y', Z)'$ does not cause X at horizon 1 (note that X and Y can be vectors). This separation criterion, when Z is univariate, coincides with the definition of non-causality proposed by Hsiao (1982). In addition to this separation property, we give various rank conditions on the matrices of the autoregressive representation of the process, which are necessary to have non-causality at all horizons. These yield a relatively simple way of showing that an assumption of non-causality at all horizons cannot hold.

In Section 6 finally, we make a number of concluding remarks and mention briefly the inference problems associated with the causality concepts discussed above. Methods for testing non-causality at different horizons and a number of applications are available in a separate paper [Dufour and Renault (1994)].

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, non-causality is defined in terms of orthogonality conditions between closed subspaces of a Hilbert space of real random variables with finite second moments. We denote $L^2 = L^2(\Omega, \mathcal{A}, Q)$ this Hilbert space of random variables defined on a common probability space $(\Omega, \mathcal{A}, Q)$, with covariance as the 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 integers. We consider a non-decreasing sequence I of such subspaces, i.e.

(2.1) $$I = \{I(t) \,:\, t \in \mathbb{Z},\, t > \omega\} \quad \text{where} \quad \omega \in \mathbb{Z} \cup \{-\infty\}\,, \quad \text{and}$$

$$t < t' \Rightarrow I(t) \subseteq I(t')\,, \ \text{for all } \ t > \omega\,.$$

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) $$X = \{X(t) \,:\, t \in \mathbb{Z}, t > \omega\}\,, \quad X(t) = (x_1(t), \ldots, x_{m_1}(t))'\,,\ x_i(t) \in L^2,$$
$$\text{for} \quad i = 1, \ldots, m_1 \quad \text{and} \quad t > \omega\,,\ t \in \mathbb{Z}\,,$$

and we suppose that the information sequence I is conformable with X according to the following definition. (Whenever each component of a vector process belongs to L^2, we will say as a shortcut that the process is in L^2.)

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 for every integer $t > \omega$,*

$$X(\omega, t] \subseteq I(t)$$

where $X(\omega, t]$ is the Hilbert space spanned by the components $x_i(\tau), i = 1, \ldots, m_1$, of $X(\tau)$, $\omega < \tau \leq t$.

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 be equal to $-\infty$ or 0 depending on whether we consider a stationary process on the integers ($t \in \mathbb{Z}$) or a process $\{X(t) : t \geq 1\}$ on the positive integers given initial values preceding date 1. The set $I(\omega) = \cap_{t>\omega} I(t)$ 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 non-negative 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 $P[x_i(t+h)|I(t)], 1 \leq i \leq m_1$, 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 non-zero 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 $Y(\tau)$, $\omega < \tau \leq t$, of Y. Here we suppose that

$$Y(t) = (y_1,(t), \ldots, y_{m_2}(t))', y_j(t) \in L^2, \text{ for } 1 \leq j \leq m_3 \text{ , } t \in \mathbb{Z} \text{ ,} \tag{2.3}$$

and we denote $I(t) + Y(\infty, t]$ the Hilbert subspace of L^2 generated by $I(t)$ and the components $y_j(\tau), j = 1, \ldots, m_2, \omega < \tau \leq t$. (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 non-causality at various horizons, with respect to a "universe" $I(t)$.

Definition 2.2 *(Non-causality 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 $Y \not\to_h X|I$) if

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

(ii) Y does not cause X up to horizon h given I (denoted $Y \not\to_{(h)} X|I$) if $Y \not\to_k X|I$ for $k = 1, 2, \ldots, h$;

(iii) Y does not cause X at any horizon given I (denoted $Y \not\to_\infty X|I$) if $Y \not\to_k X|I$ 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 non-causality proposed by Wiener (1956) and Granger (1969), in terms of predictibility, where the role of the forecast horizon h is emphasized. As noted by Lütkepohl (1993) and Renault and Szafarz (1991), non-causality at a single horizon (usually $h = 1$, the interval between available observations) 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 "non-causality" (at all horizons) considered by Lütkepohl (1993), the latter differs in an important way from the one(s) studied here, because Lütkepohl's definition is based on the absence of effect of the innovations of a variable on another variable (nullity of appropriate coefficients of the moving representation of the process) rather than the absence of the past value of a variable in the optimal forecasts of another variable ; for further discussion of this point, see Dufour and Tessier (1993). We think that that Definition 2.2 provides a more natural extension of the usual definition of Granger causality. Further, it allows one to study causality at any finite horizon h. Thirdly, Definition 2.2 allows for non-stationary processes, in which case it is important to study prediction at horizon h for each date t : since the equation in (i) above might hold for a given date t, but not for all dates, it is important to state that it holds for all $t > \omega$. 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 $I(\omega) = \cap_{t>\omega} I(t)$ 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 $I(\omega)$] not only about the process of interest X but possibly also concerning auxiliary variables $Z(t)$. It is clear that the choice of the initial date ω can affect causality properties.

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 any further detail on the variables contained by the information set I, it is already possible to derive some basic 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 non-causality between vectors) : Let the assumptions (2.1) to (2.3) hold, and suppose that I is conformable with X. Then the following properties are equivalent for any $h \in \mathbb{N}$:*

(i) $Y \not\to_h X|I$;

(ii) $Y \not\to_h x_i|I$, *for* $i = 1, \ldots, m_1$;

(iii) $y_j \not\to_h X|I$, *for* $j = 1, \ldots, m_2$;

(iv) $y_j \not\to_h x_i|I$, *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 the corresponding components of Y and X. In many cases, this can lead to important simplifications because real variables are simpler to study than vectors. Note however that the characterizations (iii) and (iv) above depend heavily on the fact that we consider here **linear** non-causality (based on projections rather than conditional expectations), since the orthogonality arguments used to prove these do not generally carry to concepts of conditional independence. Note also that 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 (in the more limited context causality at horizon 1), see Florens and Mouchart (1985, Property 3.5, p.164). Concerning the possibility of 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) $Y \not\to_h X|I \Rightarrow y_j \not\to_h X|I_{(j)}$, *for* $j = 1, \ldots, m_2$;

(ii) the converse implication is not true in general.

Part (i) of the latter proposition means that, whenever $Y \not\to_h X|I$, 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 certain conditions of linear independence (as we shall see in the next section).

We now give a general proposition which shows clearly that the issue of the causality horizon matters only in situations where the universe I involves other processes than X and Y.

Proposition 2.3 *(Exhaustivity condition for non-causality at all horizons) : Under the assumptions (2.1) to (2.3), suppose that $I(t) = H + X(\omega, t]$, for $t > \omega$, where H is a (possibly empty) Hilbert subspace of L^2. Then the following properties are equivalent :*

(i) $Y \not\to_1 X|I$;

(ii) $Y \not\to_{(h)} X|I$, $\forall h \in \mathbb{N}$;

(iii) $Y \not\to_\infty X|I$.

Proposition 2.3 gives a case where the usual notion of causality in the sense of Granger ($Y \not\to_1 X|I$) 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 that H may contain any variable in L^2 which does not depend on t (e.g., any variable known at every date $t > \omega$).

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 :

$$Z(t) = (z_1(t), \ldots, z_{m_3}(t))', \quad z_k(t) \in L^2, \tag{2.4}$$

$$\text{for} \quad k = 1, \ldots, m_3 \quad \text{and} \quad t > \omega\ , \ t \in \mathbb{Z}\ .$$

The reference information set at date t is then

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

where H is defined as in Proposition 2.3. In this case, the latter cannot be applied directly to show that $Y \not\to_\infty X|I_{XZ}$. However, an important case where the latter property holds is the one where a **separation condition** is satisfied.

Suppose indeed that can find two processes $\{Z_1(t) : t \in \mathbb{Z}, t > \omega\}$ and $\{Z_2(t) : t \in \mathbb{Z}, t > \omega\}$ of dimensions m_{31} and m_{32} respectively such that $I_{XZ}(t)$ can be obtained by adding the spaces

$$I_{XZ_1}(t) = H + X(\omega, t] + Z_1(\omega, t] \tag{2.6}$$

and $Z_2(\omega, t]$:

$$I_{XZ}(t) = I_{XZ_1}(t) + Z_2(\omega, t], \quad \forall t > \omega\ . \tag{2.7}$$

This will hold, in particular, if the vector $(Z_1(t)', Z_2(t)')'$ is an invertible linear transformation of $Z(t)$, for example 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 can show the following proposition.

Proposition 2.4 *(Separation condition for non-causality at all horizons) : Let H be a Hilbert subspace of L^2, and suppose the assumptions (2.1) to (2.7) hold. Then the* **separation condition**

$$\begin{bmatrix} Y \\ Z_2 \end{bmatrix} \not\to_1 \begin{bmatrix} X \\ Z_1 \end{bmatrix} | I_{XZ_1} \tag{2.8}$$

is a sufficient condition for $Y \not\to_\infty X|I_{XZ}$.

Intuitively, Proposition 2.4 means that whenever the separation condition holds, not only do we have $Y \not\to_1 X|I_{XZ}$ but also $Y \not\to_\infty X|I_{XZ}$. 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 non-stationary) and invertible ARIMA processes. This will allow us to obtain explicit parametric formulations of the non-causality 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 $\{W(t) : t \in \mathbb{Z}\}$ in L^2 with an autoregressive representation (possibly of infinite order) :

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

where $\mu(t)$ belongs to some Hilbert subspace H of L^2 ($\mu(t) \in E$, $\forall\, t > \omega$), $\{a(t) : t \in \mathbb{Z}\}$ 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 that the vectors $\mu(t)$ may be fixed (non-random). It is not required that the covariance matrix of

$a(t)$ be constant or non-singular. Thus $a(t)$ may not be a white noise process. Further, the vector $W(t)$ is partitioned into three subvectors,

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

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 (if any) :

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

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

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

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 non-stationary 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

$$\sum_{j=1}^{\infty} \|\pi_j\| < \infty \tag{3.5}$$

where $\|\pi_j\|^2 = tr(\pi_j \pi_j')$. Note also that model (3.1) includes as special case the model

$$X(t) - \bar{\mu}(t) = \sum_{j=1}^{\infty} \pi_j [X(t-j) - \bar{\mu}(t-j)] + a(t), \quad t > \omega, \tag{3.6}$$

where each component of $\bar{\mu}(t)$ belongs to L^2 and the series $\sum_{j=1}^{\infty} \pi_j \bar{\mu}(t-j)$ converges in q.m. In (3.6), 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. For that purpose, the following proposition will be especially useful.

Proposition 3.1 *(Recursive formulae for projection coefficients) : If the process $\{W(t) : t \in \mathbb{Z}\}$ satisfies (3.1), then*

$$P[W(t+h) | H + W(-\infty, t]] = \sum_{k=0}^{h-1} \pi_1^{(k)} \mu(t+h-k) \tag{3.7}$$

$$+ \sum_{j=1}^{\infty} \pi_j^{(h)} W(t+1-j), \quad \forall t > \omega, \quad \forall h \in \mathbb{N},$$

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

$$\pi_j^{(1)} = \pi_j, \quad \pi_j^{(h+1)} = \pi_{j+h} + \sum_{l=1}^{h} \pi_{h-l+1} \pi_j^{(l)}, \quad h = 1, 2, \ldots \tag{3.8}$$

Furthermore, any sequence of matrices $\pi_j^{(h)}$, where $j \in \mathbb{N}$ and $h \in \mathbb{N}$, which satisfies (3.8) also satisfies the recursion :

$$\pi_j^{(1)} = \pi_j, \quad \pi_j^{(h+1)} = \pi_{j+1}^{(h)} + \pi_1^{(h)} \pi_j, \quad h = 1, 2, \ldots \tag{3.9}$$

Proposition 3.1 allows one to characterize the non-causality $Y \not\rightarrow_h X | I_{XZ}$ from natural partitions of the matrices $\pi_j^{(h)}$ conformable with X, Y and Z :

$$\pi_j^{(h)} = \begin{bmatrix} \pi_{XXj}^{(h)} & \pi_{XYj}^{(h)} & \pi_{XZj}^{(h)} \\ \pi_{YXj}^{(h)} & \pi_{YYj}^{(h)} & \pi_{YZj}^{(h)} \\ \pi_{ZXj}^{(h)} & \pi_{ZYj}^{(h)} & \pi_{ZZj}^{(h)} \end{bmatrix}. \tag{3.10}$$

We can then show the following result, which is a generalization of Proposition 1 in Boudjellaba, Dufour and Roy (1992).

