

D  
Dipartimento

S  
Scienze

E  
Economiche

# Working Paper

Department  
of Economics

Ca' Foscari University of  
Venice

Monica Billio  
Silvestro Di Sanzo

Granger-causality in  
Markov Switching Models



## Granger-causality in Markov Switching Models

**Monica Billio**  
*University of Venice*

**Silvestro Di Sanzo**  
*Universidad de Alicante*

### Abstract

In this paper we propose a new parametrisation of transition probabilities that allows us to characterize and test Granger-causality in Markov switching models by means of an appropriate specification of the transition matrix. Test for independence are also provided. We illustrate our methodology with an empirical application. In particular, we investigate the causality and interdependence between financial and economic cycles using a bivariate Markov switching model. When applied to U.S. data, we find that financial variables are useful for forecasting the direction of aggregate economic activity, and vice versa.

### Keywords

Granger Causality, Markov Chains, Switching Models

### JEL Codes

C53, C32

### *Address for correspondence:*

Monica Billio  
Department of Economics  
Ca' Foscari University of Venice  
Cannaregio 873, Fondamenta S.Giobbe  
30121 Venezia - Italy  
Phone: (+39) 041 2349170  
Fax: (+39) 041 2349176  
e-mail: billio@unive.it

*This Working Paper (o QD se in inglese) is published under the auspices of the Department of Economics of the Ca' Foscari University of Venice. Opinions expressed herein are those of the authors and not those of the Department. The Working Paper series (o QD se in inglese) is designed to divulge preliminary or incomplete work, circulated to favour discussion and comments. Citation of this paper should consider its provisional character.*

# 1 Introduction

The most widely used concept of causality in time series econometrics is due to Granger (1969). Based on some information set  $\Psi$ , Granger's definition of causality states that a variable  $X$  is causal for a variable  $Y$  if the mean squared error of the 1-step ahead forecast error for  $Y$  is smaller when the history of  $X$  is included in  $\Psi$  than when it is excluded. Consequently, if these forecast error variances are equal, then  $X$  is said to be non-causal for  $Y$ .

In the great majority of practical applications, Granger-causality has been analyzed in the context of linear Vector Autoregressive (VAR) models (see for example Hamilton, 1994). For such models, the necessary and sufficient condition for  $X$  to be non-causal for  $Y$  is that all coefficients on lags of  $X$  are zero in the equation that describes  $Y$ . If the roots to the VAR model are outside the unit circle, then the Wald, LM, and LR statistics have their usual limiting  $\chi^2$ -distribution (see Lutkepohl, 1991), while the case of some unit roots implies that the limiting distribution can be nonstandard (see Sims *et al.*, 1990; Toda and Phillips, 1993).

Recently, some attempts have been made to extend Granger-causality to non-linear systems. Warne (2000) and Psaradakis *et al.* (2005) propose different definitions of causality based on Granger's ideas and provide a set of (economically and statistically meaningful) parametric Granger non-causality restrictions in the context of Markov Switching VAR models. On the other hand, van Dijk *et al.* (2000) consider out-of-sample forecasting-based tests of Granger-causality using a multivariate smooth transition autoregressive model (STAR). Finally, Mosconi and Seri (2005) properly define the meaning of Granger non-causality in the framework of

binary data models. Our work is closely related to Mosconi and Seri (2005) results and propose a different and simpler parametrisation useful for Markov chains.

In contrast to the existing literature, we describe causality relationships by working directly with multiple Markov chains. We propose a novel parametrisation of the transition probabilities that allows us to characterize and test Granger-causality using an appropriate specification of the transition matrix. More precisely, our strategy is based on a particular decomposition of the transition probabilities that allows us to test directly if a Markov chain causes another Markov chain in the Granger sense, that is, if one Markov chain helps to predict another one. The causality tests we propose are based on the transition matrix and not on the parameters of a specific model, thus they can be applied to any class of Markov Switching models.

