

# Granger causal analysis of VARMA-GARCH models <sup>☆</sup>

Tomasz Woźniak<sup>a</sup>

<sup>a</sup>*Department of Economics, European University Institute, Via della Piazzuola 43, 50133 Florence, Italy*

---

## Abstract

Recent economic developments have shown the importance of spillover and contagion effects in financial markets. Such effects are not limited to relations between the levels of financial variables but also impact on the volatility. I investigate Granger causality in conditional mean and conditional variances of time series. For this purpose a VARMA-GARCH process is used. I derive parametric restrictions for the hypothesis of noncausality in conditional variances between two groups of variables, when there are other variables in the system as well. These novel conditions are convenient for the analysis of potentially big systems of economic variables. Such systems should be considered in order to avoid the problem of omitted variable bias. Further, I apply a Bayesian procedure in order to test the restrictions. The test avoids the singularity problem that may appear in classical Wald tests. Also, it relaxes the assumption of existence of higher order moments of the residuals.

*Keywords:* Granger causality, VARMA-GARCH processes, Bayesian testing

*JEL classification:* C11, C12, C32, C53

---

## 1. Introduction

The well known concept of Granger causality (see Granger, 1969, Sims, 1972) describes relations between time series in the forecasting context. We say that one variable does not Granger cause the other if adding the former to the information set with which we forecast the latter does not improve this forecast. In this study we adapt the Granger noncausality concept to conditional variances of the time series (see also Comte & Lieberman, 2000, Robins, Granger & Engle, 1986). If one variable does not Granger cause in variance the other, then information about variability of the former is dispensable for forecasting of the

---

<sup>☆</sup>This version: January 2011

I want to thank Helmut Lutkepohl, Jacek Osiewalski, Timo Teräsvirta, Massimiliano Marcellino, Giampiero Gallo, Russell Cooper and Jérôme Adda for useful comments and remarks on the study. I am grateful to participants of the seminar at Università di Pisa. Also, I wanted to thank Michał Markun and Norbert Metiu and my colleagues from EUI Time Series Econometrics Working Group: Pierre Guérin, Matthieu Droumaguet and Stelios Bekiros for multiple discussions and suggestions.

*Email address:* tomasz.wozniak@eui.eu (Tomasz Woźniak)

*URL:* <http://bit.ly/TWecon> (Tomasz Woźniak)

later. Moreover, we investigate Granger causality in conditional mean and conditional variances of time series taken jointly.

The necessity of the joint analysis is justified with twofold reason. Firstly, as Karolyi (1995) argued, in order to have a good picture of transmissions in mean, transmissions in volatility need to be taken into account. Secondly, transmissions in volatility may be affected by transmissions in mean which have not been filtered out before, a point made by Hong (2001). The conclusion is that combined modeling of conditional mean and conditional variance processes increases reliability of the inference about the transmissions and makes it more powerful. The exposition of the phenomenon in this paper is entirely done with vector autoregressive moving average (VARMA) conditional mean process with generalized autoregressive conditional heteroskedasticity (GARCH) process for conditional variances and constant conditional correlations (CCC).

Why is information about causal relations between time series important? First of all, it gives understanding of the deep structure of the financial markets. More specifically, we learn about integration of the financial markets (assets) not only in returns, but also in risk defined as time-varying volatility. Next, modeling transmissions in volatility has significant impact on volatility forecasting. Thus, if there are Granger causal relations in conditional variances, then such modeling is important in all the applications based on volatility forecasting such as portfolio selection, Value at Risk estimation and option pricing.

Granger causality in conditional variances of exchange rates is consistent with some economic theories. Taylor (1995) shows that it is consistent with failures of exchange rates market efficiency. The arrival of news, in clusters and potentially with a lag, modeled with GARCH models explains inefficiency of the market. It is also in line with market dynamics which exhibits volatility persistence due to private information or heterogeneous beliefs (see Hong, 2001, and references therein). Finally, a *meteor showers* hypothesis for intra-daily exchange rates returns which reflects cooperative or competitive monetary policies (see Engle, Ito & Lin, 1990) can be presented as Granger causality in variance.

The term *transmissions* usually represents an intuitive interpretation of parameters reflecting the impact of one variable on the other in the dynamic systems. Karolyi (1995) and Lin, Engle & Ito (1994) used that term to describe international transmissions between stock returns and their volatilities. Further, Nakatani & Teräsvirta (2009) and Koutmos & Booth (1995) used it to describe the interactions between volatilities in multivariate GARCH models. An other term, *volatility spillovers*, was used in a similar context (see e.g. Conrad & Karanasos, 2009) as well as in many different ones. However, parameters referred to in such a way do not decide on Granger causality or noncausality themselves. In this study we present parameter conditions for precisely defined Granger noncausality in variances. In detail, we refer to the linear Granger noncausality concept of Florens & Mouchart (1985) that defined the noncausality relation in terms of orthogonality in the Hilbert space of square integrable variables.

The contribution of the study is twofold. Firstly, I derive conditions for Granger noncausality in variance that are applicable when the system of time series consists of a potentially large number of variables. The novelty of these conditions is that the

noncausality in variance between two groups of variables is analyzed when there are other variables in the system as well. So far, such conditions were derived when all the variables in the system were divided in two groups (e.g. Comte & Lieberman, 2000, Hafner & Herwartz, 2008, Woźniak, 2010). The introduced conditions reduce the dimensionality of the problem on one hand. And on the other, they allow to form and test some hypotheses that could not be tested in previous settings.