Theorem 3.1 *(Projection coefficient characterization of non-causality at horizon h) : Under the assumptions (3.1) to (3.3), the condition*

$$\pi_{XYj}^{(h)} = 0, \quad \forall j \in \mathbb{N}, \tag{3.11}$$

is sufficient for $Y \not\rightarrow_h X | I_{XZ}$, where $h \in \mathbb{N}$. If, furthermore, the process $W(t)$ is regular [assumption (3.4)], then

$$Y \not\rightarrow_h X | I_{XZ} \Leftrightarrow \pi_{XYj}^{(h)} = 0, \quad \forall j \in \mathbb{N}.$$

Proposition 3.1 and Theorem 3.1 allow 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 (3.9), we have :

$$\pi_{XYj}^{(h+1)} = \pi_{XY,j+1}^{(h)} + \pi_{XX1}^{(h)} \pi_{XYj} + \pi_{XY1}^{(h)} \pi_{YYj} + \pi_{XZ1}^{(h)} \pi_{ZYj} \tag{3.12}$$

which upon using Theorem 3.1 implies the following result.

Corollary 3.1 *Under the assumptions (3.1) to (3.4),*

$$Y \underset{(h)}{\not\rightarrow} X | I_{XZ} \Rightarrow \pi_{XYj}^{(h+1)} = \pi_{XZ1}^{(h)} \pi_{ZYj}, \quad \forall j \in \mathbb{N}.$$

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 ($\pi^{(h)}_{XZ1} \neq 0$) and Y can cause Z at horizon 1 ($\pi_{ZYj} \neq 0$). Thus the presence of Z can introduce indirect causality from Y to X going through Z. This leads to sufficient conditions for non-causality 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

$$Y \underset{(h)}{\not\to} X \Rightarrow \pi^{(h+1)}_{XYj} = \pi^{(h)}_{XZ_11}\pi_{Z_1Yj} + \pi^{(h)}_{XZ_21}\pi_{Z_2Yj} ,$$

hence sufficient conditions of the type :

$$Y \underset{1}{\not\to} \begin{bmatrix} X \\ Z_1 \end{bmatrix} | I_{XZ} \quad \text{and} \quad Z_2 \underset{(h)}{\not\to} X | I_{XYZ_1} \Rightarrow Y \underset{(h+1)}{\not\to} X | I_{XZ}$$

which can be verified easily. Theorem 3.2 below gives a complete characterization of non-causality from Y to X is terms of "causality chains". To prove it, we will need two lemmas (of separate interest) on the properties of matrix sequences $\pi^{(h)}_j$ which satisfy the recursions (3.9).

Lemma 3.1 *Let $\pi^{(h)}_j, j \in \mathbb{N}, h \in \mathbb{N}$, be any sequence of $m \times m$ matrices, which are partitioned as in (3.10) and satisfy the recursion (3.9). If*

$$\pi^{(k)}_{XYj} = 0, \quad \forall j \in \mathbb{N}, \quad k = 1, \ldots, h ,$$

then for any integer p such that $2 \leq p \leq h$, we have

(3.13)
$$\pi^{(h)}_{XZ1}\pi_{ZYj} = \sum_{l=1}^{p} \pi^{(h-p+1)}_{XZl} \left\{ \sum_{J(p-l)} \left[\prod_{i=1}^{p-l} \pi^{n_i}_{ZZi} \right] \right\} \pi_{ZYj} , \quad \forall j \in \mathbb{N} ,$$

where

$$J(l) = \left\{ (n_1, n_2, \ldots, n_l) \; : \sum_{i=1}^{l} i n_i = l \quad and \quad n_i \in \mathbb{N}_0, i = 1, \ldots, l \right\} ,$$

$$\prod_{i=1}^{p-l} \pi^{n_i}_{ZZi} = \pi^{n_1}_{ZZ1}\pi^{n_2}_{ZZ2} \ldots \pi^{n_{p-l}}_{ZZ,k-l} , \mathbb{N}_0 = \{0, 1, 2, \ldots\}$$

with the convention

$$\sum_{J(0)} \left[\prod \pi^{n_i}_{ZZi} \right] = I_{m_3} .$$

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

(3.14) $\quad \pi^{(k)}_{XYj} = 0 , \quad \forall j \in \mathbb{N}, \quad k = 1, \ldots, h ;$

(3.15) $\quad \pi_{XYj} = 0 , \; \forall j \in \mathbb{N},$ *and* $\pi^{(k)}_{XZ1}\pi_{ZYj} = 0, k = 1, \ldots, h-1, \forall j \in \mathbb{N} ;$

(3.16) $\quad \pi_{XYj} = 0 , \; \forall j \in \mathbb{N},$ *and* $R^{(k)}_{XZ}\pi_{ZYj} = 0, k = 1, \ldots, h-1, \forall j \in \mathbb{N} ,$

where

(3.17)
$$R^{(k)}_{XZ} = \sum_{l=1}^{k} \pi_{XZl} \left\{ \sum_{J(k-l)} \left[\prod_{i=1}^{k-l} \pi_{ZZi} \right] \right\} .$$

Theorem 3.2 *(Causality chain characterization of non-causality at horizon h) : Under the assumptions (3.1) to (3.3), each one of the three equivalent conditions (3.14), (3.15) and (3.16) is sufficient for $Y \underset{(h)}{\not\to} X | I_{XZ}$. If furthermore the process $W(t)$ is regular [assumption (3.4)], then each one of the conditions (3.14), (3.15) and (3.16) is necessary and sufficient for $Y \underset{(h)}{\not\to} X | I_{XZ}$.*

The interest of the criteria (3.15) and (3.16) of non-causality $Y \underset{(h)}{\not\to} X | I_{XZ}$, as opposed to (3.14) [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 (3.15) shows that non-causality at horizon h occurs when two conditions hold : (1) there is non-causality at horizon 1 ($\pi_{XYj} = 0, \;\; \forall j$), and (2) the composed effects $\pi^{(k)}_{XZ1}\pi_{ZYj}$ that runs first from Y to Z at horizon 1 (π_{ZYj}) and then from Z to X at horizons less than h ($\pi^{(k)}_{XZ1}, k = 1, \ldots, h-1$) are zero. Criterion (3.15) gives an explicit expression for these composed effects by relating them to all possible "causality chains" that runs from Y to Z, from components of Z to other components of Z, and then from Z to X [see the expression (3.16)].

Criterion (3.15) also provides a link between our concept of "non-causality at all horizons" and the coefficients of the moving average (MA) representation of the process ["impulse response coefficients", see Sims (1980)]. To do this, let us consider the formal series :

(3.18)
$$\pi(z) = I_m - \sum_{j=1}^{\infty} \pi_j z^j, \quad \psi(z) = \pi(z)^{-1} = I_m + \sum_{j=1}^{\infty} \psi_j z^j .$$

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 here to take note of another algebraic property of the matrices $\pi^{(h)}_j$.

Lemma 3.3 *Let $\pi^{(h)}_j$, $j \in \mathbb{N}, h \in \mathbb{N}$ be any sequence of $m \times m$ matrices which satisfy the recursion (3.8). Then*

(3.19)
$$\pi^{(h)}_1 = \psi_h , \quad \forall h \geq 0 ,$$

where the matrice ψ_h are defined by (3.18) with $\psi_0 = I_m$.

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 (3.7). Let us now partition each matrix ψ_h conformably with X, Y and Z :

(3.20)
$$\psi_h = \begin{bmatrix} \psi_{XXh} & \psi_{XYh} & \psi_{XZh} \\ \psi_{YXh} & \psi_{YYh} & \psi_{YZh} \\ \psi_{ZXh} & \psi_{ZYh} & \psi_{ZZh} \end{bmatrix} , \quad h \geq 0 ,$$

and similarly for $\psi(z)$ in (3.18). By combining (3.19) with Theorem 3.1, it is then easy to see that non-causality at horizon h implies zero restrictions on the impulse response coefficients of a regular process.

Corollary 3.2 *(Necessary conditions for non-causality at horizon h). Under the assumptions (3.1) to (3.4), the condition*

(3.21) $$\pi_{XYj} = 0\ ,\ \forall j \in \mathbb{N}\ ,\ \text{and } \psi_{XYh} = 0$$

is necessary for $Y \not\to_h X | I_{XZ}$ *(where* $1 \le h \le \infty$ *), and the condition*

(3.22) $$\pi_{XYj} = 0\ ,\ \forall j \in \mathbb{N}\ ,\ \text{and } \psi_{XYj} = 0\ ,\ j = 1, \ldots, h\ ,$$

is necessary for $Y \not\to_{(h)} X | I_{XZ}$ *(where* $1 \le h \le \infty$ *).*

It is important to note that condition (3.21) is **only necessary** for $Y \not\to_h X | I_{XZ}$: the necessary and sufficient condition given by Theorem 3.1 requires $\pi^{(h)}_{XYj} = 0$ for all $j \ge 1$. If we now combine (3.19) with Theorem 3.2, we can get a sufficient (or a necessary and sufficient) condition for $Y \not\to_{(h)} X | I_{XZ}$ which involves both impulse responses and autoregressive parameters.

Corollary 3.3 *(Impulse response characterization of non-causality up to horizon h). Under the assumptions (3.1) to (3.3), the condition*

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

is sufficient for $Y \not\to_{(h)} X | I_{XZ}$. *If furthermore the process* $W(t)$ *is regular [assumption (3.4)], condition (3.21) is necessary and sufficient for* $Y \not\to_{(h)} X | I_{XZ}$.

In order to gain more insight on the connection between causality at horizon 1, impulse responses and causality at longer horizons, let us partition $\pi(z)$ in (3.18), $\mu(t) = (\mu_X(t)', \mu_Y(t)', \mu_Z(t)')'$ and $a(t) = (a_X(t)', a_Y(t)', a_Z(t)')'$ conformably with $X(t), Y(t)$ and $Z(t)$. We can then write the autoregressive representation (3.1) as :

(3.24a) $$\pi_{XX}(L)X(t) + \pi_{XY}(L)Y(t) + \pi_{XZ}(L)Z(t) = \mu_X(t) + a_X(t)\ ,$$

(3.24b) $$\pi_{YX}(L)X(t) + \pi_{YY}(L)Y(t) + \pi_{YZ}(L)Z(t) = \mu_Y(t) + a_Y(t)\ ,$$

(3.24c) $$\pi_{ZX}(L)X(t) + \pi_{ZY}(L)Y(t) + \pi_{ZZ}(L)Z(t) = \mu_Z(t) + a_Z(t)\ ,$$

To simplify the exposition, let us suppose temporarily that $\mu(t) = 0, \forall t$. The usual criterion of non-causality from Y to X at horizon 1 ($Y \not\to_1 X | I_{XZ}$) is $\pi_{XY}(L) = 0$ which ensures that the past of Y does not appear in (3.24a). But from (3.24c), we can write formally :

(3.25) $$Z(t) = \pi_{ZZ}(L)^{-1}[a_Z(t) - \pi_{ZX}(L)X(t) - \pi_{ZY}(L)Y(t)]$$

where the convergence (in q.m.) of the series is ensured when the matrix $\pi_{ZZ}(z)$ is non-singular for $|z| \le 1$. Substituting (3.25) into (3.24a), we then get :

$$\left[\pi_{XX}(L) - \pi_{XZ}(L)\pi_{ZZ}(L)^{-1}\pi_{ZX}(L)\right] X(t) + \left[\pi_{XY}(L) - \pi_{XZ}(L)\pi_{ZZ}(L)^{-1}\pi_{ZY}(L)\right] Y(t)$$
$$= a_X(t) - \pi_{XZ}(L)\pi_{ZZ}(L)^{-1}a_Z(t)\ .$$

We see that, even if $\pi_{XY}(z) = 0$, the past of Y can play an indirect role in the determination of X when the formal series $\pi_{XZ}(z)\pi_{ZZ}(z)^{-1}\pi_{ZY}(z)$ is non-zero. This provides the intuition of the result in Corollary 3.4 below, which can be derived in a simple way from Theorem 3.2 and the following lemma relating the matrices $R^{(k)}_{XZ}$ to the formal series $\pi_{XZ}(z)$ and $\pi_{ZZ}(z)$.