Causality analysis based on Markov chains is attractive for a number of reasons. First, in a multi-country/multi-sector framework, this methodology allows us to explain the interactions between macro-areas. In fact, this approach is very useful to study the relationships between phases in different countries or sectors, and allows us to determine the causality of these relationships improving our comprehension of the connections among phases, and then produce a better description of how these phases evolve. Second, it is certainly useful in order to describe the relationships between leading and lagging countries or to describe the relation between business surveys and macroeconomic variables.

We illustrate our methodology with an empirical application. We investigate the causality and interdependence between financial and economic cycles using the Markov switching model proposed by Hamilton and Lin (1996, hereafter HL). HL find that economic variables may be useful in forecasting stock price volatility but

no rigorous test based on Granger's ideas is provided. We employ our methodology to determine, in a formal way, whether business cycles have predicting power for financial variables and/or vice versa.

The paper is organized as follows. In the next section, we discuss the basic assumptions and the definitions of Granger-causality we adopt in this paper. In section 3, we show how Granger-causality can be tested in Markov switching models. An application of these tests is presented in section 4. Section 5 concludes.

## 2 Basic Definitions

The aim of this section is to provide a mathematically rigorous definition of non-causality based on predictability. To do that in a general framework, we have to define the specification of the stochastic process we want to predict, the available information set, and the reduced information set. In the relevant theoretical literature, several generalization of non-causality are available. In this paper we follow Mosconi and Seri (2005) and adopt the concept of discrete-time one-step ahead strong non-causality proposed by Florens and Fougère (1996). Following their definition, one-step ahead is referred to the prediction horizon. It is opposed to global, which is valid for any horizon. On the other hand, strong, as opposed to weak, means that we focus on predicting the whole distribution, rather than only the mean. It is pointed out that the original definition proposed by Granger (1969) is specified in terms of the mean.

Let  $\{W_t = (S_t, Z_t), t \in I \subseteq \mathbb{N} = \{0, 1, \dots\}\}$ , or  $\{W_t\}$  for short<sup>1</sup>, be a discrete-time stochastic process on the probability space  $(\Omega, \Upsilon, P)$ . The usual statistical problem of non-causality is to test if  $P$  satisfies the non-causality conditions. The filtration  $\{F_t, t \in I\} = \{F_t\}$  provides the information available at time  $t$ . To make the analysis simpler, we assume  $\{F_t\}$  to be the canonical filtration<sup>2</sup> associated to a large stochastic process  $\{(W_t, D_t)\} = \{(S_t, Z_t, D_t)\}$ , where  $\{S_t\}$ ,  $\{Z_t\}$  and  $\{D_t\}$  may be either scalar or vector processes. Finally, we introduce the reduced information set, which is represented by the canonical filtrations  $\{G_t^1\} = \{\sigma\{S_\tau, D_\tau\}, 0 \leq \tau \leq t\}$  and  $\{G_t^2\} = \{\sigma\{Z_\tau, D_\tau\}, 0 \leq \tau \leq t\}$ . Then, let  $\{\Gamma_t^1\}$ ,  $\{\Gamma_t^2\}$  and  $\{\Gamma_t\}$  be the canonical filtrations associated with the processes  $\{S_t\}$ ,  $\{Z_t\}$  and  $\{W_t\}$ , respectively. Note that  $\Gamma_t^1 \subseteq G_t^1 \subseteq F_t, \forall t \in I$ . Similarly,  $\Gamma_t^2 \subseteq G_t^2 \subseteq F_t, \forall t \in I$ .

Through the paper the following set of definitions, which are fixed in terms of conditional independence of sub- $\sigma$ -fields of  $\Upsilon$ , will be adopted (see Florens and Mouchart, 1982, for further details):

**Definition 1.** *Strong one-step ahead Granger non-causality:*  $\{Z_t\}$

does not strongly cause  $\{S_t\}$  one-step ahead given  $\{G_{t-1}^1\}$  if:

$$\Gamma_t^1 \perp\!\!\!\perp \Gamma_{t-1}^2 | G_{t-1}^1 \quad \forall t \in I.$$

Similarly,

---

<sup>1</sup>Through the paper the following notation will be taken for discrete-value variables, capital letters denote the random variable, while small letters denote a particular realization. Moreover,  $\{Q_t\}$  denotes a stochastic process, while  $Q_t$  represents the value of the process at time  $t$ .