Secondly, I propose a Bayesian procedure for joint testing of the conditions for Granger noncausality in conditional mean and noncausality in conditional variance processes. It is easily applicable and solves some drawbacks of classical testing. In comparison to the Wald test of Boudjellaba, Dufour & Roy (1992) adapted to testing noncausality relations in GARCH models, the Bayesian test does not have the problem of singularities. In a classical test the singularities appear due to the construction of the asymptotic covariance matrix of the nonlinear parametric restrictions. In Bayesian analysis on the contrary, the posterior distribution of the restrictions is available, and thus, a well defined covariance matrix as well. Additionally, in this study the existence of only fourth order moments of time series is assumed, which is an improvement in comparison to assumptions of available classical tests.

The reminder of this paper is organized as follows: the notation for the whole article and parameter restrictions for Granger noncausality in VARMA models are presented in Section 2. A GARCH model used in the analysis is set in Section 3. Further in this Section, I present the main theoretical findings of the paper deriving the conditions for Granger noncausality in conditional variance process. In Section 4, I start with a discussion of classical testing for noncausality in VARMA-GARCH models, and then propose Bayesian testing with appealing properties. Section 5 presents the empirical illustration on the example of daily exchange rates of Swiss franc, British pound and U.S. dollar all denominated in Euro, and Section 6 concludes. All the proofs of the theorems as well as additional necessary theoretical findings are given in Appendix A. Tables and graphs with results of the empirical illustration are reported in Appendix B and Appendix C.

## 2. Granger noncausality in VARMA models

First, we set the notation following Boudjellaba, Dufour & Roy (1994). Let  $\{y_t : t \in \mathbb{Z}\}$  be a  $N \times 1$  multivariate square integrable stochastic process on the integers  $\mathbb{Z}$ . Write

$$y_t = (y'_{1t}, y'_{2t}, y'_{3t})', \quad (1)$$

where  $y_{it}$  is a  $N_i \times 1$  vector such that  $y_{1t} = (\mathbf{y}_{1t}, \dots, \mathbf{y}_{N_1,t})'$ ,  $y_{2t} = (\mathbf{y}_{N_1+1,t}, \dots, \mathbf{y}_{N_1+N_2,t})'$  and  $y_{3t} = (\mathbf{y}_{N_1+N_2+1,t}, \dots, \mathbf{y}_{N_1+N_2+N_3,t})'$  ( $N_1, N_2 \geq 1, N_3 \geq 0$  and  $N_1 + N_2 + N_3 = N$ ). Variables of interest are contained in vectors  $y_1$  and  $y_2$  between which we want to study causal relationships. Vector  $y_3$  (that for  $N_3 = 0$  is empty) contain auxiliary variables that are also used for forecasting purposes. Further, let  $I(t-1)$  be the Hilbert space generated by the components of  $y_\tau$ , for  $\tau \leq t-1$ , i.e. an information set generated by the past realizations of  $y_t$ . Let  $I^2(t-1)$  be the Hilbert space generated by the variables  $\mathbf{y}_{i,\tau} \mathbf{y}_{j,\tau}$ ,  $1 \leq i, j \leq N$

for  $\tau \leq t-1$ .  $I_{-1}(t-1)$  is the closed subspace of  $I(t-1)$  generated by the components of  $(y'_{2\tau}, y'_{3\tau})'$  and  $I_{-1}^2(t-1)$  is the closed subspace of  $I^2(t-1)$  generated by the variables  $y_{i,\tau}y_{j,\tau}$ ,  $N_1+1 \leq i, j \leq N$  for  $\tau \leq t-1$ . For any subspace  $I_{t-1}$  of  $I(t-1)$  and for  $N_1+1 \leq i \leq N_1+N_2$ , we denote by  $P(y_{it}|I_{t-1})$  the affine projection of  $y_{it}$  on  $I_{t-1}$ , i.e. the best linear prediction of  $y_{it}$  based on the variables in  $I_{t-1}$  and a constant term.

For the Granger causal analysis of stochastic processes we propose to consider modeling framework of VARMA-GARCH processes. This approach is practical for empirical work, in contrast to Florens & Mouchart (1985) where the problem of causality is treated without any particular process assumed. Granger noncausality from  $y_1$  to  $y_2$  is defined as follows.

**Definition 1.**  $y_1$  does not Granger cause  $y_2$  given  $y_3$ , denoted by  $y_1 \xrightarrow{G} y_2|y_3$ , if each component of the error vector  $y_{2t} - P(y_{2t}|I_{-1}(t-1))$  is orthogonal to  $I(t-1)$  for all  $t \in \mathbb{Z}$ .

Definition 1 states simply that the forecast of  $y_2$  cannot be improved by adding to the information set past realizations of  $y_1$ . We now translate this notion to modeling framework of vector autoregressive moving average processes.