Lemma 3.4 *If we denote by*

$$R_{XZ}(z) = \sum_{k=1}^{\infty} R^{(k)}_{XZ} z^k$$

the formal series associated with the coefficients $R^{(k)}_{XZ}$ *defined in Lemma 3.2, then the following relationship holds :*

$$R_{XZ}(z) \equiv -\pi_{XZ}(z)\pi_{ZZ}(z)^{-1},$$

where the symbol $\equiv$ *means that the two formal series (in z) considered are identical (i.e., the equality of the coefficients of the corresponding powers of z on both sides of* $\equiv$*).*

Corollary 3.4 *(Necessary conditions for non-causality at all horizons) : Let the assumptions (3.1) to (3.4) hold, and suppose that the power series* $\pi(z)$ *converges when* $z \in C$ *and* $|z| < \rho$*, for some* $\rho > 0$*. Then the three following conditions are equivalent and each one of them yields a necessary condition for* $Y \not\to_\infty X | I_{XZ}$ *:*

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

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

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

where $\psi(z)$ *is defined by (3.18) and (3.20).*

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 $\|\pi_j\|, j \ge 1$, is bounded and even if $\|\pi_j\|$ grows at an exponential rate ($\|\pi_j\| \le C\rho^j$) as j increases. Note that condition (3.28) is formally identical with (3.26) on permuting π and ψ, which is possible because π and ψ play symmetric roles in (3.27). These (necessary) conditions eliminate both the **direct** effect of Y on X (by cancelling 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 Theorem 3.2 that they are not **when** Z **is multivariate**. Indeed, the condition (3.26) can be written :

(3.29) $$\pi_{XYj} = 0, \quad \forall j \ge 1\ , \quad \text{and} \quad \sum_{j=1}^{k-1} R^{(k-j)}_{ZYj}\pi_{ZYj} = 0\ , \quad \forall k > 1\ .$$

On the other hand, the necessary and sufficient condition (3.16) from Theorem 3.2 is :

(3.30) $$\pi_{XYj} = 0\ , \quad \forall j \ge 1\ , \quad \text{and} \quad R^{(k)}_{XZ}\pi_{ZYj} = 0\ , \quad \forall k \ge 1\ , \quad \forall j \ge 1\ .$$

which is clearly not generally equivalent to (3.29). Comparing (3.29) and (3.30) shows in particular that the condition $\pi_{XY}(z) \equiv \psi_{XY}(z) \equiv 0$ is generally insufficient for $Y \not\to_\infty X | I_{XZ}$.

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 non-causality at all horizons for Z univariate) : Under the assumptions of Corollary 3.4, suppose that the process $Z(t)$ is univariate ($m_3 = 1$). Then the four following properties are equivalent :*

(3.31) $Y \underset{\infty}{\not\to} X | I_{XZ}$;

(3.32) $\pi_{ZY}(z) \equiv 0$ *and* $\pi_{XZ}(z)\pi_{ZZ}(z)^{-1}\pi_{ZY}(z) \equiv 0$;

(3.33) $\pi_{XY}(z) \equiv 0$ *and* $\psi_{XY}(z) \equiv 0$;

(3.34) $\psi_{XY}(z) \equiv 0$ *and* $\psi_{XZ}(z)\psi_{ZZ}(z)^{-1}\psi_{ZY}(z) \equiv 0$.

The case where Z is multivariate is more complicated because causality relations internal to Z must be taken into account. This problem will be studied in greater detail in Section 5.[1]

4 Non-causality 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 parametrized 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 4.1 *(Truncation rule for non-causality 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*

$$\pi^{(h)}_{XY j} = 0 \ , \forall j \in \mathbb{N} \ , \ h = 1, \ldots, m_3 p + 1 \Rightarrow Y \underset{\infty}{\not\to} X | I_{XZ} \ .$$

If, furthermore, the process $W(t)$ is regular [(assumption (3.4)], then

$$Y \underset{\infty}{\not\to} X | I_{YZ} \Leftrightarrow \quad Y \underset{(m_3 p+1)}{\not\to} X | I_{XZ}$$

In other words, for autoregressive processes of order p, it is sufficient to have non-causality up to horizon $m_3 p + 1$ for non-causality at all horizons to occur. 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 non-causality up to horizon q is sufficient to have non-causality at all horizons.

A truncation result similar to the one of Proposition 4.1 also holds for the necessary condition of Corollary 3.4.

[1] For the special case of finite order VAR models, with $X(t), Y(t)$ and $Z(t)$ all univariate ($m_1 = m_2 = m_3 = 1$), Bruneau and Nicolaï (1992) have independently derived the conditions (3.27) and (3.33). It is clear that Corollaries 3.4 and 3.5 provide considerably more general results, as they apply to infinite order autoregressive models with $X(t), Y(t)$ [as well as $Z(t)$ in Corollary 3.4] possibly multivariate.

Proposition 4.2 *(Necessary condition for non-causality at all horizons in VAR processes) : Under the assumptions (3.1) to (3.4) with $\pi_k = 0$ for $k > p$, the necessary condition*

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

for $Y \not\to_{\infty} X | I_{XZ}$ is equivalent to the following finite set of conditions :

(4.1a) $\pi_{XYj} = 0, j = 1, \ldots, p$;

(4.1b) *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 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 4.1 entails that testing $Y \not\to_{\infty} X | I_{XZ}$ requires one to test $pm_1m_2 + (m_3p)(pm_1m_2) = pm_1m_2(m_3p+1)$ restrictions, because testing causality up to horizon m_3p+1 leads one to check that the m_3pp matrices $R^{(k)}_{XZ}\pi_{ZYj}, k = 1, \ldots, m_3p, j = 1, \ldots, p$ are all zero.

As expected, we see that the necessary and sufficient conditions for non-causality 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 $\pi^{(h)}_1 = \pi^h_1 = \psi_h$ 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 non-causality at all horizons. Further, it is important to note that $pm_1m_2(m_3p+1)$ is an **upper bound** on the effective number of restrictions which characterize non-causality at all horizons. For example, when $m_3 = 1$, we get from Corollary 3.5 and Proposition 4.2 that $Y \not\to_{\infty} X | I_{XZ}$ is ensured by $3pm_1m_2$ restrictions, instead of $(p+1)pm_1m_2$ restrictions.

5 Separating causal decompositions and rank conditions for non-causality

Except for causality at horizon 1, the conditions for non-causality given in the two previous sections are generally nonlinear. Furthermore the transformations of the autoregressive coefficients considered are multilinear [i.e., they are sums of products of the autoregressive coefficients, as in (3.15) and (3.16)], and so the rank of the corresponding jacobian matrix may easily not be constant ; for other examples of such problems, see Boudjellaba, Dufour and Roy (1992, 1994). This of course can complicate the interpretation and the execution of the corresponding tests. It would certainly be useful here to find necessary conditions for non-causality from Y to X (up to horizon h or at all horizons) which are easier to test in practice.

When a test for such a necessary condition is significant, it allows one to unambiguously reject the null hypothesis of non-causality (up to a given horizon) considered. In particular, it would be both theoretically illuminating and useful from a practical standpoint to reduce properties of non-causality up to horizon h ($h \geq 2$) to non-causalities at horizon one.

In Proposition 2.4, we saw that the concept of **separation** provides a **sufficient** condition for non-causality at all horizons which is based on non-causalities at horizon 1. We will now use generalizations of the concept of separation to obtain necessary conditions of non-causality. To do this we will consider appropriate linear transformations of $Z(t)$, say $Z_1(t) = Q_1 Z(t)$ and $Z_2(t) = Q_2 Z(t)$ which automatically satisfy a weaker separation condition. We shall call such transforms "separating causal factors".

To define these, we consider the subspace

(5.1) $$E_{ZY} = span\{\pi_{Z_{Yk}j} : 1 \leq k \leq m_2, j \geq 1\}$$

of $\mathbb{R}^{m_3}$ (with $m_3 \geq 1$) generated by the columns $\pi_{Z_{Yk}j}$, $k = 1, 2, \ldots, m_2$, of all the $m_3 \times m_2$ submatrices

$$\pi_{ZYj} = \left[\pi_{Z_{Y1}j}, \ldots, \pi_{Z_{Ym_2}j}\right], \quad j \geq 1,$$

of π_j in model (3.1) [see (3.10)], as well as its orthogonal complement $E_{ZY}^{\perp}$ in $\mathbb{R}^{m_3}$. So, by definition, $a \in E_{ZY}^{\perp}$ if and only if $a'\pi_{ZYj} = 0$ for all $j \geq 1$. Let

(5.2) $$n_{31} = dim(E_{ZY}^{\perp}) \; , \; n_{32} = dim(E_{ZY})$$

be the dimensions of these subspaces ($0 \leq n_{31} \leq m_3$, $n_{32} = m_3 - n_{31}$), $m_{31} = max\{1, n_{31}\}$, $m_{32} = \max\{1, n_{32}\}$, and define Q_1 and Q_2 to be $m_{31} \times m_3$ and $m_{32} \times m_3$ matrices such that the columns of Q_1' and Q_2' constitute bases of $E_{ZY}^{\perp}$ and E_{ZY} : i.e.

(5.3) $$Im(Q_1') = E_{ZY}^{\perp}, \qquad Im(Q_2') = E_{ZY} \; ;$$

when $n_{31} = 0$ (i.e., $E_{ZY}^{\perp} = \{0\}$), we set $Q_1' = 0$, while for $n_{32} = 0$, we set $Q_2' = 0$. We consider also the $m_3 \times m_3$ matrix Q defined by

(5.4) $$\begin{aligned} Q &= [Q_1', Q_2']', \quad \text{if} \quad n_{31} \geq 1 \quad \text{and} \quad n_{32} \geq 1 \, , \\ &= Q_1 \, , \quad \text{if} \quad n_{31} = m_3 \, , \\ &= Q_2 \, , \quad \text{if} \quad n_{31} = 0 \, . \end{aligned}$$

By construction, the matrix Q is invertible. The appropriate "separating causal factors" are then defined as follows.

Definition 5.1 *(Separating causal factors) : Under the assumptions (3.1) and (3.2) with $m_3 \geq 1$, we say that the vectors*

(5.5) $$Z_1(t) = Q_1 Z(t) \; , Z_2(t) = Q_2 Z(t) \; ,$$

where Q_1 and Q_2 satisfy (5.3), are **separating vectors of causal factors** *associated with Z for the relationship $Y \to X$. If furthermore the matrix Q defined by (5.4) is orthogonal, i.e.*

(5.6) $$Q'Q = QQ' = I_{m_3} \; ,$$

we say that the separating vectors $Z_1(t)$ and $Z_2(t)$ have an orthogonal base.

It is clear from the latter definition that the vectors $Z_1(t)$ and $Z_2(t)$ are not uniquely defined. In particular, we can always choose Q_1 and Q_2 so that the matrix Q is orthogonal. Note that Definition 5.1 also implies that

$$Z(t) = Q_1^{-1} Z_1(t) \qquad \text{and} \qquad Z_2(t) = Q_2 = 0$$

when E_{ZY} contains only the zero vector (so that $n_{32} = 0$ and $n_{31} = m_{31} = m_3$) while we have $Z(t) = Q_2^{-1} Z_2(t)$ and $Z_1(t) = Q_1 = 0$ when $E_{ZY} = \mathbb{R}^{m_3}$ (so that $n_{31} = 0$ and $n_{32} = m_{32} = m_3$). If furthermore the matrix Q is orthogonal, we have $Z(t) = Z_1(t)$ in the first case, and $Z(t) = Z_2(t)$ in the second case. In these two extreme cases, there is only one non-degenerate separating vector of causal factors, which is identical to $Z(t)$ up to a non-singular linear transformation.