<sup>2</sup>Recall that a canonical filtration associated with a general process  $\{Q_t\}$  defined on  $(\Omega, \Upsilon, P)$  is a family  $\{F_t\}$  of *sub- $\sigma$*  fields of  $\Upsilon$ , where  $F_t = \sigma\{Q_s, 0 \leq s \leq t\}$ . In a more intuitive way,  $F_t$  represents the knowledge of the history of  $\{Q_t\}$  up to time  $t$ .

*Strong one-step ahead Granger non-causality:*  $\{S_t\}$  does not strongly cause  $\{Z_t\}$  one-step ahead given  $\{G_{t-1}^2\}$  if:

$$\Gamma_t^2 \perp\!\!\!\perp \Gamma_{t-1}^1 | G_{t-1}^2 \quad \forall t \in I.$$

**Definition 2.** *Strong simultaneous independence:*  $\{S_t\}$  and  $\{Z_t\}$  are strongly simultaneous independent given  $\{F_{t-1}^1\}$  if:

$$\Gamma_t^1 \perp\!\!\!\perp \Gamma_t^2 | F_{t-1}^1 \quad \forall t \in I.$$

As in Mosconi and Seri (2005) we use the term simultaneous in the latter definition: it has exactly the same meaning as instantaneous in Geweke (1984) and Granger (1988) but in discrete time it is preferable.

We next show how we can apply definitions 1 and 2 to a specific stochastic process and information set. To be concrete, we assume that  $W_t = (S_t, Z_t)$  is a binary random variable that takes values on  $\{0, 1\}$ , and restrict the information set to the canonical filtration associated with  $\{W_t\}$ . We assume the stochastic process  $\{W_t\}$  to be a first-order Markov process (or Markov chain) with transition probabilities:

$$P(w_t | w_{t-1}, \dots, w_0) = P(w_t | w_{t-1}) = P(s_t, z_t | s_{t-1}, z_{t-1}). \quad (1)$$

Then, all the information from the history of the process which is relevant for the transition probabilities in  $t$  is represented by the state of the process in  $(t-1)$ . Under the additional assumption that transition probabilities do not vary over time, the process is defined as a Markov chain with stationary transition probabilities. For particular realizations of  $S_t$  and  $Z_t$ , this process can be represented with the notation  $P_{s_t z_t | s_{t-1} z_{t-1}}$ . For example,  $P_{10|01}$  would correspond to

$P(S_t = 1, Z_t = 0 | S_{t-1} = 0, Z_{t-1} = 1)$ . These transition probabilities are summarized in the transition matrix  $\Sigma$ . Finally, notice that we can decompose the transition probabilities as follows:

$$P(w_t | w_{t-1}) = P(s_t, z_t | s_{t-1}, z_{t-1}) = P(s_t | z_t, s_{t-1}, z_{t-1}) P(z_t | s_{t-1}, z_{t-1}). \quad (2)$$

We can now define Granger non-causality for a Markov chain:

**Definition 3.** *Strong one-step ahead non-causality for a Markov chain with stationary transition probabilities:*  $S_{t-1}$  does not strongly cause  $Z_t$  one step ahead given  $Z_{t-1}$  if:

$$P(z_t | z_{t-1}, s_{t-1}) = P(z_t | z_{t-1}) \quad \forall t.$$

Similarly,

*Strong one-step ahead non-causality for a Markov chain with stationary transition probabilities :*  $Z_{t-1}$  does not strongly cause  $S_t$  one step ahead given  $S_{t-1}$  if:

$$P(s_t | s_{t-1}, z_{t-1}) = P(s_t | s_{t-1}) \quad \forall t.$$

**Definition 4.** *Strong simultaneous independence for a Markov chain with stationary transition probabilities:*  $\{S_t\}$  and  $\{Z_t\}$  are strongly simultaneous independent given  $\{W_{t-1}^1\}$  if:

$$P(s_t, z_t | s_{t-1}, z_{t-1}) = P(s_t | s_{t-1}, z_{t-1}) P(z_t | s_{t-1}, z_{t-1}) \quad \forall t;$$

or equivalently:

$$\begin{aligned} P(s_t | z_t, s_{t-1}, z_{t-1}) &= P(s_t | s_{t-1}, z_{t-1}) \\ P(z_t | s_t, s_{t-1}, z_{t-1}) &= P(z_t | s_{t-1}, z_{t-1}) \quad \forall t. \end{aligned}$$



