

Article

# A Pitfall in Using the Characterization of Granger Non-Causality in Vector Autoregressive Models

Umberto Triacca

Department of Computer Engineering, Computer Science and Mathematics, University of L'Aquila,  
Via Vetoio I-67010 Coppito, L'Aquila, Italy; E-Mail: umberto.triacca@ec.univaq.it;  
Tel.: +0039-06-4456100

Academic Editor: Kerry Patterson

Received: 3 February 2015 / Accepted: 1 April 2015 / Published: 9 April 2015

---

**Abstract:** It is well known that in a vector autoregressive (VAR) model Granger non-causality is characterized by a set of restrictions on the VAR coefficients. This characterization has been derived under the assumption of non-singularity of the covariance matrix of the innovations. This note shows that if this assumption is violated, then the characterization of Granger non-causality in a VAR model fails to hold. In these situations Granger non-causality test results must be interpreted with caution.

**Keywords:** covariance matrix; Granger causality; time series

**JEL classifications:** C1; C32

---

## 1. Introduction

A variable  $y$  is said to cause another variable  $x$  in Granger sense [1], if future  $x$ -values can be better predicted using past values of  $x$  and  $y$  rather than using the past of  $x$  alone.

The concept of Granger causality is a fundamental tool for the empirical investigation of dynamic interactions in multivariate time-series. It is well known that in a vector autoregressive (VAR) model Granger non-causality is characterized by a set of restrictions on the VAR coefficients. Since the seminal paper of Sims [2], this characterization forms the basis of various tests for Granger non-causality. Thus, it is important to investigate the conditions under which this characterization holds.

In particular, we show that if the assumption concerning the non-singularity of the covariance matrix of the VAR innovations is violated, then the characterization of Granger non-causality fails to hold.

In literature, the assumption concerning the non-singularity of the covariance matrix of the VAR innovations is considered a modest requirement. However, we sustain that dynamic systems with singular covariance matrix are not infrequent in economics. Thus the VAR practitioners have to be careful about this possibility, since in these situations we cannot conclude that there is Granger causality if the tests reject the null hypothesis of non-causality.

The rest of the paper is organized as follows. Section 2 presents the characterization of Granger non-causality condition within the framework of bivariate autoregressive models. Section 3 presents two illustrative examples. Section 4 provides a theoretical result establishing under which conditions the characterization of Granger non-causality for VAR models fails to hold. Section 5 concludes.

## 2. Granger Non-Causality in VAR Models

In order to simplify matters, we will focus on the bivariate case.

Let  $\{(x_t, y_t)'; t \in \mathbb{Z}\}$  be a purely non-deterministic zero-mean covariance stationary bivariate stochastic process. Suppose that  $\{(x_t, y_t)'; t \in \mathbb{Z}\}$  admits the following autoregressive representation.

$$\begin{bmatrix} \pi_{11}(L) & \pi_{12}(L) \\ \pi_{21}(L) & \pi_{22}(L) \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \epsilon_{xt} \\ \epsilon_{yt} \end{bmatrix} \quad (1)$$

where  $L$  is the lag operator such that  $L^n x_t = x_{t-n}$ ,

$$\begin{aligned} \pi_{ij}(L) &= 1 - \sum_{h=1}^p \pi_{ij}(h) L^h \quad \text{for } i = j \\ \pi_{ij}(L) &= - \sum_{h=1}^p \pi_{ij}(h) L^h \quad \text{for } i \neq j \end{aligned}$$

and  $\{(\epsilon_{xt}, \epsilon_{yt})'; t \in \mathbb{Z}\}$  is a bi-dimensional white noise process with non singular covariance matrix  $\Sigma$ .

For any information set  $I(t)$  available at time  $t$  the optimal (minimum mean square error) predictor of  $x_{t+1}$ , based on the information in  $I(t)$ , is denoted  $P(x_{t+1}|I(t))$ ,  $e(x_{t+1}|I(t)) = x_{t+1} - P(x_{t+1}|I(t))$  is the corresponding prediction error, and  $\sigma^2(x_{t+1}|I(t))$  is the variance of  $e(x_{t+1}|I(t))$ .

**Definition 1.** (Granger non-causality) Consider the information sets  $I_{xy}(t) = \{x_s, y_s; s \leq t\}$  and  $I_x(t) = \{x_s; s \leq t\}$ .  $y$  does not cause  $x$  in Granger sense, with respect to the information set  $I_{xy}(t)$ , if

$$\sigma^2(x_{t+1}|I_{xy}(t)) = \sigma^2(x_{t+1}|I_x(t)) \quad \forall t \in \mathbb{Z}$$

The following theorem proved by [1,3] provides a useful characterization of non-causality in a VAR model.