By introducing separating causal factors, we are in fact defining a new process

$$\tilde{W}(t) = [X(t)', Y(t)', \tilde{Z}(t)']' \; , \quad \text{where} \quad \tilde{Z}(t) = QZ(t) \; ,$$

which is obtained by a non-singular linear transformation of $W(t)$:

(5.7) $$\tilde{W}(t) = \tilde{Q}W(t) \; , \qquad \tilde{Q} = \begin{bmatrix} I_{m_1+m_2} & O' \\ O & Q \end{bmatrix} ,$$

so that $I_{X\tilde{Z}}(t) = I_{XZ}(t)$ and $I_{\tilde{W}}(t) = I_W(t)$. It is then clear that $\tilde{W}(t)$ admits an autoregressive representation similar to (3.1) with π_j replaced by

(5.8) $$\tilde{\pi}_j = \tilde{Q}\pi_j\tilde{Q}^{-1} = \begin{bmatrix} I_{m_1+m_2} & O' \\ O & Q \end{bmatrix} \pi_j \begin{bmatrix} I_{m_1+m_2} & O' \\ O & Q^{-1} \end{bmatrix} , \; j \geq 1 \, .$$

In particular, the coefficient matrix of $Y(t-j)$ in the autoregressive representation of $Z_i(t)$ is $\tilde{\pi}_{Z_iYj} = Q_i\pi_{ZYj}$ $(i = 1, 2)$. Using the fact that the columns of Q_1' and Q_2' constitue bases of the subspaces $E_{ZY}^{\perp}$ and E_{ZY} in $\mathbb{R}^{m_3}$, it is then easy to prove the following proposition which motivates Definition 5.1 and the terminology "separating causal factors".

Proposition 5.1 *(Separating causal decomposition) : Let the assumptions (3.1) to (3.3) hold with $m_3 \geq 1$, let $Z_1(t)$ and $Z_2(t)$ be defined as in (5.5). Then we can write :*

(5.9) $$Z(t) = P_1 Z_1(t) + P_2 Z_2(t)$$

where P_1 and P_2 are fixed matrices and

(5.10) $$Y \underset{1}{\not\to} Z_1 | I_{XZ} \, .$$

If furthermore the process $W(t)$ is regular [assumption (3.4)] and $dim(E_{ZY}) \geq 1$, we also have :

(5.11) $$Y \underset{1}{\to} Z_2 | I_{XZ} \, .$$

From the latter proposition, we see that the auxiliary variable vector $Z(t)$ of a regular process can always be represented as a linear combination of two processes $Z_1(t)$ and $Z_2(t)$, such that $Y \not\to_1 Z_1|I_{XZ}$ and $Y \to_1 Z_2|I_{XZ}$. Thus Y can cause Z at the horizon 1 only through Z_2.

To obtain a characterization of $Y \not\to_{(h)} X|I_{XZ}$ based on a generalized separation condition, consider the coefficient matrices $\tilde{\pi}_j^{(h)}$ and $\tilde{\psi}_j$ obtained from the autoregressive representation of $\tilde{W}(t)$ by equations analogous to (3.7)-(3.9) and (3.18). It is then easy to see that $\tilde{\pi}_j^{(h)}$ and $\tilde{\psi}_j$ are related to $\pi_j^{(h)}$ and ψ_j by the equations :

$$(5.12) \qquad \tilde{\pi}_j^{(h)} = \tilde{Q}\pi_j^{(h)}\tilde{Q}^{-1} = \begin{bmatrix} \pi_{XXj}^{(h)} & \pi_{XYj}^{(h)} & \tilde{\pi}_{XZj}^{(h)} \\ \pi_{YXj}^{(h)} & \pi_{YYj}^{(h)} & \tilde{\pi}_{YZj}^{(h)} \\ \tilde{\pi}_{ZXj}^{(h)} & \tilde{\pi}_{ZYj}^{(h)} & \tilde{\pi}_{ZZj}^{(h)} \end{bmatrix},$$

$$\tilde{\psi}_j = \tilde{Q}\psi_j\tilde{Q} = \begin{bmatrix} \psi_{XXj} & \psi_{XYj} & \tilde{\psi}_{XZj} \\ \psi_{YXj} & \psi_{YYj} & \tilde{\psi}_{YZj} \\ \tilde{\psi}_{ZXj} & \tilde{\psi}_{ZYj} & \tilde{\psi}_{ZZj} \end{bmatrix},$$

for $j \geq 1$ and $h \geq 1$, where the matrices $\tilde{\pi}_j^{(h)}$ and $\tilde{\psi}_j$ have been partitioned conformably with $(X(t)', Y(t)', \tilde{Z}(t)')'$. To simplify the exposition, we will now suppose that Q is orthogonal (so that $\tilde{Q}$ is also orthogonal) and define the coefficient matrices associated with Z_1 and Z_2 as follows :

$$(5.13) \qquad \tilde{\pi}_{Z_iXj}^{(h)} = Q_i\pi_{ZXj}^{(h)}, \qquad \tilde{\pi}_{Z_iYj}^{(h)} = Q_i\pi_{ZYj}^{(h)},$$

$$\tilde{\pi}_{XZ_ij}^{(h)} = \pi_{XZ_ij}^{(h)}Q_i', \qquad \tilde{\pi}_{YZ_ij}^{(h)} = \pi_{YZ_ij}^{(h)}Q_i',$$

$$\tilde{\pi}_{Z_1Z_2j}^{(h)} = Q_1\pi_{Z_1Z_2j}^{(h)}Q_2', \qquad \tilde{\pi}_{Z_2Z_1j}^{(h)} = Q_2\pi_{Z_2Z_1}^{(h)}Q_1',$$

for all $h \geq 1$, $j \geq 1$, and $i = 1, 2$, and similarly for $\tilde{\psi}_j$. We then have the following proposition.

Proposition 5.2 *(Characterization of non-causality by the separating causal decomposition) : Let the assumptions (3.1) to (3.4) hold with $m_3 \geq 1$, and let $Z_1(t), Z_2(t)$ and the corresponding $\tilde{\pi}_{.j}$ matrices be defined as in (5.5), (5.6), (5.12) and (5.13). Then the three following properties are equivalent for any horizon $h \geq 2$:*

$$(5.14) \qquad Y \underset{(h)}{\not\to} X|I_{XZ} ;$$

$$(5.15) \qquad Y \underset{1}{\not\to} X|I_{XZ}, \quad \textit{and} \quad \tilde{\pi}_{XZ_21}^{(k)} = 0 \quad \textit{for} \quad k = 1, 2, \ldots, h-1 ;$$

$$(5.16) \qquad Y \underset{1}{\not\to} X|I_{XZ} \quad \text{and} \quad \tilde{\psi}_{XZ_2k} = 0 \quad \text{for} \quad k = 1, 2, \ldots, h-1 .$$

Given that $Y \not\to_1 X|I_{XZ}$, we see from Proposition 5.2 that the coefficients $\tilde{\pi}_{YZ_21}^{(k)}$, $k = 1, \ldots, h-1$ are those which determine causality up to any further horizon h (for a regular process). In the extreme case where $Q_2 = 0$ (i.e. $E_{ZY} = \{0\}$), the condition

$$(5.17) \qquad \tilde{\pi}_{XZ_21}^{(k)} = 0, \qquad k = 1, \ldots, h-1,$$

holds automatically from the definition $\tilde{\pi}_{XZ_21}^{(k)} = \pi_{XZ_21}^{(k)}Q_2'$ and the property (5.15) reduces to $Y \not\to_1 X|I_{XZ}$: the latter is then sufficient for $Y \not\to_{(h)} X|I_{XZ}$ because we have $Z(t) = Z_1(t)$, so that $Y \not\to_1 Z|I_{XZ}$, and the separation condition of Proposition 2.4 applies. In the other cases, the condition (5.17) is a weakened form of non-causality from Z_2 to X, as it means that $P[X(t+k)|I_{XZ}(t) + Y(\omega, t]]$ does not depend on $Z_2(t)$ for $1 \leq k \leq h-1$. But these forecasts may depend on $Z_2(t-j), j \geq 1$.

We may ask here how it occurs that the presence of $Z_2(t-j), j \geq 1$, in the optimal forecasts of $X(t+k), 1 \leq k \leq h-1$, does not entail indirect causality from Y to X going through Z_2. The answer to this question is contained in the following proposition.

Proposition 5.3 *Under the assumptions of Proposition 5.2, the property $Y \not\to_{(h)} X|I_{XZ}$ with $h \geq 2$ implies :*

$$(5.18) \qquad \tilde{\pi}_{XZ_2j} + \sum_{\ell=1}^{j-1} \tilde{\pi}_{XZ_1,j-\ell}\tilde{\pi}_{Z_1Z_21}^{(\ell)} + \sum_{\ell=1}^{j-1} \tilde{\pi}_{XZ_2,j-\ell}\tilde{\pi}_{Z_2Z_21}^{(\ell)} = 0, \ \textit{for}\ 1 \leq j \leq h-1,$$

where the latter condition reduces to $\tilde{\pi}_{XZ_21} = 0$ when $h = 2$.

Proposition 5.3 indeed shows that, even if $Z_2(t-j), j \geq 1$, appear in the optimal forecast of $X(t+1)$ based on $I_{XZ}(t) + Y(\omega, t]$, this causality is compensated (when $Y \not\to_{(h)} X|I_{XZ}$ with $h > j$) by causality chains in opposite directions of the following forms :

$$Z_2(t+1-j) \to Z_1(t+1-j+l) \to X(t+1),$$
$$Z_2(t+1-j) \to Z_2(t+1-j+l) \to X(t+1).$$

We will now show that $Z_2 \not\to_1 X|I_{XYZ_1}$ must indeed hold to have $Y \not\to_\infty X|I_{XZ}$ (i.e. the separation property must hold), unless a rather special compensation scheme of the form $Z_2 \to Z_1 \to X$ operates.

Proposition 5.4 *(Generalized separation condition for non-causality) : Let the assumptions of Proposition 5.2 hold. Then, for $h \geq 2$,*

$$(5.19) \qquad Y \underset{(h)}{\not\to} X|I_{XZ} \quad \textit{and} \quad \sum_{l=1}^{j-1} \tilde{\pi}_{XZ_1,j-l}\tilde{\pi}_{Z_1Z_21}^{(l)} = 0 \quad \textit{for} \quad j = 2, 3, \ldots, h-1$$

$$\Longrightarrow \quad \tilde{\pi}_{XZ_2j} = 0, \quad \textit{for} \quad j = 1, 2, \ldots, h-1.$$

In particular, if $Y \not\to_\infty X|I_{XZ}$, at least one of the two following properties must hold :

$$(5.20) \qquad Y \underset{1}{\not\to} Z_1|I_{XZ} \quad \textit{and} \quad Z_2 \underset{1}{\not\to} X|I_{XYZ_1} \ \textit{(separation condition)} ;$$

$$(5.21) \qquad Y \underset{1}{\not\to} Z_1|I_{XZ} \quad \textit{and there is at least one} \quad j \geq 2 \ \textit{such that} \quad \sum_{l=1}^{j-1} \tilde{\pi}_{XZ_1,j-l}\tilde{\pi}_{Z_1Z_21}^{(l)} \neq 0 .$$

From the latter proposition, we see that the separation condition is necessary for $Y \not\to_{\infty} X|I_{XZ}$ in two relatively simple cases of interest [when the process $W(t)$ is regular] :

(1) $Z_1 \not\to_1 X|I_{XZ_2Y}$, i.e. $\tilde{\pi}_{XZ_1j} = 0$ for $j \geq 1$; or

(2) for any $l \geq 1$, the matrix $\pi^{(l)}_{ZZ_1}$ is proportional to an identity matrix of dimension m_3 (scalar matrix), because then $\tilde{\pi}^{(l)}_{Z_1Z_1} = Q_1\pi^{(l)}_{ZZ_1}Q_2' = 0$.

In particular, the latter condition is automatically satisfied when $Z(t)$ is univariate ($m_3 = 1$), in which case we have either $Q_1 = 0$ or $Q_2 = 0$. We can thus state the following useful corollary.