It is easy to show the equivalence between definitions 1 and 3. In fact, note that  $P(z_t, s_{t-1}/z_{t-1}) = P(z_t/s_{t-1}, z_{t-1})P(s_{t-1}/z_{t-1})$ . Under the hypotheses described above, definition 1 implies that  $P(z_t, s_{t-1}/z_{t-1}) = P(z_t/z_{t-1})P(s_{t-1}/z_{t-1})$ , which in turn implies definition 3. Similarly, it is possible to proceed for definitions 2 and 4.

The non-causality definition involves the marginal distributions of  $S_t$  and  $Z_t$  conditional on  $W_{t-1}$ . Then, to study causality it is necessary to consider the transition probabilities of the Markov process whereas testing for simultaneous independence requires the joint distribution to be fully specified, and compared to the product of the marginal distributions. Since  $W_{t-1}$  as well as  $W_t$  can belong to a finite set of four states, the most general model representing  $P(w_t|w_{t-1})$  involves 16 parameters, corresponding to the transition probabilities from each of the states in  $t-1$  to each of the states in  $t$ . More precisely, since the sum of the transition probabilities for transitions starting from each of the states is equal to 1, only 12 parameters are enough to describe the conditional distribution completely.

Let us now show how we can construct a transition matrix starting from the non-causality definition. Note that the state of the system in  $(t-1)$  can be defined by the four possibilities of the joint Markov chain  $W_{t-1}$  :

$$X_t = (1, s_{t-1}, z_{t-1}, s_{t-1}z_{t-1})' = (1, s_{t-1}) \otimes (1, z_{t-1})'$$

where  $\otimes$  denotes the Kronecker product. In fact,  $X_t$  is an invertible linear transformation of:

$$X_t^* = [(1 - s_{t-1})(1 - z_{t-1}), s_{t-1}(1 - z_{t-1}), (1 - s_{t-1})z_{t-1}, s_{t-1}z_{t-1}].$$

where  $X_t^*$  is characterized by four mutually exclusive dummies representing the

four states of the process in  $(t - 1)$ . As in Mosconi and Seri (2005), we employ  $X_t$  instead of  $X_t^*$  (or directly  $s_{t-1}$  and  $z_{t-1}$ ) to describe the state in  $(t - 1)$  since by means of this specification non-causality restrictions are more easily written and interpreted. Given the decomposition (2), we can consider the logistic function to represent the two probabilities and then it is simply to verify that we can represent the joint probability of  $S_t$  and  $Z_t$  as follows:

$$P(s_t, z_t | s_{t-1}, z_{t-1}) = P(s_t / z_t, s_{t-1}, z_{t-1}) P(z_t / s_{t-1}, z_{t-1}) \quad (3)$$

$$= \frac{\exp(\alpha \odot Y_t)}{1 + \exp(\alpha \odot Y_t)} * \frac{\exp(\beta \odot X_t)}{1 + \exp(\beta \odot X_t)}, \quad (4)$$

where

$$\begin{aligned} Y_t &= (1, z_t)' \otimes (1, s_{t-1})' \otimes (1, z_{t-1})' \\ &= (1, s_{t-1}, z_{t-1}, s_{t-1}z_{t-1}, z_t, z_t z_{t-1}, z_t s_{t-1}, z_t z_{t-1} s_{t-1})', \end{aligned}$$

the vectors  $\alpha$  and  $\beta$  have dimensions  $(8 \times 1)$  and  $(4 \times 1)$ , respectively<sup>3</sup>, and  $\odot$  denotes element-by-element multiplication, while  $X_t$  has already been defined. Note that  $\alpha$  and  $\beta$  represent 8 and 4 parameters, respectively. Then, we simply have an alternative parameterization of the transition matrix.