Suppose that  $y_t$  follows a  $N$ -dimensional VARMA(p,q) process:

$$\alpha(L)y_t = \beta(L)\epsilon_t \quad (2)$$

where  $L$  is a lag operator such that  $L^i y_t = y_{t-i}$ ,  $\alpha(z) = I_N - \alpha_1 z - \dots - \alpha_p z^p$ ,  $\beta(z) = I_N - \beta_1 z - \dots - \beta_q z^q$  are matrix polynomials with  $I_N$  denoting the identity matrix of order  $N$ , and  $\{\epsilon_t : t \in \mathbb{Z}\}$  is a white noise process with nonsingular covariance matrix  $\mathbf{V}$ . Comte & Lieberman (2000) mention that all the results in this section hold also if  $E[\epsilon_t \epsilon_t' | I(t-1)] = H_t$ , i.e. if the conditional covariance matrix of  $\epsilon_t$  is time varying, provided that unconditional covariance matrix,  $E[H_t] = \mathbf{V}_\epsilon$ , is constant and nonsingular. Without loss of generality we assumed in (2) that  $E[y_t] = 0$ . Further, we assume for the process (2) that:

**Assumption 1.** All the roots of  $|\alpha(z)| = 0$  and all the roots of  $|\beta(z)| = 0$  are outside the unit circle.

**Assumption 2.** The terms  $\alpha(z)$  and  $\beta(z)$  are left coprime and satisfy other identifiability conditions given in Lütkepohl (2005).

These assumptions guarantee that the VARMA(p,q) process is stationary, invertible and identified. Let the vector  $y_t$  be partitioned as in (1), then we can write (2) as:

$$\begin{bmatrix} \alpha_{11}(L) & \alpha_{12}(L) & \alpha_{13}(L) \\ \alpha_{21}(L) & \alpha_{22}(L) & \alpha_{23}(L) \\ \alpha_{31}(L) & \alpha_{32}(L) & \alpha_{33}(L) \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{bmatrix} = \begin{bmatrix} \beta_{11}(L) & \beta_{12}(L) & \beta_{13}(L) \\ \beta_{21}(L) & \beta_{22}(L) & \beta_{23}(L) \\ \beta_{31}(L) & \beta_{32}(L) & \beta_{33}(L) \end{bmatrix} \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \\ \epsilon_{3t} \end{bmatrix} \quad (3)$$

Given assumptions 1 - 2 and the VARMA(p,q) process in the form as in (3) we repeat after Theorem 4 of Boudjellaba et al. (1994) the conditions for Granger noncausality. We

say that  $y_1$  does not Granger cause  $y_2$  given  $y_3$  ( $y_1 \not\stackrel{G}{\rightarrow} y_2|y_3$ )<sup>1</sup> if and only if

$$\Gamma_{ij}(z) = \det \begin{bmatrix} \alpha_{11}^j(z) & \beta_{11}(z) & \beta_{13}(z) \\ \tilde{\alpha}_{N_1+i,j}(z) & \beta_{21}^i(z) & \beta_{23}^i(z) \\ \alpha_{31}^j(z) & \beta_{31}(z) & \beta_{33}(z) \end{bmatrix} = 0 \quad \forall z \in \mathbb{C} \quad (4)$$

for  $i = 1, \dots, N_2$  and  $j = 1, \dots, N_1$ ; where  $\alpha_{lk}^j(z)$  is the  $j$ th column of  $\alpha_{lk}(z)$ ,  $\beta_{lk}^i(z)$  is the  $i$ th row of  $\beta_{lk}(z)$ , and  $\tilde{\alpha}_{N_1+1,j}(z)$  is the  $(i, j)$ -element of  $\alpha_{21}(z)$ .

In general, the condition (4) leads to  $N_1 N_2$  determinant conditions. Each of them can be represented in a form of polynomial in  $z$  of degree  $p + q(N_1 + N_3)$ :  $\Gamma_{ij}(z) = \sum_{i=1}^{p+q(N_1+N_3)} a_i z^i$ , where  $a_i$  are nonlinear functions of parameters of the VARMA process. Notice that  $\Gamma_{ij} = 0 \Rightarrow a_i = 0$  for  $i = 1, \dots, p + q(N_1 + N_3)$  which gives restrictions for Granger noncausality.

**Example 1.** Let  $y_t$  be  $N = 3$  dimensional VARMA(1,0) process,  $N_1 = N_2 = N_3 = 1$  and assume we are interested to know whether  $y_1$  Granger causes  $y_2$ . The restriction for such a case is

$$\mathbf{R}^I(\psi) = \alpha_{21} = 0 \quad (5)$$

where  $\psi$  is a vector containing all the parameters of the model and  $\psi \in \Psi \subset \mathbb{R}^k$  and  $k$  denote the dimension of  $\psi$ .

**Example 2.** Let now  $y_t$  be VARMA(1,1) process of the same dimension and partitioning as before. The determinant condition (4) leads to the following set of restrictions:

$$\mathbf{R}_1^{\text{II}}(\psi) = \alpha_{11}(\beta_{23}\beta_{31} - \beta_{21}\beta_{33}) + \beta_{21}(\beta_{11}\beta_{33} - \beta_{13}\beta_{31}) + \alpha_{31}(\beta_{13}\beta_{21} - \beta_{11}\beta_{23}) = 0 \quad (6a)$$

$$\mathbf{R}_2^{\text{II}}(\psi) = \beta_{21}(\alpha_{11} - 2\beta_{33} - \beta_{11}) + \beta_{23}(\alpha_{31} - \beta_{31}) = 0 \quad (6b)$$

$$\mathbf{R}_3^{\text{II}}(\psi) = \alpha_{21} - \beta_{21} = 0 \quad (6c)$$

and let  $\mathbf{R}^{\text{II}}(\psi) = (\mathbf{R}_1^{\text{II}}(\psi), \mathbf{R}_2^{\text{II}}(\psi), \mathbf{R}_3^{\text{II}}(\psi))'$ .