Corollary 5.1 *(Separation criterion when Z is univariate) : Under the assumptions (3.1) to (3.4) with $Z(t)$ univariate ($m_3 = 1$), $Y \not\to_{\infty} X|I_{XZ}$ if and only if at least one of the two following conditions is satisfied :*

(5.22) $$Y \not\to_1 \begin{bmatrix} X \\ Z \end{bmatrix} |I_{XZ} \; ;$$

(5.23) $$\begin{bmatrix} Y \\ Z \end{bmatrix} \not\to_1 X|I_X \; .$$

From Proposition 2.4, we know that the conditions (5.22) and (5.23) are each sufficient to have $Y \not\to_{\infty} X|I_{XZ}$, irrespective of the dimension of $Z(t)$. When $Z(t)$ is univariate, at least one of these two conditions must be satisfied to have $Y \not\to_{\infty} X|I_{XZ}$, because **internal** compensations analogous to those represented by (5.18) cannot take place. It is also interesting to note that the verification of at least one of the two conditions (5.22) and (5.23) corresponds to the definition of "non-causality" proposed by Hsiao (1982, p.247, Definition 3) for the case where X, Y and Z are univariate. From the above results, we see that Hsiao's definition is not generally equivalent to non-causality at all horizons (except precisely for a trivariate process).

When $Z(t)$ is multivariate, we may wish in a first step to clarify the underlying causality structure by looking at relatively simple necessary conditions for non-causality. Though a number of such conditions are given by Corollary 3.4, these apply only to causality at all horizons and may not be the most convenient for testing purposes. It is possible to obtain such necessary conditions by considering **rank conditions** on the autoregressive coefficient matrices π_j [as in Ahn and Reinsel (1988)] and $\pi_j^{(h)}$.

For that purpose, let us consider the lines $\pi^{(k)}_{x_iZ_1} = \psi_{x_iZk}, i = 1, \ldots, m_1$, and $R^{(k)}_{x_iZ}, i = 1, \ldots, m_1$, of the matrices $\pi^{(k)}_{XZ_1}$ and $R^{(k)}_{XZ}$ respectively [see Lemmas 3.1 and 3.2], for $1 \leq k \leq h$, and define $E^{(h)}_{XZ}$ and $F^{(h)}_{XZ}$ as the subspaces generated by the transposed of these lines :

(5.24) $$\begin{aligned} E^{(h)}_{XZ} &= span\{\pi^{(k)\prime}_{x_iZ_1} \; : \; 1 \leq i \leq m_1, \; 1 \leq k \leq h\} \\ &= span\{\psi'_{x_iZk} \; : \; 1 \leq i \leq m_1, \; 1 \leq k \leq h\}, \\ F^{(h)}_{XZ} &= span\{R^{(k)\prime}_{x_iZ} \; : \; 1 \leq i \leq m_1, \; 1 \leq k \leq h\}, \end{aligned}$$

where $h \geq 1$. Consider also the associated matrices :

(5.25) $$P^{(h)}_{XZ} = \begin{bmatrix} \pi^{(1)}_{XZ_1} \\ \pi^{(2)}_{XZ_1} \\ \vdots \\ \pi^{(h)}_{XZ_1} \end{bmatrix}, Q^{(h-1)}_{XZ} = \begin{bmatrix} R^{(1)}_{XZ} \\ R^{(2)}_{XZ} \\ \vdots \\ R^{(h)}_{XZ} \end{bmatrix},$$

$$S^{(k)}_{ZY} = [\pi_{ZY1}, \pi_{ZY2}, \ldots, \pi_{ZYk}] \; .$$

Using the orthogonality condition given by (3.15)-(3.16), it is then easy to see that the following proposition must hold.

Proposition 5.5 *(Rank conditions for non-causality up to horizon h) : Under the assumptions (3.1) to (3.4), with the notations (5.1) and (5.24)-(5.25), each one of the following conditions is necessary for $Y \not\to_{(h)} X|I_{XZ}$, where $h \geq 2$:*

(i) $dim(E^{(h-1)}_{XZ}) + dim(E_{ZY}) \leq m_3$;

(ii) $dim(F^{(h-1)}_{XZ}) + dim(E_{ZY}) \leq m_3$;

(iii) $rank(P^{(h-1)}_{XZ}) + dim(E_{ZY}) \leq m_3$;

(iv) $rank(P^{(h-1)}_{XZ}) + rank(S^{(k)}_{ZY}) \leq m_3$, *for all* $k \geq 1$;

(v) $rank(Q^{(h-1)}_{XZ}) + dim(E_{ZY}) \leq m_3$;

(vi) $rank(Q^{(h-1)}_{XZ}) + rank(S^{(k)}_{ZY}) \leq m_3$, *for all* $k \geq 1$.

Rejection of any one of the conditions (i) to (vi) above implies that the hypothesis $Y \not\to_{(h)} X|I_{XZ}$ has to be rejected. For example, each one of the following conditions implies that $Y \not\to_{(2)} X|I$ cannot hold : (1) $dim(E^{(1)}_{XZ}) \geq 1$ and $dim(E_{ZY}) = m_3$; (2) $dim(E^{(1)}_{XY}) = m_3$ and $dim(E_{ZY}) \geq 1$; (3) $rank(P^{(1)}_{XZ}) \geq 1$ and $rank(S^{(m_3)}_{ZY}) = m_3$; (4) $rank(P^{(1)}_{XZ}) = m_3$ and $rank(S^{(m_3)}_{ZY}) \geq 1$. It is of course easy to find a large number of similar examples.

6 Concluding remarks

In this paper, we have proposed a general notion of causality (on non-causality) at various horizons which allows one to distinguish between short-run (short horizon) and long-run (long horizon) causality, and thus separates the direct effects of a variable (horizon one) from its indirect effects at longer horizons. Under quite general conditions which allow both stationary and non-stationary processes, we observed that non-causality at a given horizon h between a vector X and a vector Y can be reduced to non-causality between each component of X and each component of Y (**componentwise characterization** of non-causality), and we showed that non-causality at horizon one entails non-causality at all horizons in two cases: when X and Y include all the variables considered in the analysis (**exhaustivity** condition), or when the variables of the system can be separated in two subvectors which do not cause each other at horizon one (**separation** condition). The separation condition is especially convenient because it reduces non-causality at all horizons to simple horizon one non-causalities. In other cases, non-causality at horizon one does not generally entail non-causality at longer horizons.

For multivariate processes possessing an autoregressive representation (possibly of infinite order), we derived general parametric necessary and sufficient conditions for non-causality up to any given horizon. In particular, we gave a "causality chain" interpretation of these conditions which throws light on the relationship between causalities at horizons greater than one by showing how such causalities are related to "indirect causal effects" going through auxiliary variables. We also studied the relationship between causality at various horizons and impulse coefficients, and we observed that zero restrictions on the latter are necessary (but not sufficient) for non-causality at all horizons to hold. For finite order VAR processes, we showed that non-causality up to a given finite horizon is sufficient to have non-causality at all horizons, so that hypothesis of non-causality at various horizons can be tested in such cases by considering only a finite set of restrictions. Finally, we introduced the notion of **separating causal decomposition** by which the auxiliary variables Z of a system are decomposed in two components Z_1 and Z_2, such that Y does not cause Z_1 at horizon one while Y does cause Z_2 at horizon one, and we showed that the indirect effects (i.e. effects with horizons greater than one) between Y and X must go through Z_2 (but not Z_1). From the separating causal decomposition, it follows that a generalized separation condition is necessary (as well as sufficient) for $Y \not\rightarrow_{(h)} X$ to hold. Furthermore, when there is only one auxiliary variable (Z univariate), the disjunction of two simple separation conditions is both necessary and sufficient for non-causality at all horizons to hold (for example, this condition applies to trivariate autoregressions).

An obvious application of the above results is the development of tests for hypothesis of the form $Y \not\rightarrow_{(h)} X$ or $Y \not\rightarrow_{\infty} X$. 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, as in Corollary 5.1), it is possible to reduce these nonlinear conditions to combinations of linear conditions which can be tested by testing separately causality hypothesis 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 non-standard 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 (1994)].

APPENDIX

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, $i=1,\ldots,m_1$. Consider now the equivalence between (i) and (iii). If $Y \not\rightarrow_h X|I$, we have by definition :

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

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

$$P\left[X(t+h)|I(t)+y_j(\omega,t]\right] = P\left[P\left[X(t+h)|I(t)+Y(\omega,t]\right]|I(t)+y_j(\omega,t]\right] \ ,$$

$$= P\left[P[X(t+h)|I(t)]|I(t)+y_j(\omega,t]\right] = P[X(t+h)|I(t)] \ , \quad \forall t > \omega \ ,$$

for $j=1,\ldots,m_2$, which means that $y_j \not\rightarrow_h X|I$ for $j=1,\ldots,m_2$. Thus (i) $\Rightarrow$ (iii). Conversely, if (iii) holds, we have

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

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. $Y \not\rightarrow_h X|I$, and we can conclude that (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 $Y \not\rightarrow_h X|I$, 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 $j=1,\ldots,m_2$:

$$P\left[X(t+h)|I_{(j)}(t)\right] = P\left[P[X(t+h)|I(t)+Y(\omega,t]]|I_{(j)}(t)\right]$$

$$= P\left[P[X(t+h)|I(t)]|I_{(j)}(t)\right] = P[X(t+h)|I(t)] \ , \quad \forall t > \omega \ ,$$

which means that $y_j \not\rightarrow_h X|I_{(j)}$. On the other hand, we can have

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

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 $Y \not\rightarrow_h X|I$ may not hold.

Q.E.D.

Proof of Proposition 2.3 : Since (ii) $\Rightarrow$ (i) by definition, we need to show first that $Y \nrightarrow_1 X|I \Rightarrow Y \nrightarrow_{(h)} X|I$.

The proof is done by induction. Suppose that $Y \nrightarrow_{(h)} X|I$. Then, by the properties of iterated projections,

(A.3) $$P[X(t+h+1)|I(t)+Y(\omega,t]]$$
$$= P[P[[X(t+h+1)|I(t+h)+Y(\omega,t+h]]\,|\,|I(t)+Y(\omega,t]] ,$$

$\forall t > \omega$. Further, since $Y \nrightarrow_1 X|I$,

(A.4) $$P[X(t+h+1)|I(t+h)+Y(\omega,t+h]] = P[X(t+h+1)|I(t+h)] ,$$

$\forall t > \omega$. But $P[X(t+h+1)|I(t+h)]$ 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 $x_i(\tau), i = 1,\ldots,m_1, t+1 \le \tau \le t+h$, and thus we can write

(A.5) $$P[X(t+h+1)|I(t+h)] = a_h(t) + b_h(t)$$

where $a_h(t) \in I(t)$ and $b_h(t) \in X[t+1,t+h]$. Further, since $Y \nrightarrow_{(h)} X|I$, each component $x_i(\tau), i = 1,\ldots,m_1, t+1 \le \tau \le t+h$, satisfies

$$P[x_i(\tau)|I(t)+Y(\omega,t]] = P[x_i(\tau)|I(t)] , \quad \forall t > \omega ,$$

which implies that

(A.6) $$P[b_h(t)|I(t)+Y(\omega,t]] = P[b_h(t)|I(t)] , \quad \forall t > \omega .$$

From (A.3)-(A.6), we then get

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

so that each component of $P[X(t+h+1)|I(t)+Y(\omega,t]]$ belongs to $I(t)$. Consequently

$$P[X(t+h+1)|I(t)+Y(\omega,t]] = P[X(t+h+1)|I(t)] , \quad \forall t > \omega ,$$

which means that $Y \nrightarrow_{h+1} X|I$. Thus (i) $\Leftrightarrow$ (ii). The equivalence between (ii) and (iii) follows trivially from Definition 2.2.

Q.E.D.