Such parameterization is very useful since it allows us to impose the non-causality restrictions in a very simple way by easily restricting the transition matrix to be described by a number of parameters comprised between 4 and 12.

---

<sup>3</sup>In the following we denote with  $\alpha_j$  and  $\beta_j$  the  $j$ 'th element of the vectors  $\alpha$  and  $\beta$ , respectively.

### 3 Granger non-causality tests

Given the parametrization (4) the conditions for strong one step ahead non-causality and strong simultaneous independence are easily stated as restrictions on the parameter space.

To impose the Granger non-causality (as in definition 3) it is necessary that the dependence on  $s_{t-1}$  disappears in the second term of the decomposition (2), thus it simply requires that the parameters of the terms of  $X_t$  depending on  $s_{t-1}$  are equal to zero:

$$H_{1\not\Rightarrow 2} \quad (S_t \not\Rightarrow Z_t) : \quad \beta_2 = \beta_4 = 0.$$

Under  $H_{1\not\Rightarrow 2}$ ,  $S_{t-1}$  does not strongly cause one-step ahead  $Z_t$  given  $Z_{t-1}$ . The terms  $s_{t-1}$  and  $s_{t-1}z_{t-1}$  are excluded from  $X_t$ , thus  $p(z_t/s_{t-1}, z_{t-1}) = p(z_t/z_{t-1})$ .

On the other hand, if we want to test if  $Z_{t-1}$  does not strongly cause  $S_t$ , we can differently decompose the joint probability and define it as follows:

$$P(s_t, z_t | s_{t-1}, z_{t-1}) = P(z_t/s_t, s_{t-1}, z_{t-1})P(s_t/s_{t-1}, z_{t-1}) \quad (5)$$

$$= \frac{\exp(\alpha^* \odot Y_t^*)}{1 + \exp(\alpha^* \odot Y_t^*)} * \frac{\exp(\beta^* \odot X_t)}{1 + \exp(\beta^* \odot X_t)}, \quad (6)$$

where

$$\begin{aligned} Y_t^* &= (1, s_t)' \otimes (1, s_{t-1})' \otimes (1, z_{t-1})' \\ &= (1, s_{t-1}, z_{t-1}, s_{t-1}z_{t-1}, s_t, s_t z_{t-1}, s_t s_{t-1}, s_t z_{t-1} s_{t-1})'. \end{aligned}$$

In this case, the test becomes:

$$H_{1\not\Leftarrow 2} \quad (S_t \not\Leftarrow Z_t) : \quad \beta_3^* = \beta_4^* = 0.$$

Under  $H_{1\not\Leftarrow 2}$ ,  $Z_{t-1}$  does not strongly cause one-step ahead  $S_t$  given  $S_{t-1}$ . In fact,  $z_{t-1}$  and  $s_{t-1}z_{t-1}$  are excluded from  $X_t$ , so that  $P(s_t/s_{t-1}, z_{t-1}) = P(s_t/s_{t-1})$ .

It is important to note that the parameters  $(\alpha, \beta)$  and  $(\alpha^*, \beta^*)$  are bijective transformations of the transition probabilities. That is,

$$P_{s_t z_t | s_{t-1} z_{t-1}} = f(Y_t, X_t, \alpha, \beta) = f^*(Y_t^*, X_t, \alpha^*, \beta^*). \quad (7)$$

Hence, we can obtain the parameters  $(\alpha^*, \beta^*)$  as:

$$(\alpha^*, \beta^*) = f^{*-1}(P_{s_t z_t | s_{t-1} z_{t-1}}) = f^{*-1}(f(Y_t, X_t, \alpha, \beta)).$$

Consequently we can obtain the estimates of  $(\alpha^*, \beta^*)$  from the estimates of  $(\alpha, \beta)$ . Finally, to perform the test we need their variance-covariance matrix: given the variance-covariance matrix of  $(\hat{\alpha}, \hat{\beta})$  we can compute it by the usual sandwich formula.