The problem of testing restrictions (5) and (6) is dealt with in Section 4.

### 3. Parameter restrictions for second-order Granger noncausality in GARCH models

The subsequent section consists of two parts. In the first one, we present a multivariate GARCH model with constant conditional correlations. For this model, we discuss conditions for stationarity, asymptotic properties, classical and Bayesian estimation and how it was used to model and test volatility transmissions. Finally, we present its VARMA and VAR formulations in order to derive parametric conditions for second-order Granger noncausality in the second part of the section.

---

<sup>1</sup>The word *given* denoted by  $|$  in descriptions of noncausality relations does not mean probabilistic conditioning. It should be read *when there are other variables in the system grouped in* . . . .

*The GARCH(r,s) model and its properties.* The conditional mean part of the model is described with VARMA process (2) whose residual term  $\epsilon_t$  has the following conditional variance process:

$$\epsilon_t = D_t r_t \quad (7a)$$

$$r_t \sim i.i.d.(0, P) \quad (7b)$$

where  $D_{ii,t} = [\sqrt{h_{i,t}}]$  for  $i = 1, \dots, N$  is a  $N \times N$  diagonal matrix with square roots of conditional variances on the diagonal,  $r_t$  is a vector of standardized residuals that follows *i.i.d.* with zero mean and scale matrix  $P$ .

Conditional variances of  $\epsilon_t$  follow the multivariate GARCH(r,s) process of Jeantheau (1998):

$$h_t = \omega + A(L)\epsilon_t^{(2)} + B(L)h_t \quad (8)$$

where  $h_t$  is a  $N \times 1$  vector of conditional variances of  $\epsilon_t$ ,  $\omega$  is a  $N \times 1$  vector of constant terms,  $\epsilon_t^{(2)} = (\epsilon_{1t}^2, \dots, \epsilon_{Nt}^2)'$  is a vector of squared residuals,  $A(L) = \sum_{i=1}^r A_i L^i$  and  $B(L) = \sum_{i=1}^s B_i L^i$  are matrix polynomials of ARCH and GARCH effect respectively. All the matrices in  $A(L)$  and  $B(L)$  are of dimension  $N \times N$  and allow for volatility transmissions from one series to another.  $P$  is a positive definite constant conditional correlation matrix with ones on the diagonal.

The conditional covariance matrix of the residual term  $\epsilon_t$  is decomposed into  $E[\epsilon_t \epsilon_t' | I(t-1)] = H_t = D_t P D_t$ . For the matrix  $H_t$  to be well defined positive definite covariance matrix,  $h_t$  have to be positive for all  $t$  and  $P$  positive definite (see Bollerslev, 1990). Vector of conditional variances,  $h_t = E[\epsilon_t^{(2)} | I(t-1)]$ , is also the best linear predictor of  $\epsilon_t^{(2)}$ ,  $h_t = P(\epsilon_t^{(2)} | I(t-1))$ . The vector of unconditional variances of  $\epsilon_t$  is given by  $\sigma^2 = E[\epsilon_t^{(2)}] = E[h_t] = [I_N - A(1) - B(1)]^{-1} \omega$  for stationary GARCH(r,s) process.

The VARMA(p,q)-GARCH(r,s) model described by (2), (7) and (8) that is an object of analysis in this study has its origins in constant conditional correlation GARCH (CCC-GARCH) model proposed by Bollerslev (1990). That model consisted of  $N$  univariate GARCH equations describing vector of conditional variances  $h_t$ . CCC-GARCH model is equivalent to equations (7) and (8) with diagonal matrices  $A(L)$  and  $B(L)$ . Its extended version, with non-diagonal matrices  $A(L)$  and  $B(L)$ , was used in Karolyi (1995) and analyzed by Jeantheau (1998). He & Teräsvirta (2004) call this model extended CCC-GARCH (ECCC-GARCH).

Jeantheau (1998) proves that GARCH(r,s) model, as in (8), has unique, ergodic, weakly and strictly stationary solution when  $\det[I_N - A(z) - B(z)] = 0$  has its unit roots outside the unit circle. He & Teräsvirta (2004) give sufficient conditions for the existence of the fourth moments and derives complete forth moments structure. For instance, they give the conditions for existence and analytical form of  $E[\epsilon_t^{(2)} \epsilon_t^{(2)'}]$ , as well as for the  $n$ th order autocorrelation matrix of  $\epsilon_t^{(2)}$ ,  $\mathbf{R}_N(n) = D_N^{-1} \Gamma_N(n) D_N^{-1}$ , where  $\Gamma_N(n) = [\gamma_{ij}(n)] = E[(\epsilon_t^{(2)} - \sigma^2)(\epsilon_{t-n}^{(2)} - \sigma^2)']$  and  $D_N = \text{diag}(\sqrt{\sigma^2})$ .