Proof of Proposition 2.4 : By Proposition 2.3, condition (2.8) implies that

$$\begin{bmatrix} Y \\ Z_2 \end{bmatrix} \nrightarrow_\infty \begin{bmatrix} Y \\ Z_1 \end{bmatrix} | I_{XZ_1} ,$$

hence, by Proposition 2.1,

$$\begin{bmatrix} Y \\ Z_2 \end{bmatrix} \nrightarrow_\infty X | I_{XZ_1} .$$

In other words, for any $h \in \mathbb{N}$,

$$P[X(t+h)|I_{XZ_1}(t) + Z_2(\omega,t] + Y(\omega,t]] = P[X(t+h)|I_{XZ_1}(t)] , \quad \forall t > \omega ,$$

or consequently,

$$P[X(t+h)|I_{XZ}(t) + Y(\omega,t]] = P[X(t+h)|I_{XZ_1}(t)] , \quad \forall t > \omega .$$

Thus each component of $P[X(t+h)|I_{XZ}(t)+Y(\omega,t]]$ is an element of $I_{XZ_1}(t) \subseteq I_{XZ}(t)$, hence

$$P[X(t+h)|I_{XZ}(t) + Y(\omega,t]] = P[X(t+h)|I_{XZ}(t)] , \quad \forall t > \omega .$$

for any $h \in \mathbb{N}$, which means that $Y \nrightarrow_\infty X|I_{XZ}$.

Q.E.D.

Proof of Proposition 3.1 : Let $t \in \mathbb{Z}$ and $t > \omega$. The proof is done by recurrence (mathematical induction) on h. From (3.1), we have

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

which yields (3.7) for $h = 1$ with $\pi_j^{(1)} = \pi_j$, because $a(t+1)$ is orthogonal to the space spanned by $\{W(\tau) : \tau \le t\}$ and $\mu(t+1)$, so that $P[a(t+1)|H + W(-\infty,t]] = 0$. Suppose now that (3.7) and (3.8) hold up to h, where $h \ge 1$. By (3.1), we have

$$W(t+h+1) = \mu(t+h+1) + \sum_{j=1}^{\infty} \pi_j W(t+h+1-j) + a(t+h+1) ,$$

hence, denoting $\bar{I}(t) = H + W(-\infty,t]$,

$$\begin{aligned} P[W(t+h+1)|\bar{I}(t)] &= \mu(t+h+1) + \sum_{j=1}^{\infty} \pi_j P[W(t+h+1-j)|\bar{I}(t)] \\ &= \mu(t+h+1) + \sum_{j=1}^{h} \pi_j P[W(t+h+1-j)|\bar{I}(t)] \\ &\quad + \sum_{j=h+1}^{\infty} \pi_j W(t+h+1-j) . \end{aligned}$$

Then, by the recurrence assumption, we have for $j = 1,\ldots,h$,

$$P[W(t+h+1-j)|\bar{I}(t)] = \sum_{k=0}^{h-j} \pi_1^{(k)} \mu(t+h+1-j-k)$$

$$+\sum_{k=1}^{\infty}\pi_k^{(h+1-j)}W(t+1-k)\,,$$

hence

$$P[W(t+h+1)|\tilde{I}(t)] = \mu(t+h+1)+\sum_{j=1}^{h}\pi_j\left[\sum_{k=0}^{h-j}\pi_1^{(k)}\mu(t+h+1-j-k)\right]$$

$$+\sum_{j=1}^{h}\pi_j\left[\sum_{k=1}^{\infty}\pi_k^{(h+1-j)}W(t+1-k)\right]+\sum_{j=h+1}^{\infty}\pi_jW(t+h+1-j)$$

$$=\pi_1^{(0)}\mu(t+h+1)+\sum_{l=1}^{h}\left[\sum_{k=0}^{l-1}\pi_{l-k}\pi_1^{(k)}\right]\mu(t+h+1-l)$$

$$+\sum_{k=1}^{\infty}\left[\sum_{l=1}^{h}\pi_{h+1-l}\pi_k^{(l)}\right]W(t+1-k)+\sum_{k=1}^{\infty}\pi_{k+h}W(t+1-k)$$

$$=\pi_1^{(0)}\mu(t+h+1)+\sum_{l=1}^{h}\left[\pi_l+\sum_{k=1}^{l-1}\pi_{l-k}\pi_1^{(k)}\right]\mu(t+h+1-l)$$

$$+\sum_{k=1}^{\infty}\left[\pi_{k+h}+\sum_{l=1}^{h}\pi_{h+1-l}\pi_k^{(l)}\right]W(t+1-k)$$

where we set $\pi_1^{(0)}=I_m$ and $\sum_{k=1}^{0}\pi_{l-k}\pi_1^{(k)}=0$. We thus have :

$$\text{(A.7)}\qquad P[W(t+h+1)|\tilde{I}(t)] = \sum_{k=0}^{h}\pi_1^{(k)}\mu(t+h+1-k)$$

$$+\sum_{j=1}^{\infty}\pi_j^{(h+1)}W(t+1-j)$$

with

$$\pi_j^{(h+1)}=\pi_{j+h}+\sum_{l=1}^{h}\pi_{h+1-l}\pi_j^{(l)},$$

which proves (3.7) and (3.8) by recurrence.

Since (3.7) holds for any process of the form (3.1) irrespective of the covariance matrices of $a(t)$, let us consider a process $W(t)$ which satisfies (3.1) for $t\geq 1$, with $E[a(t)a(t)']=\Delta$ for $t\geq 1$, such that $\det(\Delta)\neq 0$ and $W(t)=0$ for $t\leq 0$. It is clear that such a process always exists. Then by the principle of iterated projections, we have for all $t\in I\!N$ and $h\in I\!N$,

$$\text{(A.8)}\qquad P[W(t+h+1)|\tilde{I}(t)] = P[P[W(t+h+1)|\tilde{I}(t+1)]|\tilde{I}(t)]$$

$$\begin{aligned}
&= P\left[\textstyle\sum_{j=1}^{\infty}\pi_j^{(h)}W(t+2-j)|\tilde{I}(t)\right]+\textstyle\sum_{k=0}^{h}\pi_1^{(k)}\mu(t+h+1-k)\\
&= \pi_1^{(h)}P[W(t+1)|\tilde{I}(t)]+\textstyle\sum_{k=1}^{\infty}\pi_{k+1}^{(h)}W(t+1-k)\\
&\quad+\textstyle\sum_{k=0}^{h}\pi_1^{(k)}\mu(t+h+1-k)\\
&= \textstyle\sum_{j=1}^{\infty}\left[\pi_1^{(h)}\pi_j+\pi_{j+1}^{(h)}\right]W(t+1-j)+\textstyle\sum_{k=0}^{h}\pi_1^{(k)}\mu(t+h+1-k).
\end{aligned}$$

Subtracting (A.8) from (A.7), we see that

$$\text{(A.9)}\qquad \sum_{j=1}^{\infty}\left[\pi_j^{(h+1)}-\pi_1^{(h)}\pi_j-\pi_{j+1}^{(h)}\right]W(t+1-j)=0\,,\qquad \forall t,h\in I\!N\,.$$

Then, multiplying (A.9) by $a(t)$ and taking the expected value for $t>0$, we find :

$$\left[\pi_1^{(h+1)}-\pi_1^{(h)}\pi_1-\pi_2^{(h)}\right]\Delta=0\,,\quad \forall h\in I\!N\,,$$

hence, since Δ is invertible,

$$\pi_1^{(h+1)}-\pi_1^{(h)}\pi_1-\pi_2^{(h)}=0\,,\qquad \forall h\in I\!N\,.$$

We thus have

$$\sum_{j=2}^{\infty}\left[\pi_1^{(h+1)}-\pi_1^{(h)}\pi_j-\pi_{j+1}^{(h)}\right]W(t+1-j)=0\,,\qquad \forall t,h\in I\!N\,.$$

Similarly, multiplying by $a(t-1)$, for $t>2$, and taking the expected value, we find :

$$\pi_2^{(h+1)}-\pi_1^{(h)}\pi_2-\pi_3^{(h)}=0\,,\qquad \forall h\in I\!N\,,$$

and, upon pursuing this process,

$$\pi_j^{(h+1)}-\pi_1^{(h)}\pi_j-\pi_{j+1}^{(h)}=0\,,\qquad \forall j,h\in I\!N\,.$$

This completes the proof of (3.9) by recurrence.

Q.E.D.

Proof of Theorem 3.1 : Let $t\in Z\!\!Z, t>\omega$, and denote $I(t)=I_{XZ}(t)$. We deduce from Proposition 3.1 that

$$P[X(t+h)|I(t)+Y(-\infty,t]]=\sum_{j=1}^{\infty}\left[\pi_{XX_j}^{(h)}X(t+1-j)+\pi_{XY_j}^{(h)}Y(t+1-j)\right.$$

$$\left.+\,\pi_{XZ_j}^{(h)}Z(t+1-j)\right]+\mu_h(t)$$

where $\mu_h(t)=\sum_{k=0}^{h-1}\pi_1^{(k)}\mu(t+h-k)$. Condition (3.11) then implies that

$$P[X(t+h)|I(t)+Y(-\infty,t]]=\sum_{j=1}^{\infty}\left[\pi_{XX_j}^{(h)}X(t+1-j)+\pi_{XZ_j}^{(h)}Z(t+1-j)\right]+\mu_h(t)$$

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 (3.11) is sufficient for $Y \not\to_h X|I$.

Suppose now that the matrices $E[a(t)a(t)']$ are non-singular for $t > \omega$. If $Y \not\to_h X|I$, 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 $\pi_{X.j}^{(h)} = \left[\pi_{XXj}^{(h)}, \pi_{XYj}^{(h)}, \pi_{XZj}^{(h)}\right]$, can also expressed as the limit in quadratic mean (q.m.) of a sequence

$$U_T = \sum_{j=1}^{T} \phi_j^{(T)} W(t+1-j) \; , \; T \in \mathbb{N} \, ,$$

where the components of U_T all belong to $I(t)$:

$$U_T = \sum_{j=1}^{T} \left[\phi_{Xj}^{(T)} X(t+1-j) + \phi_{Zj}^{(T)} Z(t+1-j)\right] .$$

Consequently, defining $\phi_j^{(T)} = 0$ for $j > T$, and

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

we see that the vector $\bar{U}_T(t)$ converges in q.m. to zero, hence

$$E[\bar{U}_T(t)a(t)'] = \left[\pi_{X.1}^{(h)} - \phi_1^{(T)}\right] E[a(t)a(t)'] \underset{T\to\infty}{\longrightarrow} 0 \, ,$$

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 non-singular, we must have $\pi_{X.1}^{(h)} - \phi_1^{(T)} \underset{T\to\infty}{\longrightarrow} 0$. 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 $t-1 > \omega$) allows one to conclude that $\pi_{XY2}^{(h)} = 0$. Proceeding analogously for increasing j, we see that $\pi_{XYj}^{(h)} = 0$ for $j = 1, 2, 3 \ldots$

Q.E.D.

Proof of Lemma 3.1 : We shall prove (3.13) 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 (3.9), it follows that (3.12) must hold. In particular, we must have :

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

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

From the assumption

(A.10) $$\pi_{XYj}^{(k)} = 0 \, , \qquad \forall j \in \mathbb{N} \, , \; k = 1, \ldots, h,$$

we then have $\pi_{XYj} = \pi_{XYj}^{(2)} = \pi_{XYj}^{(h-1)} = 0$ for all j, so that

$$\pi_{XYj}^{(2)} = \pi_{XZ1}\pi_{ZYj} = 0 \, , \qquad \forall j \in \mathbb{N} \, ,$$

and

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

which is identical with (3.13) with $p = 2$.

Let us now call $P(h)$ the property obtained when (3.13) 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 (3.13) for all integers p such that $2 \leq p \leq h$.