In the same way it is possible to test for the strong simultaneous independence:

$$H_{1 \not\leftrightarrow 2} \quad (S_t \not\leftrightarrow Z_t) : \quad \alpha_5 = \alpha_6 = \alpha_7 = \alpha_8 = 0$$

or equivalently:

$$H_{1 \not\leftrightarrow 2} \quad (S_t \not\leftrightarrow Z_t) : \quad \alpha_5^* = \alpha_6^* = \alpha_7^* = \alpha_8^* = 0.$$

In the present framework it is also possible to test for the independence of the Markov Chains. When independence holds, the transition matrix of the joint Markov process is given by the Kronecker product of the transition matrix of each specific chain. In fact, we can write the transition probabilities in equation (1) as follows:

$$P(s_t, z_t | s_{t-1}, z_{t-1}) = P(s_t | s_{t-1})P(z_t | z_{t-1}),$$

and the transition matrix  $\Sigma$  will be:

$$\Sigma = P_s \otimes P_z,$$

where  $P_i$ ,  $i = s, z$  is the  $(2 \times 2)$  transition matrix of the specific Markov chain.

The test for independence is thus given by:

$$H_{1\perp\perp 2} (S_t \perp\!\!\!\perp Z_t) : \beta_2 = \beta_4 = 0; \alpha_3 = \alpha_4 = \alpha_5 = \alpha_6 = \alpha_7 = \alpha_8 = 0$$

or equivalently:

$$H_{1\perp\perp 2} (S_t \perp\!\!\!\perp Z_t) : \beta_3^* = \beta_4^* = 0; \alpha_2^* = \alpha_4^* = \alpha_5^* = \alpha_6^* = \alpha_7^* = \alpha_8^* = 0.$$

## 4 An application to the relationship between financial and business cycles

To illustrate our methodology, we consider the model proposed by Hamilton and Lin (1996) to study the relationship between financial and business cycles. We employ the same data set as HL to facilitate the comparison between the results. Data is taken from Citibase.

HL propose the following specification for the business cycle:

$$y_t - \mu_{s_t} = \phi(y_{t-1} - \mu_{s_{t-1}}) + \epsilon_t \quad (8)$$

where  $y_t$  is defined as 100 times the monthly change in the natural logarithm of the Federal Reserve Board's index of industrial production for 1965:1 to 1993:6. Here  $\epsilon_t$  is supposed to be i.i.d.  $N(0, \sigma^2)$ , while  $S_t$  is an unobserved latent variable that takes values in the set  $\{0, 1\}$ , and represents the state of the business cycle. It is assumed to follow a first-order Markov process with transition probabilities given by:

$$P(S_t = 0 | S_{t-1} = 0) = p_{00}$$

$$P(S_t = 1 | S_{t-1} = 1) = p_{11}$$

The model for the stock return takes the form:

$$r_t = \delta_0 + \delta_1 r_{t-1} + e_t \quad (9)$$

$$e_t = \sqrt{g_{z_t}} u_t \quad (10)$$

$$u_t = \sqrt{h_t} w_t \quad (11)$$

$$h_t = \gamma_0 + \gamma_1 u_{t-1}^2 + \gamma_2 u_{t-1}^2 I_{t-1} \quad (12)$$

where  $r_t$  is 100 times the change in natural logarithm of the S&P 500 stock index plus the dividend yield on the S&P 500 minus the yield on 3-month Treasury bills, with both yields at monthly rates. The random variable  $w_t$  is assumed to be i.i.d.  $N(0, 1)$ , and  $Z_t$  is an unobserved latent variable that reflects the volatility phase of the stock market. As before, this variable follows a two-state first-order Markov process.