**Theorem 1.** (Characterization of Granger non-causality for VAR models) Let  $\{(x_t, y_t)'; t \in \mathbb{Z}\}$  be a purely non-deterministic zero-mean covariance stationary bivariate stochastic process. If  $\{(x_t, y_t)'; t \in \mathbb{Z}\}$  admits an autoregressive representation as in Equation (1) with non-singular white noise matrix  $\Sigma$ , then

$$\sigma^2(x_{t+1}|I_{xy}(t)) = \sigma^2(x_{t+1}|I_x(t)) \quad \forall t \in \mathbb{Z}$$

if and only if

$$\pi_{12}(h) = 0 \quad \text{for } h = 1, \dots, p$$

This characterization is very useful since it forms the basis of various test procedures for Granger non-causality. Therefore, it is important to understand the conditions under which this theorem holds. In particular, we consider the role played by the assumption of non-singularity of the covariance matrix  $\Sigma$ .

What does this assumption mean? We know that  $\Sigma$  is non-singular if and only if the rank of  $\Sigma$ ,  $\rho(\Sigma)$ , is full. The rank of  $\Sigma$  indicates an important structural characteristic of the bivariate process  $\{(x_t, y_t)'; t \in \mathbb{Z}\}$ . We can have three cases:

1.  $\rho(\Sigma) = 0$ . In this case  $\Sigma$  is the null matrix and the process is deterministic and may be perfectly predicted from its past.
2.  $\rho(\Sigma) = 1$ . This is a degenerate case in which the bivariate innovation  $\{(\epsilon_{xt}, \epsilon_{yt})'; t \in \mathbb{Z}\}$  is essentially univariate.
3.  $\rho(\Sigma) = 2$ . The full rank case.

We will not consider the case 1 here since our process is supposed to be purely non-deterministic. If  $\rho(\Sigma) = 2$ ,  $\Sigma$  is non-singular, two genuine (linearly independent) shocks perturb the system each period. If  $\rho(\Sigma) = 1$ ,  $\Sigma$  is singular, there is only one genuine shock that perturbs the system. It is important to note that the case 2 is not infrequent in economics. Many dynamic models deliver solutions for the endogenous variables whose covariance matrix is singular because there is a number of endogenous variables larger than shocks.

In general, we have that  $y$  does not Granger cause  $x$  if and only if the past values of  $y$  does not appear in the  $x$  equation of the VAR model. However, in this section, we present an example in which  $y$  does not cause  $x$  and the past values of  $y$  are present in the expression for  $x$ . Consider a bivariate stochastic process  $\{(x_t, y_t)'; t \in \mathbb{Z}\}$  that admits the following representation:

$$\begin{bmatrix} 1 - \alpha L & -\beta L \\ -\gamma L & 1 + \beta L \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \epsilon_{xt} \\ -\epsilon_{xt} \end{bmatrix} \quad (2)$$

where  $\epsilon_{xt} \sim WN(0, \sigma^2)$ .

Premultiplying both sides of Equation (2) by the following adjoint matrix

$$\begin{bmatrix} 1 + \beta L & \beta L \\ \gamma L & 1 - \alpha L \end{bmatrix}$$

we get the “final equations”:

$$[(1 - \alpha L)(1 + \beta L) - \gamma L \beta L] \begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 1 + \beta L & \beta L \\ \gamma L & 1 - \alpha L \end{bmatrix} \begin{bmatrix} \epsilon_{xt} \\ \epsilon_{yt} \end{bmatrix}$$

The implied univariate ARMA model for the subprocess  $\{x_t; t \in \mathbb{Z}\}$  is

$$x_t = (\alpha - \beta)x_{t-1} + (\gamma + \alpha)\beta x_{t-2} + \epsilon_{xt}$$

This is in an autoregression form, and can be used to forecast  $x_{t+1}$  from  $x_{t-j}$ ,  $j \geq 0$ . In particular, we have that

$$P(x_{t+1}|I_t^x) = (\alpha - \beta)x_t + (\gamma + \alpha)\beta x_{t-1}$$

and hence

$$\sigma^2(x_{t+1}|I_t^{xy}) = \sigma^2$$

On the other hand we have that

$$P(x_{t+1}|I_t^{xy}) = \alpha x_t + \beta y_t$$

with

$$\sigma^2(x_{t+1}|I_t^{xy}) = \sigma^2$$

Thus

$$\sigma^2(x_{t+1}|I_t^x) = \sigma^2(x_{t+1}|I_t^{xy})$$

Hence we can conclude that  $y$  does not Granger cause  $x$  but  $\beta$  can be different from zero. In other terms, can happen that  $y$  does not Granger cause  $x$  and the past values of  $y$  can be present in the  $x$  equation of the VAR model.