From the first step, we know that (3.13) holds for $p = 2$. By the mathematical induction principle, it will sufficient to show that (3.13) must hold for $p = \bar{p}+1$ whenever it does for $p = 1, \ldots, \bar{p}$ (where $\bar{p} < h$). If we assume (3.13) for $p = 2, \ldots, \bar{p} < h$ and take p to be any integer such that $2 \leq p \leq \bar{p}$, we have

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

where we write (to simplify the notation)

$$\sum_{J(\ell)} \prod_i \pi_{ZZi}^{n_i} = \sum_{J(\ell)} \left[\prod_{i=1}^{\ell} \pi_{ZZi}^{n_i}\right] .$$

But, by (3.9), we have

$$\pi_\ell^{(h-p+1)} = \pi_1^{(h-p)}\pi_\ell + \pi_{\ell+1}^{(h-p)}$$

and, in particular,

$$\pi_{XZ\ell}^{(h-p+1)} = \pi_{XX1}^{(h-p)}\pi_{XZ\ell} + \pi_{XY1}^{(h-p)}\pi_{YZ\ell} + \pi_{XZ1}^{(h-p)}\pi_{ZZ\ell} + \pi_{XZ,\ell+1}^{(h-p)} .$$

By (A.10), we have $\pi_{XY1}^{(h-p)} = 0$, hence

(A.11) $$\left[\sum_{\ell=1}^{p} \pi_{XZ\ell}^{(h-p+1)} \sum_{J(p-\ell)} \prod_i \pi_{ZZi}^{n_i}\right] \pi_{ZYj}$$

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

$$+ \pi_{XX1}^{(h-p)} \left[\sum_{\ell=1}^{p} \pi_{XZ\ell} \left[\sum_{J(p-\ell)} \prod_i \pi_{ZZi}^{n_i}\right]\right] \pi_{ZYj} \, .$$

We will now show that the second term of the latter sum is zero. Since $p \le h-1$, we know that $P(p)$ must hold (from the recurrence assumption), i.e.

$$\pi_{XZ1}^{(p)}\pi_{ZYj} = \left[\sum_{\ell=1}^{q} \pi_{XZ\ell}^{(p-q+1)} \sum_{J(q-\ell)} \prod_i \pi_{ZZi}^{n_i}\right] \pi_{ZYj}\,, \quad \text{for } 2 \le q \le p\,.$$

In particular, by taking $q = p$, we get

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

Since $\pi_{XYj}^{(k)} = 0$, $\forall j \in \mathbb{N}$, for $k = 1, \ldots, p+1$ (for $p+1 \le h$), we must have $\pi_{XZ1}^{(p)}\pi_{ZYj} = 0$ [see (3.12)], which implies that the second term of (A.11) is zero. Consequently

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

$$= \left[\sum_{\ell=2}^{p+1}\pi_{XZ\ell}^{(h-p)}\sum_{J(p+1-\ell)}\prod_i \pi_{ZZi}^{n_i}\right]\pi_{ZYj} + \pi_{XZ1}^{(h-p)}\left[\sum_{\ell=1}^{p}\pi_{ZZ\ell}\sum_{J(p-\ell)}\prod_i \pi_{ZZi}^{n_i}\right]\pi_{ZYj}\,.$$

But it is clear from the definition of $J(p)$ that

$$\sum_{\ell=1}^{p}\pi_{ZZ\ell}\sum_{J(p-\ell)}\prod_i \pi_{ZZi}^{n_i} = \sum_{J(p)}\prod_i \pi_{ZZi}^{n_i}\,,$$

hence

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

which means that (3.13) holds with p replaced by $p+1$. In particular, we can take $p = \bar{p}$, so that (3.13) holds for $p = 2, 3, \ldots, \bar{p}+1$, hence (by recurrence) for $p = 2, \ldots, h$. Property $P(h)$ is thus established and the proof is complete.

Q.E.D.

Proof of Lemma 3.2 : We need to show that the equivalence between (3.14), (3.15) and (3.16) holds for any $h \ge 2$. Again we shall proceed by recurrence considering first the equivalence between (3.14) and (3.15), and then the one between (3.14) and (3.16).

The equivalence between (3.14) and (3.15) follows by applying the recursion (3.12), which is implied by (3.9). For $h = 2$, the result clearly holds since $\pi_{XYj}^{(1)} = \pi_{XYj}$, and $\pi_{XYj} = 0$ for all j implies [on applying (3.12)] :

$$\pi_{XYj}^{(2)} = \pi_{XZ1}^{(1)}\pi_{ZYj}\,, \qquad \forall j \in \mathbb{N}\,.$$

Suppose now that the equivalence holds for some $h \ge 2$. Then, given that $\pi_{XYj}^{(k)} = 0$, $\forall j \in \mathbb{N}$, for $k = 1, \ldots, h$, it follows from (3.12) that

$$\pi_{XYj}^{(h+1)} = \pi_{XZ1}^{(h)}\pi_{ZYj}\,, \quad \forall j \in \mathbb{N}\,,$$

and the equivalence holds for $h+1$. The equivalence between (3.14) and (3.15) for any $h \ge 2$ follows by recurrence.

The equivalence between (3.14) and (3.16) holds for $h = 2$ because the criteria (3.15) and (3.16) are then identical. Suppose now that the equivalence between (3.14) and (3.16) holds for some $h \ge 2$. Then given that $\pi_{XYj}^{(k)} = 0$, $\forall j \in \mathbb{N}$, $k = 1, \ldots, h$, we see from (3.12) and Lemma 3.1 that

$$\pi_{XYj}^{(h+1)} = \pi_{XZ1}^{(h)}\pi_{ZYj} = R_{XZ}^{(h)}\pi_{ZYj}\,, \quad \forall j \in \mathbb{N}\,,$$

and the equivalence between (3.14) and (3.16) also holds for $h+1$. The equivalence between (3.14) and (3.16) thus holds for all $h \ge 2$.

Q.E.D.

Proof of Theorem 3.2 : Under the assumptions (3.1) to (3.3), it follows from Theorem 3.1 that condition (3.14) is sufficient for $Y \not\to_{(h)} X | I_{XZ}$. Further, under these conditions, the recursion (3.9) applies so that (3.15) and (3.16) are each equivalent to (3.14) by Lemma 3.2. When $W(t)$ is a regular process, Theorem 3.1 also entails that (3.14) is necessary and sufficient for $Y \not\to_{(h)} X | I_{XZ}$, and each one of the three conditions (3.14), (3.15) and (3.16) is thus necessary and sufficient.

Q.E.D.

Proof of Lemma 3.3 : From the identity $\pi(z)\psi(z) = I_m$, we see easily that the matrices ψ_h satisfy the recursion

(A.12) $$\psi_0 = I_m, \quad \psi_h = \sum_{k=1}^{h}\pi_k\psi_{h-k}\,, \quad h = 1, 2, \ldots\,.$$

Now the recursion (3.8) can be rewritten

$$\pi_1^{(1)} = \pi_1\,,\ \pi_1^{(h)} = \pi_h\pi_1^{(0)} + \sum_{\ell=1}^{h-1}\pi_{h-\ell}\pi_1^{(\ell)},\ h = 2, 3, \ldots$$

with $\pi_1^{(0)} = I_m$, or equivalently

(A.13) $$\pi_1^{(0)} = I_m, \quad \pi_1^{(h)} = \sum_{k=1}^{h}\pi_k\pi_1^{(h-k)},\ h = 1, 2, \ldots\,.$$

Clearly the recursions (A.12) and (A.13) are equivalent and so they define the same sequence of matrices, i.e. we must have $\pi_1^{(h)} = \psi_h$, for all $h \ge 0$.

Q.E.D.

Proof of Lemma 3.4 : From the definitions (3.18), we see that

$$\pi_{ZZ}(z) = I_{m_3} - \sum_{k=1}^{\infty}\pi_{ZZk}z^k,\ \pi_{XZ}(z) = -\sum_{j=1}^{\infty}\pi_{XZj}z^j,$$

hence

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

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

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

$$= \sum_{k=1}^{\infty} R_{XZ}^{(k)} z^k = R_{XZ}(z)\,.$$

Q.E.D.

Proof of Corollary 3.4 : By Theorem 3.2, the condition

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

is necessary and sufficient for $Y \not\to_\infty X | I_{XZ}$. This condition can be written in terms of generating functions, as follows :

$$\text{(A.14)} \qquad \pi_{XY}(z) \equiv 0 \quad \text{and} \quad R_{XZ}(z)\pi_{ZYj} \equiv 0\ ,\quad \forall j \geq 1\ .$$

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 $Y \not\to_\infty X | I_{XZ}$:

$$\pi_{XY}(z) \equiv 0 \quad \text{and} \quad R_{XZ}(z)\pi_{ZY}(z) \equiv 0\ .$$

Using Lemma 3.4, the latter can then be written :

$$\pi_{XY}(z) \equiv \quad \text{and} \quad \pi_{XZ}(z)\pi_{ZZ}(z)^{-1}\pi_{XY}(z) \equiv 0\ .$$

This yields condition (3.26).

Let us now assume that $z \in \mathbb{C}$ and $|z| < \delta$ so that $\pi(z)$ converges in $\mathbb{C}$ for $|z| < \delta$. To show the equivalence between (3.26) and (3.27), it is then sufficient to use standard formulae for the inversion of partitioned matrices (where we omit the symbol z to simplify the notation) :

$$\psi_{XY} = -\pi_{XX.Z}^{-1}\left[\pi_{XY} - \pi_{XZ}\pi_{ZZ}^{-1}\pi_{ZY}\right]\tilde{\pi}_{YY}^{-1}$$

where $\pi_{XX.Z} = \pi_{XX} - \pi_{XZ}\pi_{ZZ}^{-1}\pi_{ZX}$ and

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

Since $\pi(0) = I_m$ (a non-singular 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

$$\pi_{XY}(z) \equiv 0 \text{ and } \pi_{XZ}(z)\pi_{ZZ}(z)^{-1}\pi_{ZY}(z) \equiv 0 \Leftrightarrow \pi_{XY}(z) \equiv 0 \text{ and } \psi_{XY}(z) \equiv 0,$$

which shows the equivalence between (3.26) and (3.27). The equivalence between (3.27) and (3.28) is deduced from the one between (3.26) and (3.27) on permuting π and ψ, as these play symmetric roles in (3.27).

Q.E.D.

Proof of Corollary 3.5 : When $Z(t)$ is univariate,

$$R_{XZ}(z) = -\pi_{XZ}(z)\pi_{ZZ}(z)^{-1} = [R_{x_1Z}(z), \ldots, R_{x_{m_1}Z}(z)]'$$

is an $m_1 \times 1$ vector of scalar formal series $R_{x_iZ}(z)$ 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

$$\text{(A.15)} \qquad R_{x_iZ}(z)\pi_{Zy_j}(z) \equiv 0$$

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 $\pi(0) = I_m$ and so it is non-singular in an open neighborhood of $z = 0$]. Then the series $R_{x_iZ}(z)$ and $\pi_{Zy_j}(z)$ represent functions which are analytic in the circle $|z| < \bar{\delta}$, and (A.15) can hold only if either $R_{x_iZ}(z) \equiv 0$ or $\pi_{Zy_j}(z) \equiv 0$. Consequently, in all cases, we must have

$$\text{(A.16)} \qquad R_{x_iZ}^{(k)}\pi_{Zy_\ell j} = 0\ , \qquad \forall k \in \mathbb{N}\ ,\ \forall j \in \mathbb{N}\ ,$$

for $i = 1, \ldots, m_1$ and $\ell = 1, \ldots, m_2$, or equivalently

$$R_{XZ}^{(k)}\pi_{ZYj} \equiv 0\ , \qquad \forall k \in \mathbb{N}\ ,\ \forall j \in \mathbb{N}\ .$$

Thus the condition (3.32) implies (3.16) when $Z(t)$ is univariate, and Corollary 3.5 then follows from Theorem 3.2 and Corollary 3.4.

Q.E.D.

Proof of Proposition 4.1 : 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 (1994).