If the parameter  $g_{z_t}$  in equation (3) does not switch between regimes it simply equals unity for all  $t$ . If this is the case, equations (2)-(5) describe stock returns with an autoregression whose residual  $e_t$  follows a 1st-order ARCH-L process. The "L" stands for the leverage effect, which means that stock price increases and decreases can have asymmetric effects on subsequent volatility (see Nelson, 1991). The dummy variable  $I_{t-1}$  takes value 1 if  $e_{t-1}$  is negative and zero otherwise. This means that if parameter  $\delta_2$  is different from zero, a stock price decrease has a different effect on subsequent volatility than would a stock price increase of the same magnitude. HL normalize  $g_0 = 1$ , thus  $g_1$  can be interpreted as the ratio of the average variance of stock returns when  $z_t = 1$  compared to that observed when  $z_t = 0$ .

In their work, HL show that economic variables may be useful in forecasting stock price volatility, but no rigorous test based on Granger's ideas is provided.

We next employ the Granger-causality tests described in the previous section for investigating, in a formal way, whether business cycle has a predicting power for financial variables and/or vice versa.

## 4.1 Estimation and testing

The model is estimated by Maximum Likelihood using the Hamilton's (Hamilton, 1989). It is well-known that maximum likelihood estimation of regime-switching models is plagued for complicated likelihood functions with numerous local maxima. To provide some reassurance of the robustness of our result, we estimate the model with several sets of starting values. Maximum likelihood estimates for the parameters and transition probabilities are reported in Tables 1 and 2, respectively. As in HL, we find that industrial production tends to fall by 2/3 of a percent per month as long as the economy remains in regime 1 ( $\hat{\mu}_1 = -0.631$ ), and that the unforecastable component of stock returns (the residual  $e_t$  in equation (9)) has a variance that is over nine times as large in regime 0 as it is in regime 1 ( $\hat{g}_1 = 0.112$ ).

Despite of no imposing constraints when estimating  $\Sigma$ , some of the ML estimates of the transition probabilities fall in the boundary of the allowable parameter space  $[0, 1]$ , which violates the standard regularity conditions that motivate the usual formula for asymptotic standard errors. In such cases, a value of zero is imposed a priori to calculate standard errors with respect to the remaining parameters.

The non-causality relationships between  $S_t$  and  $Z_t$  are tested employing Wald tests. Results are depicted in Table 3:

- 1) The hypothesis  $H_{1 \nrightarrow 2}$ , related to the non-causality of  $S_t$  towards  $Z_t$ , is

strongly rejected. Therefore, macroeconomic variables, such as the index of industrial production, have a predicting power for financial variables. This result supports the empirical evidence provided by Chen (1991) and Billio and Pelizzon (2003) among others, that macroeconomic variables are key determinants of stock returns.

2) The hypothesis  $H_{1 \nrightarrow 2}$ , related to the non-causality of  $Z_t$  towards  $S_t$ , is also rejected. This means that stock volatility is useful for forecasting the direction of aggregate economic activity. Our finding is in accordance with the empirical works of Perez-Quiros and Timmermann (2001), Chauvet (1999) and Hamilton and Lin (1996) among others, which show that financial variables lead the business cycle and seem to be generated from expectations about changes in future economic activity.

3) The hypothesis  $H_{1 \perp 2}$  and  $H_{1 \nrightarrow 2}$ , concerning the independence between  $Z_t$  and  $S_t$ , are rejected. This result is in contrast with general attitude in empirical works studying relationships between financial and economic cycles, to impose a priori the independence of the Markov chains (see Chauvet, 1999, among others).

## 5 Conclusions

In this paper we propose a new technique for characterizing and testing Granger-causality, which is well-equipped to handle those models where the change in regime evolves according to multiple Markov chains. While in the existing literature causality tests are based on testing restrictions involving the coefficients of the Markov switching VAR model, we propose a method for analyzing causal links that specifically takes into account the Markov chains of the model. Dif-



ferently from Mosconi and Seri (2005), we use an appropriate specification of the transition matrix.

We apply our methodology to check for the causal relationships between financial and business cycles in the United States. We use the bivariate model with Markov switching proposed by Hamilton and Lin (1996). The causality tests we perform suggest that financial variables are useful for forecasting the direction of aggregate economic activity, and vice versa.