The VARMA-GARCH model with constant conditional correlations has well established asymptotic properties. They can be set under the following assumptions:

**Assumption 3.** 1. All the roots of  $|I_N - A(z) - B(z)| = 0$  are outside the unit circle.

2. All the roots of  $|I_N - B(z)| = 0$  are outside the unit circle.

**Assumption 4.** The multivariate GARCH(r,s) model is minimal in the sense of Jeantheau (1998).

Under the assumptions 3.1 and 4 GARCH(r,s) model is stationary and identifiable. Jeantheau (1998) showed that the minimum contrast estimator for multivariate GARCH model is strongly consistent under, among others, stationarity and identifiability conditions. Ling & McAleer (2003) proved strong consistency of QMLE for VARMA-GARCH model under assumptions 1 - 4. Moreover, they have set asymptotic normality of QMLE provided that  $E\|y_t\|^6 < \infty$ .

It was already mentioned that for positive definiteness of conditional covariance matrix,  $H_t$ ,  $h_t$  has to be positive for all  $t$ . Usual parameter conditions for  $h_t$  to be positive are  $\omega > 0$  and  $[A_i]_{jk}, [B_l]_{jk} \geq 0$  for  $i = 1, \dots, r, l = 1, \dots, s$  and  $j, k = 1, \dots, N$ . Conrad & Karanasos (2009) derived such conditions that some elements of  $A_i, B_l$  ( $i = 1, \dots, r; l = 1, \dots, s$ ) and even  $\omega$  are allowed to be negative. Still, it is not known whether asymptotic results hold under these conditions. However, their empirical usefulness has been proven, as Conrad & Karanasos (2009) found that some parameters of the model responsible for volatility transmissions are negative.

Classical estimation with maximum likelihood method has been presented in Bollerslev (1990). The maximum likelihood estimator is an argument maximizing the likelihood function,  $\hat{\psi} = \arg \max_{\psi \in \Psi} L(\psi; y)$ . The log-likelihood functions for Normal and Student-t distributions are respectively:

$$l_N(\psi; y) = -\frac{TN}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T (\ln |H_t| + \epsilon_t' H_t^{-1} \epsilon_t) \quad (9a)$$

$$l_t(\psi; y) = T \left[ \ln \Gamma\left(\frac{N+\nu}{2}\right) - \ln \Gamma\left(\frac{\nu}{2}\right) - \frac{N}{2} \ln(\pi(\nu-2)) \right] - \frac{1}{2} \sum_{t=1}^T \ln |H_t| - \frac{N+\nu}{2} \sum_{t=1}^T \ln \left[ 1 + \frac{1}{\nu-2} \epsilon_t' H_t^{-1} \epsilon_t \right] \quad (9b)$$

where  $\Gamma(\cdot)$  is Euler's gamma and  $|\cdot|$  a matrix determinant. Algorithms maximizing the log-likelihood function, like BHHH algorithm (see Berndt, Hall, Hall & Hausman, 1974), use analytical derivatives. Fiorentini, Sentana & Calzolari (2003) provide analytical expressions for the score, Hessian, and information matrix of multivariate GARCH models with Student-t conditional distributions of residuals. In Bayesian estimation of vector GARCH models numerical integration methods are used. Vrontos, Dellaportas & Politis (2003) propose Metropolis-Hastings algorithm (see Chib & Greenberg, 1995, and references therein) for the estimation of model (8).

A family of GARCH models has already been used in volatility spillovers literature. More specifically, the empirical works of Worthington & Higgs (2004) and Caporale, Pittis & Spagnolo (2006) used the BEKK-GARCH model of Engle & Kroner (1995) to prove volatility transmissions between stock exchange indices. The issue of causality in variance or second-order causality (both defined in the next paragraph) was treated by Comte & Lieberman (2000) who derived necessary and sufficient conditions on parameters of the model for second-order noncausality between two vectors of variables. No testing procedure was, however, available due to the lack of asymptotic results. Comte & Lieberman (2003) filled in the gap deriving asymptotic normal distribution for QMLE under the assumption of bounded moments of order 8 for  $\epsilon_t$ . Hafner & Herwartz (2008) use the results of these two papers and propose a Wald statistics for noncausality in variance hypothesis. As a consequence of using asymptotic derivations of Comte & Lieberman (2003), the test also requires finiteness of 8th order moments of the error term. Hafner (2009) presented the conditions under which temporal aggregation in GARCH models does not influence testing of causality in variance.

Karolyi (1995) used VARMA-CCC-GARCH model to show the necessity of modeling volatility spillovers for inference about transmissions in returns of stock exchange indexes. The assumption of constant conditional correlation may be too strong for stock exchange data. CCC-GARCH model proved, however, its usefulness in modeling exchange rates volatility. In their recent study Omrane & Hafner (2009) use trivariate model for volatility spillovers between exchange rates. Conrad & Karanasos (2009) and Nakatani & Teräsvirta (2008) showed important case that volatility transmissions may be negative, the former for the system containing inflation rate and output growth, and the later for Japanese stock returns. The formal test for volatility transmissions was proposed by Nakatani & Teräsvirta (2009). Their Lagrange multiplier test statistics for the hypothesis of no volatility transmissions ( $A(L)$  and  $B(L)$  diagonal) versus volatility transmissions ( $A(L)$  and  $B(L)$  non-diagonal) assumes the existence of the 6th order moments of residual term,  $E|\epsilon_t^6| < \infty$ . Woźniak (2010) introduce the notion of Granger second-order causality and causality in variance for ECCC-GARCH models for the setting in which the vector of variables is partitioned in two parts. This was the case also for Comte & Lieberman (2000) and Hafner & Herwartz (2008) in BEKK-GARCH modeling. In this paper we extend the analysis such that an inference about causality between two (vectors of) variables is performed when in the system there are also other variables used for forecasting.