However, this is not a counter example for the characterization of non-causality. In fact, in the proof of Theorem 1 it is assumed that the covariance matrix  $\Sigma$  is non singular, while in our example

$$\Sigma = \begin{bmatrix} \sigma^2 & -\sigma^2 \\ -\sigma^2 & \sigma^2 \end{bmatrix}$$

is clearly singular.

It is important to remember that the problem of singularity has been investigated by a number of authors. For example, Gonzalo and Lee [4] treat singularity of the error covariance matrix as a major concern. While their analysis concerns the pitfalls in testing for cointegration in systems with singular covariance matrices, the goal of this paper is different. Here, we are interested to clarify the role of the non singularity assumption for the characterization of non-causality in a VAR model.

### 3. The Result

In this section we present the main result of this paper.

**Proposition 1.** *Let  $\{(x_t, y_t)'; t \in \mathbb{Z}\}$  be a purely non-deterministic zero-mean covariance stationary bivariate stochastic process. If  $\{(x_t, y_t)'; t \in \mathbb{Z}\}$  admits an autoregressive representation as in Equation (1) with singular white noise matrix  $\Sigma$  given by*

$$\Sigma = \begin{bmatrix} \sigma^2 & \lambda\sigma^2 \\ \lambda\sigma^2 & \lambda^2\sigma^2 \end{bmatrix}$$

*then  $y$  does not cause  $x$  in Granger sense.*

**Proof.** First we note that the condition

$$\Sigma = \begin{bmatrix} \sigma^2 & \lambda\sigma^2 \\ \lambda\sigma^2 & \lambda^2\sigma^2 \end{bmatrix}$$

implies that  $\epsilon_{yt} = \lambda\epsilon_{xt}$ . Premultiplying both sides of Equation (1) by the following adjoint matrix

$$\begin{bmatrix} \pi_{22}(L) & \pi_{12}(L) \\ \pi_{21}(L) & \pi_{11}(L) \end{bmatrix}$$

we get the “final equations”:

$$[\pi_{11}(L)\pi_{22}(L) - \pi_{12}(L)\pi_{21}(L)] \begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \pi_{22}(L) & \pi_{12}(L) \\ \pi_{21}(L) & \pi_{11}(L) \end{bmatrix} \begin{bmatrix} \epsilon_{xt} \\ \lambda \epsilon_{xt} \end{bmatrix}$$

The implied univariate ARMA(2p, p) models for the subprocesses  $\{x_t; t \in \mathbb{Z}\}$  and  $\{y_t; t \in \mathbb{Z}\}$  are given by

$$[\pi_{11}(L)\pi_{22}(L) - \pi_{12}(L)\pi_{21}(L)] x_t = [\pi_{22}(L) + \lambda\pi_{12}(L)] \epsilon_{xt}$$

and

$$[\pi_{11}(L)\pi_{22}(L) - \pi_{12}(L)\pi_{21}(L)] y_t = [\pi_{21}(L) + \lambda\pi_{11}(L)] \epsilon_{xt}$$

It follows that

$$y_t = \frac{[\pi_{21}(L) + \lambda\pi_{11}(L)]}{[\pi_{22}(L) + \lambda\pi_{12}(L)]} x_t$$

and, thus,  $I_{xy}(t) = I_x(t)$ . We can conclude that

$$\sigma^2(x_{t+1}|I_{xy}(t)) = \sigma^2(x_{t+1}|I_x(t)) \quad \forall t \in \mathbb{Z}$$

that is  $y$  does not cause  $x$  in Granger sense.  $\square$

Proposition 1 tells us that  $y$  does not Granger cause  $x$  but  $\pi_{12}$  can be different from zero. In other terms, it can happen that  $y$  does not Granger cause  $x$  and the past values of  $y$  can be present in the  $x$  equation of the VAR model. The characterization provided by Theorem 1 fails to hold because the assumption concerning the non-singularity of the covariance matrix of the VAR innovations is violated.

We close this section by observing that the primary goal of assuming non-singularity for the covariance matrix of the error term in the VAR model is to avoid the possibility of including variables with redundant information (in terms of prediction) in the model. Formally, we have that if  $\Sigma$  is singular, then  $I_{xy}(t) = I_x(t)$ . The variable  $y$  is totally redundant as it does not provide any additional information useful for predicting  $x$ .