## References

Billio M. and L. Pelizzon (2003), *Volatility and shocks spillover before and after EMU in Europe stock markets*, Journal of Multinational Financial Management, 13, 323-340.

Chauvet M.(1999), *Stock Market Fluctuations and the Business cycle*, Journal of Economic and Social Measurement, 25, 235-57.

Chen N. (1991), *Financial Economic Opportunities and the Macroeconomy*, Journal of Finance, 46, 529-554.

Florens J.P. and Fougère D. (1996), *Non-Causality in Continuous Time*, Econometrica, 1195-1212.

Florens J. P. and Mouchart M. (1982), *A Note on Non-Causality*, Econometrica, 583-91.

Granger C.W.J. (1969), *Investigating Causal Relations by Econometric Models and Cross-Spectral Methods*, Econometrica, 357-84.

Hamilton J.D. and Lin J. (1996), *Stock Market Volatility and Business Cycle*, Journal of Applied Econometrics, 573-93.

Hamilton J.D. (1994), *Time Series Analysis*, Princeton, NJ: Princeton University Press.

Hamilton J.D. (1989), *A new approach to the economic analysis of nonstationary time series and the business cycle*, *Econometrica*, 357-84.

Kim C.J. and Nelson C.R. (1999), *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*, Cambridge, Mass.: MIT Press.

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

Mosconi R. and Seri R. (2005), *Non-Causality in Bivariate Binary Time Series*, forthcoming in *Journal of Econometrics*.

Nelson D. (1991), *Conditional heteroskedasticity in asset returns: a new approach*, *Econometrica*, 347-70.

Perez-Quiros, G. and Timmermann, A. (2001), *Business cycle asymmetries in stock returns : evidence from higher order moments and conditional densities*, Working paper series 58, European Central Bank.

Psaradakis Z., Morten O.R. and Sola M. (2005), *Markov-Switching causality and the Money-Output Relationship*, *Journal of Applied Econometrics*, 665-83.

Sims C.A., Stock J.H. and Watson M.W. (1990), *Inference in Linear Time Series Models with Some Unit Roots*, *Econometrica*, 113-44.

Toda H.Y. and Phillips P.C.B. (1993), *Vector Autoregression and Causality*, *Econometrica*, 1367-93.

van Dijk D., Teräsvirta T. and Franses P.H. (2000), *Smooth Transition Autoregressive Models - A Survey of Recent Developments*, Working Paper Series in Economics and Finance 380, Stockholm School of Economics.

Warne A. (2000), *Causality and Regime Inference in a Markov Switching VAR*,  
 Research Department, Sveriges Riksbank, Stockholm.

## Tables

TABLE 1

**Maximum likelihood estimates (standard errors in parentheses)**

$\hat{\mu}_0$	0.293	(0.043)	$\hat{\phi}$	0.255	(0.063)	$\hat{\delta}_1$	0.186	(0.055)
$\hat{\mu}_1$	-0.631	(0.126)	$\hat{\gamma}_2$	0.143	(0.013)	$\hat{\gamma}_0$	13.30	(2.8)
$\hat{g}_1$	0.112	(0.033)	$\hat{\delta}_0$	0.423	(0.125)	$\hat{\sigma}^2$	0.450	(0.116)
$\hat{\gamma}_1$	0.123	(0.036)						

TABLE 2

**Transition Matrix (standard errors in parentheses)**

$\hat{\Sigma} =$	0.821	0.000	1.000	0.000
	(0.046)			
	0.179	0.042	0.000	0.156
	(0.038)	(0.011)		(0.190)
	0.000	0.000	0.000	0.011
				(0.135)
	0.000	0.958	0.000	0.833
		(0.128)		(0.232)

TABLE 3

**Causality and Independence Tests**

Hypothesis	Wald test	p-value
$H_{1 \rightarrow 2}$	47.11	0.00
$H_{1 \leftarrow 2}$	57.23	0.00
$H_{1 \perp\!\!\!\perp 2}$	32.29	0.00
$H_{1 \leftrightarrow 2}$	60.33	0.00