Before we introduce the notion of Granger noncausality for conditional variances, first, we present GARCH(r,s) model, (8), in VARMA and VAR formulations. Define a process  $v_t = \epsilon_t^{(2)} - h_t$ . Then  $\epsilon_t^{(2)}$  follows a VARMA process given by:

$$\phi(L)\epsilon_t^{(2)} = \omega + \theta(L)v_t \quad (10)$$

where  $\phi(L) = I_N - A(L) - B(L)$  and  $\theta(L) = I_N - B(L)$  are matrix polynomials of VARMA representation of GARCH(r,s) process. Suppose  $\epsilon_t^{(2)}$  and  $v_t$  are partitioned analogously as



$y_t$  in (1). Then (10) can be written in a form:

$$\begin{bmatrix} \phi_{11}(L) & \phi_{12}(L) & \phi_{13}(L) \\ \phi_{21}(L) & \phi_{22}(L) & \phi_{23}(L) \\ \phi_{31}(L) & \phi_{32}(L) & \phi_{33}(L) \end{bmatrix} \begin{bmatrix} \epsilon_{1t}^{(2)} \\ \epsilon_{2t}^{(2)} \\ \epsilon_{3t}^{(2)} \end{bmatrix} = \begin{bmatrix} \omega_{1t} \\ \omega_{2t} \\ \omega_{3t} \end{bmatrix} + \begin{bmatrix} \theta_{11}(L) & \theta_{12}(L) & \theta_{13}(L) \\ \theta_{21}(L) & \theta_{22}(L) & \theta_{23}(L) \\ \theta_{31}(L) & \theta_{32}(L) & \theta_{33}(L) \end{bmatrix} \begin{bmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{bmatrix}. \quad (11)$$

Given assumption 3.2 VARMA process (10) is invertible and can be written in a VAR form:

$$\Pi(L)\epsilon_t^{(2)} - \sigma^* = v_t \quad (12)$$

where  $\Pi(L) = \theta(L)^{-1}\phi(L) = [I_N - B(L)]^{-1}[I_N - A(L) - B(L)]$  is a matrix polynomial of VAR representation of GARCH(r,s) process and  $\sigma^* = \theta(1)^{-1}\omega$  is a constant term. Again, partitioning the vectors we rewrite (12) in the form:

$$\begin{bmatrix} \Pi_{11}(L) & \Pi_{12}(L) & \Pi_{13}(L) \\ \Pi_{21}(L) & \Pi_{22}(L) & \Pi_{23}(L) \\ \Pi_{31}(L) & \Pi_{32}(L) & \Pi_{33}(L) \end{bmatrix} \begin{bmatrix} \epsilon_{1t}^{(2)} \\ \epsilon_{2t}^{(2)} \\ \epsilon_{3t}^{(2)} \end{bmatrix} - \begin{bmatrix} \sigma_{1t}^* \\ \sigma_{2t}^* \\ \sigma_{3t}^* \end{bmatrix} = \begin{bmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{bmatrix}. \quad (13)$$

Under assumption 3 processes (10) and (12) are both stationary.

*Parameter restrictions.* In this paragraph, we present the main theoretical findings of the paper, which derive the conditions for second-order Granger noncausality for ECCC-GARCH model. We start with defining two concepts: Granger noncausality in variance and second-order Granger noncausality. Further, we derive the parametric conditions in Theorems 1 and 2 and discuss their novelty.

Robins et al. (1986) introduced the concept of Granger causality for conditional variances. Comte & Lieberman (2000) called this concept *second-order Granger causality* and distinguished it from *Granger causality in variance*. We define both of them in the following forms:

**Definition 2.**  $y_1$  does not second-order Granger cause  $y_2$  given  $y_3$ , denoted by  $y_1 \xrightarrow{so} y_2|y_3$ , if

$$P([y_{2t} - P(y_{2t}|I(t-1))]^{(2)}|I^2(t-1)) = P([y_{2t} - P(y_{2t}|I(t-1))]^{(2)}|I_{-1}^2(t-1)) \quad \forall t \in \mathbb{Z}$$

**Definition 3.**  $y_1$  does not Granger cause  $y_2$  in variance given  $y_3$ , denoted by  $y_1 \xrightarrow{v} y_2|y_3$ , if

$$P([y_{2t} - P(y_{2t}|I(t-1))]^{(2)}|I^2(t-1)) = P([y_{2t} - P(y_{2t}|I_{-1}(t-1))]^{(2)}|I_{-1}^2(t-1)) \quad \forall t \in \mathbb{Z} \quad (14)$$

where  $[.]^{(2)}$  means that we square every element of a vector. The difference between the two definitions is in Hilbert spaces on which  $y_{2t}$  is projected. On the right-hand side of definition 2 we take the affine projection of  $y_{2t}$  on  $I(t-1)$ , whereas on the right-hand side of definition 3 we take the affine projection of  $y_{2t}$  on  $I_{-1}(t-1)$ . In other words, before one consider whether there is *second-order Granger noncausality* she/he needs first

to filter out Granger causality in mean. Further, an implicit assumption in the definition of *Granger noncausality in variance* is that  $y_1$  does not Granger cause  $y_2$ ,  $y_1 \not\stackrel{G}{\rightarrow} y_2|y_3$ . The relations between Granger noncausality, Granger noncausality in variance and second-order Granger noncausality were established in Comte & Lieberman (2000) and are as follows:

$$y_1 \xrightarrow{v} y_2|y_3 \Leftrightarrow (y_1 \xrightarrow{G} y_2|y_3 \text{ and } y_1 \xrightarrow{so} y_2|y_3).$$