Lemma 6.1 *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 $|z| < \delta$, where $\delta > 0$, and such that*

$$\sum_{j=0}^{\infty} a_j z^j = \left(\sum_{k=0}^{p} b_k z^k\right)\left(\sum_{\ell=0}^{\infty} c_\ell z^\ell\right)$$

for $|z| < \delta$, where $0 \leq p < \infty, c_0 = 1$ and the power series $\sum_{\ell=0}^{\infty} c_\ell z^\ell$ converges for $|z| < \delta$. Then

$$a_j = 0\ ,\quad \forall j \geq 0 \quad \Leftrightarrow \quad a_j = 0\ ,\ \textit{for } j = 0, 1, \ldots, p\ .$$

Given Theorem 3.2, we will now prove Proposition 4.1 by showing that

$$R^{(k)}_{XZ}\pi_{ZYj} = 0, \quad \forall j \geq 1, k = 1, 2, \ldots, m_3 p ,$$

implies

$$R^{(k)}_{XZ}\pi_{ZYj} = 0 , \quad \forall j \geq 1 , \forall k \geq 1 ,$$

where the latter identity is equivalent to

$$R_{XZ}(z)\pi_{ZYj} \equiv 0 , \quad \forall j \geq 1 .$$

From Lemma 3.4, 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 $(x_i(t), y_\ell(t))'$ 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 $1 \leq \ell \leq m_2$. We have :

$$R_{x_i Z}(z)\pi_{Z y_\ell j} = -\{det[\pi_{ZZ}(z)]\}^{-1}\pi_{x_i Z}(z)\pi^{*}_{ZZ}(z)\pi_{Z y_\ell j} ,$$

where $\{det[\pi_{ZZ}(z)]\}^{-1} = \sum_{\ell=0}^{\infty} c_\ell z^\ell$ is a formal series with $c_0 = 1$ and $\pi_{x_i Z}(z)\ \pi^{*}_{ZZ}(z)\pi_{Z y_\ell j}$ is a polynomial of degree not greater than $m_3 p$. Using Lemma 4.1, we can then state that

$$R_{x_i Z}(z)\pi_{Z y_\ell j} = 0 , \quad \forall j \geq 1 ,$$

if and only if, for $j \geq 1$, the coefficients associated with powers of z not greater than $m_3 p$ in $R_{x_i Z}(z)\pi_{Z y_\ell j}$ are equal to zero :

$$R^{(k)}_{x_i Z}\pi_{Z y_\ell j} = 0 , \quad \text{for} \quad k = 1, 2, \ldots, m_3 p .$$

The result then follows from Theorem 3.2.

Q.E.D.

Proof of Proposition 4.2 : For each pair

$$(x_i(t), y_\ell(t))', \quad 1 \leq i \leq m_1 , 1 \leq \ell \leq m_2 ,$$

we have :

$$\pi_{x_i Z}(z)\pi_{ZZ}(z)^{-1}\pi_{Z y_\ell}(z) \equiv 0 \Leftrightarrow \{det[\pi_{ZZ}(z)]\}^{-1}\pi_{x_i Z}(z)\pi^{*}_{ZZ}(z)\pi_{Z y_\ell}(z) \equiv 0 ,$$

where $\pi_{x_i Z}(z)\pi^{*}_{ZZ}(z)\pi_{Z y_\ell}(z)$ is a polynomial of degree not greater than $(m_3 + 1)p$. The result then follows on applying Lemma 4.1.

Q.E.D.

Proof of Proposition 5.1 : By definition, the columns of Q_1' and Q_2' constitute bases for the subspaces $E^{\perp}_{ZY}$ and E_{ZY} respectively. Consequently, we must have :

$$\tilde{\pi}_{Z_1 Y j} = Q_1\pi_{ZYj} = 0 , \quad \forall j \geq 1 ,$$

which implies that $Y \not\to_1 Z_1 | I_{XZ}$. Furthermore, by definition again, $E^{\perp}_{ZY}$ includes all vectors $a \in R^{m_3}$ with the property $a'\pi_{ZYj} = 0$ for all $j \geq 1$, so that $Z_1(t)$ has maximal dimension : any matrix $\bar{Q}_1$ such that $\bar{Q}_1\pi_{ZYj} = 0, \forall j \geq 1$, can be written $\bar{Q}_1 = AQ_1$ for some matrix A and so any other process $\bar{Z}_1(t) = \bar{Q}_1 Z(t)$ such that $Y \not\to_1 \bar{Z}_1 | I_{XZ}$ is a linear transformation of $Z_1(t)$, i.e. $\bar{Z}_1(t) = AZ_1(t)$. Consequently, when $dim(E_{ZY}) \geq 1$, we must have :

$$\tilde{\pi}_{Z_2 Y j} = Q_2\pi_{ZYj} \neq 0 \quad \text{for some } j \geq 1 ,$$

which implies (for a regular process) that

$$Y \underset{1}{\to} Z_2 | I_{XZ} \qquad (\text{negation of } Y \underset{1}{\not\to} Z_2 | I_{XZ}) .$$

Finally, to obtain (5.9), we observe the following facts : when $dim(E_{ZY}) = 0$, we have $Z_2(t) = Q_2 = 0$ and $E^{\perp}_{ZY} = R^{m_3}$, so that we can take $Z(t) = Z_1(t)$; when $dim(E_{ZY}) = m_3$, we have $Z_1(t) = Q_1 = 0$ and $E_{ZY} = R^{m_3}$ so that we can take $Z(t) = Z_2(t)$; and when $1 \leq dim(E_{ZY}) < m_3$, we have $[Z_1(t)', Z_2(t)']' = QZ(t)$, where Q is non-singular [see (5.4)], so that

$$Z(t) = Q^{-1}\begin{bmatrix} Z_1(t) \\ Z_2(t) \end{bmatrix} = P_1 Z_1(t) + P_2 Z_2(t)$$

where $Q^{-1} = [P_1, P_2]$ is partitioned conformably with $Z_1(t)$ and $Z_2(t)$. This completes the proof.

Q.E.D.

Proof of Proposition 5.2 : Since the columns of Q_2' form a base of the space E_{ZY}, the property

$$\pi^{(k)}_{XZ1}\pi_{ZYj} = 0 , \quad \text{for } k = 1, \ldots, h-1 , \quad \forall j \geq 1 ,$$

in the characterization (3.15) of Theorem 3.2 is equivalent to

$$\pi^{(k)}_{XZ1}Q_2' = 0 , \quad \text{for } k = 1, \ldots, h-1 .$$

Since $\tilde{\pi}^{(k)}_{XZ_2 1} = \pi^{(k)}_{XZ1}Q_2'$, the result then follows from Theorem 3.2.

Q.E.D.

Proof of Proposition 5.3 : When $h = 2$, the result follows trivially from Proposition 5.2. Suppose now that $h \geq 3$, and consider an integer j such that $2 \leq j \leq h-1$. For $k = j-1$, we can then apply the recurrence (3.8) :

$$\pi_1^{(k+1)} = \pi_{k+1} + \sum_{\ell=1}^{k}\pi_{k-\ell+1}\pi_1^{(\ell)},$$

or equivalently,

$$\pi_1^{(j)} = \pi_j + \sum_{\ell=1}^{j-1}\pi_{j-\ell}\pi_1^{(\ell)}.$$

In particular,

$$\pi^{(j)}_{XZ1} = \pi_{XZj} + \sum_{\ell=1}^{j-1}\left[\pi_{XXj-\ell}\pi^{(\ell)}_{XZ1} + \pi_{XZj-\ell}\pi^{(\ell)}_{ZZ1}\right]$$

where we took into account the fact that $\pi_{XYj-\ell} = 0$ (since $Y \not\to_1 X|I_{XZ}$). Multiplying on the right by Q'_2, we find :

$$\tilde{\pi}^{(j)}_{XZ_21} = \tilde{\pi}_{XZ_2j} + \sum_{\ell=1}^{j-1}\left[\pi_{XXj-\ell}\tilde{\pi}^{(\ell)}_{XZ_21} + \pi_{XZj-\ell}\pi^{(\ell)}_{ZZ1}Q'_2\right].$$

But, since $Y \not\to_{(h)} X|I_{XZ}$ and $j \leq h-1$, we know from Proposition 5.2 that

$$\tilde{\pi}^{(\ell)}_{XZ_21} = 0\,, \quad \text{for } \ell = 1,2,\ldots,j\,,$$

hence

$$0 = \tilde{\pi}_{XZ_2j} + \sum_{\ell=1}^{j-1}\pi_{XZj-\ell}\pi^{(\ell)}_{ZZ1}Q'_2 = \tilde{\pi}_{XZ_2j} + \sum_{\ell=1}^{j-1}\pi_{XZj-\ell}[Q'_1, Q'_2]\begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}\pi^{(\ell)}_{ZZ1}Q'_2$$

$$= \tilde{\pi}_{XZ_2j} + \sum_{\ell=1}^{j-1}\left[\tilde{\pi}_{XZ_1j-\ell}\tilde{\pi}^{(\ell)}_{Z_1Z_21} + \tilde{\pi}_{XZ_2j-\ell}\tilde{\pi}^{(\ell)}_{Z_2Z_21}\right],$$

which is the derived result.

Q.E.D.

Proof of Proposition 5.4 : When $Y \not\to_{(h)} X|I_{XZ}$ and

$$\sum_{\ell=1}^{j-1}\tilde{\pi}_{XZ_1j-\ell}\tilde{\pi}^{(\ell)}_{Z_1Z_21} = 0\,, \qquad \text{for } j = 2,3,\ldots,h-1\,,$$

we get from Proposition 5.3 that

$$\tilde{\pi}_{XZ_2j} = -\sum_{\ell=1}^{j-1}\tilde{\pi}_{XZ_2j-\ell}\tilde{\pi}^{(\ell)}_{Z_2Z_21}\,, \qquad \text{for } j = 2,3,\ldots,h-1\,.$$

Since $\tilde{\pi}_{XZ_21} = 0$ (by Proposition 5.2), it is then easy to see that

$$\tilde{\pi}_{XZ_2j} = 0\,, \qquad \text{for } j = 2,3,\ldots,h-1\,.$$

The fact that $Y \not\to_\infty X|I_{XZ}$ implies (5.20) or (5.21) is a direct consequence of (5.18).

Q.E.D.

REFERENCES

Ahn, S.K. and G.C. Reinsel (1988), "Nested Reduced-Rank Autoregressive Models for Multiple Time Series", *Journal of the American Statistical Association* 83, 849-856.

Andrews, D.W.K. (1987), "Asymptotic Results for Generalized Wald Tests", *Econometric Theory* 3, 348-350.

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.

Boudjellaba, H., J.-M. Dufour and R. Roy (1994), "Simplified Conditions for Non-Causality Between Vectors in Multivariate ARMA Models", *Journal of Econometrics* (forthcoming).

Bruneau, C. and J.P. Nicolai (1992), "Probabilistic Foundations of Causal Analysis in a Stationary Vectorial Autoregressive Model", *Working Paper, CREST-ENSAE.*

Dufour, J.-M. and E. Renault (1994), "Short Run and Long Run Causality in Time Series: Inference", *Working Paper*, C.R.D.E., Université de Montréal, and GREMAQ, Université de Toulouse I.

Dufour, J.-M. and D. Tessier (1993), "On the Relationship Between Impulse Response Analysis, Innovation Accounting and Granger Causality", *Economics Letters* 42, 327-333.

Dufour, J.-M. and D. Tessier (1994), "Parsimonious Autoregressive Conditions for Noncausality in Multivariate ARMA Models", *Working Paper*, C.R.D.E., Université de Montréal.

Florens, J.-P. and M. Mouchart (1985), "A Linear Theory for Non-Causality", *Econometrica* 53, 157-175.

Geweke, J. (1984), "Inference and Causality in Economic Time Series Models" in: Z. Griliches and M. Intrilligator (ed.), *Handbook of Econometrics II*, Amsterdam, North-Holland, 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, 424-259.

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

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

Lütkepohl, H. (1993), "Testing for Causation Between Two Variables in Higher Dimensional VAR Models", in H. Schneeweiss and K. Zimmermann (eds.), *Studies in Applied Econometrics*, Heidelberg: Physica-Verlag, 1993.

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*, Amsterdam, North-Holland, 701-716.

Pierce, D.A. and L.D. Haugh (1977), "Causality in Temporal Systems : Characterizations and Survey", *Journal of Econometrics* 5, 265-293.

Renault, E. and A. Szafarz (1991), "True Versus Spurious Instantaneous Causality", *Cahier de recherche n° 9103*, Centre d'économie mathématique et d'économétrie, Université Libre de Bruxelles.

Sims, C. (1980), "Macroeconomics and Reality", *Econometrica* 48, 1-48.

Wiener, N. (1956), "The Theory of Prediction", in: E.F. Beckenback (ed.), *Modern Mathematics for Engineers (Series 1)*, Chapter 8, New York, McGraw-Hill.