#### 4. Some Economic Examples

To illustrate how theoretical models can result in VAR representations with singular covariance matrix in economic settings we will consider two simple examples.

**Example 1.** We begin with the well-known cobweb market model. Demand for a product depends on price,  $p_t$ . Supply depends linearly on  $m_t$ , the price expected the previous period, when production decisions are made. Both supply and demand also depend on unobserved white noise random shocks. The model can be solved to yield the reduced form of the price level

$$p_t = \alpha m_{t-1} + \eta_t, \quad \alpha < 0$$

where  $\eta_t \sim WN(0, \sigma^2)$ . This equation describes how the endogenous variable actually evolve. Further, we assume that the agents form their expectations following an adaptive scheme, that is

$$m_t = m_{t-1} + \beta(p_t - m_{t-1}), \quad 0 \leq \beta \leq 1$$

Thus we have that the bivariate stochastic process  $\{(p_t, m_t)'; t \in \mathbb{Z}\}$  admits the following VAR representation:

$$\begin{bmatrix} 1 & -\alpha L \\ 0 & 1 - (1 + \alpha\beta - \beta)L \end{bmatrix} \begin{bmatrix} p_t \\ m_t \end{bmatrix} = \begin{bmatrix} \eta_t \\ \beta\eta_t \end{bmatrix}$$

**Example 2.** We consider the following present value model of a stock price:

$$p_t = E_t \left[ \sum_{i=1}^{\infty} b^i d_{t+i} \right], \quad 0 < b < 1$$

Here  $E_t$  denotes an expectation conditional on all information available to the stock market participants at time  $t$  and  $b$  is the discount factor. The present value model relates the asset's price,  $p_t$ , to the sum of its discounted expected future dividends,  $d_{t+i}$ . A similar model is presented in [5] (p. 306). We assume that

$$d_t = e_t + ce_{t-1} \quad e_t \sim i.i.d.N(0, \sigma_e^2)$$

with  $|c| < 1$ . Then

$$E_t(d_{t+i}) = \begin{cases} ce_t & \text{for } i = 1 \\ 0 & \text{for } i > 1 \end{cases}$$

which implies

$$p_t = bce_t$$

and hence

$$\delta e_t = b^{-1}p_t$$

It follows that process  $\{(d_t, p_t)'; t \in \mathbb{Z}\}$  admits the following VAR representation:

$$\begin{bmatrix} 1 & -b^{-1}L \\ 0 & 1 \end{bmatrix} \begin{bmatrix} d_t \\ p_t \end{bmatrix} = \begin{bmatrix} e_t \\ bce_t \end{bmatrix}$$

These examples show that dynamic systems with singular covariances matrix can arise in models where the number of unobserved shocks is less than the number of observed variables. In these cases the characterization of Granger non-causality for VAR models fails to hold.

## 5. Conclusions

In this paper we have investigated a potential problem with application of the Granger-causality tests based on the VAR modeling. In particular, we have showed that if the covariance matrix of the VAR innovations is singular, then the characterization of Granger non-causality for VAR models fails to hold. In literature the assumption concerning the non-singularity of the covariance matrix of the VAR innovations is considered a modest requirement. However, dynamic systems with singular covariance matrices are not uncommon in economics. In these VARs the condition that the past values of  $y$  are not present in the  $x$  equation is not a necessary condition for non-causality. Hence, if we reject the null hypothesis of non-causality expressed in terms of restrictions on the VAR coefficients, we cannot conclude that  $y$  Granger causes  $x$ . The tests are not able to capture the presence of non-causality.

## Acknowledgments

I gratefully acknowledge the comments and suggestions of the referees.

## Conflicts of Interest

The author declares no conflict of interest.

## References

1. Granger, C.W.J. Investigating causal relations by econometric models and cross spectral methods. *Econometrica* **1969**, *37*, 424–438.
2. Sims, C.A. Macroeconomics and reality. *Econometrica* **1980**, *48*, 1–48.
3. Sims, C.A. Money, income, and causality. *Am. Econ. Rev.* **1972**, *62*, 540–552.
4. Gonzalo, J.; Lee, T. Pitfalls in testing for long run relationships. *J. Econom.* **1998**, *86*, 129–154.
5. Hamilton, J.D. *Time Series Analysis*; Princeton University Press: Princeton, NJ, USA, 1994.

© 2015 by the author; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).