Two implications of this statement are that definitions 2 and 3 are equivalent when  $y_1$  does not Granger cause  $y_2$ , and conversely, that when  $y_1$  Granger causes  $y_2$  then Granger noncausality in variance is excluded, but still  $y_1$  may not second-order cause  $y_2$ .

Under the assumptions 1 - 4 the VARMA-GARCH models with constant conditional correlations are stationary identifiable and invertible in both of their parts: a VARMA process for  $y_t$  and for  $\epsilon_t^{(2)}$ . One more assumption is needed in order to state noncausality relations in conditional variances process:

**Assumption 5.** The process  $v_t$  is covariance stationary with covariance matrix  $V_v$ .

We now introduce a theorem in which second-order Granger noncausality relations are set:

**Theorem 1.** Let  $\epsilon_t^{(2)}$  follow a stationary vector autoregressive process as in (12) partitioned as in (13) that is identifiable (assumptions 1 - 5). Then,  $y_1$  does not second-order Granger cause  $y_2$  given  $y_3$  (denoted by  $y_1 \xrightarrow{so} y_2|y_3$ ) if and only if

$$\Pi_{21}(z) \equiv 0 \quad \forall z \in \mathbb{C}. \quad (15)$$

The Theorem 1 is an adaptation of Proposition 1 of Boudjellaba et al. (1992) to ECCC-GARCH models in VAR representation for  $\epsilon_t^{(2)}$ . It sets the conditions for second-order noncausality between two (vectors of) variables when in the system there are other auxiliary variables  $y_{3t}$ . Parametric condition (15), however, is unfit for the practical use. This is due to the fact that  $\Pi_{21}(L)$  is highly nonlinear function of parameters of the original GARCH(r,s) process (8). Moreover, it is a polynomial of potentially infinite order. Therefore, evaluation of the matrix polynomial  $\Pi(z)$  is further presented in Lemma 1 and Theorem 2. The Lemma, proof of the Theorem 1 as well as proofs of other theorems from this paper can be found in Appendix A.

**Theorem 2.** Let  $\epsilon_t^{(2)}$  follow a stationary vector autoregressive moving average process as in (10) partitioned as in (11) that is identifiable and invertible (assumptions 1 - 5). Then  $y_1$  does not second-order Granger cause  $y_2$  given  $y_3$  (denoted by  $y_1 \xrightarrow{so} y_2|y_3$ ) if and only if

$$\Gamma_{ij}^{so}(z) = \det \begin{bmatrix} \phi_{11}^j(z) & \theta_{11}(z) & \theta_{13}(z) \\ \varphi_{n_1+i,j}(z) & \theta_{21}^i(z) & \theta_{23}^i(z) \\ \phi_{31}^j(z) & \theta_{31}(z) & \theta_{33}(z) \end{bmatrix} = 0 \quad \forall z \in \mathbb{C} \quad (16)$$

for  $i = 1, \dots, N_2$  and  $j = 1, \dots, N_1$ ; where  $\phi_{lk}^j(z)$  is the  $j$ th column of  $\phi_{lk}(z)$ ,  $\theta_{lk}^i(z)$  is the  $i$ th row of  $\theta_{lk}(z)$ , and  $\varphi_{n_1+1,j}(z)$  is the  $(i, j)$ -element of  $\phi_{21}(z)$ .

Similarly as this was the case for restrictions (4), condition (16) leads to  $N_1 N_2$  determinant conditions. Each of them can be represented in a form of polynomial in  $z$  of degree  $\max(r, s) + s(N_1 + N_3)$ :  $\Gamma_{ij}^{so}(z) = \sum_{i=1}^{\max(r,s)+(N_1+N_3)s} b_i z^i$ , where  $b_i$  are nonlinear functions of parameters of GARCH process. We obtain parameter restrictions for the hypothesis of second-order Granger noncausality setting  $b_i = 0$  for  $i = 1, \dots, \max(r, s) + s(N_1 + N_3)$ . Such restrictions are ready to be tested.

The innovation of the condition (16) is that the second-order noncausality from  $y_{1t}$  to  $y_{2t}$  is analyzed when there are other variables in the system collected in the vector  $y_{3t}$ . The restrictions can even be used for big systems of variables. In Granger causality analysis this is particularly important to consider sufficiently large set of variables. Sims (1980) on the example of vector moving average model showed that the Granger causal relation may appear in the model due to omitted variables problem. Further, Lütkepohl (1982) showed that because of the omitted variables problem a noncausality relation may arrive. The conclusion of these two papers is maintained for second-order causality in multivariate GARCH models: one should consider sufficiently large set of variables in order to avoid omitted variables bias problem.

The condition (16) generalizes results from other studies. Comte & Lieberman (2000) derive alike restriction for BEKK-GARCH model with the difference that vector  $y_t$  is partitioned only into two sub-vectors  $y_{1t}^*$  and  $y_{2t}^*$ . Woźniak (2010) does the same for ECCC-GARCH model. The fact that the vector of variables is partitioned in three and not only two sub-vectors has serious implications for testing Granger causality relations in conditional variances. Notice that under such conditions, the formulation of some hypotheses is not even possible. This is because, in general, the fact that  $y_{1t}^* \xrightarrow{so} y_{2t}^*$  (which can be written as  $y_{1t} \xrightarrow{so} (y_{2t}, y_{3t})$ ) does not imply that  $y_{1t} \xrightarrow{so} y_{2t}|y_{3t}$  or that  $y_{1t} \xrightarrow{so} y_{3t}|y_{2t}$ . Moreover, the results of Woźniak (2010) are nested in the condition (16) by setting  $N_3 = 0$ .

To conclude this section we illustrate the derivation of the parameter restrictions for several processes that are often used in empirical works.

**Example 3.** Let  $y_t$  be trivariate GARCH(1,1) process ( $N = 3$  and  $r = s = 1$ ). Then VARMA process for  $\epsilon_t^{(2)}$  is as follows:

$$\begin{bmatrix} 1 - (A_{11} + B_{11})L & -(A_{12} + B_{12})L & -(A_{13} + B_{13})L \\ -(A_{21} + B_{21})L & 1 - (A_{22} + B_{22})L & -(A_{23} + B_{23})L \\ -(A_{31} + B_{31})L & -(A_{32} + B_{32})L & 1 - (A_{33} + B_{33})L \end{bmatrix} \begin{bmatrix} \epsilon_{1t}^2 \\ \epsilon_{2t}^2 \\ \epsilon_{3t}^2 \end{bmatrix} = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} + \begin{bmatrix} 1 - B_{11}L & -B_{12}L & -B_{13}L \\ -B_{21}L & 1 - B_{22}L & -B_{23}L \\ -B_{31}L & -B_{32}L & 1 - B_{33}L \end{bmatrix} \begin{bmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{bmatrix}. \quad (17)$$

If an investigator is interested in testing the hypothesis  $y_1 \xrightarrow{so} y_2|y_3$  then applying Theorem 2 she/he obtains the following set of restrictions:

$$\mathbf{R}_1^{III}(\psi) = A_{11}(B_{23}B_{31} - B_{21}B_{33}) + A_{31}(B_{13}B_{21} - B_{11}B_{23}) = 0 \quad (18a)$$

$$\mathbf{R}_2^{III}(\psi) = A_{11}B_{21} + A_{31}B_{23} = 0 \quad (18b)$$

$$\mathbf{R}_3^{III}(\psi) = A_{21} = 0. \quad (18c)$$

If an investigator is interested in testing the hypothesis  $y_1 \xrightarrow{so} (y_2, y_3)$ , then from Theorem 2 the conditions are given by:

$$\det \begin{bmatrix} 1 - (A_{11} + B_{11})z & 1 - B_{11}z \\ -(A_{i1} + B_{i1})z & -B_{i1}z \end{bmatrix} = 0 \quad \text{for } i = 2, 3,$$

which results in restrictions:

$$\mathbf{R}_1^{IV}(\psi) = A_{11}B_{21} = 0 \quad \text{and} \quad \mathbf{R}_2^{IV}(\psi) = A_{21} = 0 \quad (19a)$$

$$\mathbf{R}_3^{IV}(\psi) = A_{11}B_{31} = 0 \quad \text{and} \quad \mathbf{R}_4^{IV}(\psi) = A_{31} = 0. \quad (19b)$$

**Example 4.** Let  $\epsilon_t^{(2)}$  follow  $N = 3$  dimensional ARCH(r) process and one is interested whether  $y_1$  second-order Granger causes  $y_2$  (given  $y_3$ ). The restriction for such a case is

$$\mathbf{R}^V(\psi) = A_{i,21} = 0 \quad \text{for } i = 1, \dots, r. \quad (20)$$

#### 4. Bayesian testing of noncausality in VARMA-GARCH models

In the following section, the problem of testing restrictions on parameters of the original VARMA-GARCH model is considered. Apart from deriving tests for Granger causality and second-order Granger causality hypotheses separately, we propose a joint test of the parametric restrictions from conditions (4) and (16). Thus, not only do we emphasize the role of joint modeling of transmissions in conditional mean and conditional variance processes, but we present a complete set of tools for the underlying analysis as well. Moreover, a Bayesian testing procedure is proposed as a solution for some of the drawbacks of the classical tests.

We start with presenting a classical Wald test of Boudjellaba et al. (1992) for the parameter restrictions for Granger noncausality in VARMA process. The Wald test has this desirable feature that it requires estimating only the most general model. What is not required, is the estimation of restricted models. Thus, estimating one model one can do both: perform testing procedure and analyze the parameters responsible for transmissions. Before a test can be performed an investigator should first estimate the VARMA-GARCH model and derive a set of parametric restrictions from condition (4). The Wald statistics is given by

$$\mathbf{W}(\hat{\psi}_m) = T\mathbf{R}(\hat{\psi}_m)' [\mathbf{T}(\hat{\psi}_m)' \mathbf{V}(\hat{\psi}_m) \mathbf{T}(\hat{\psi}_m)]^{-1} \mathbf{R}(\hat{\psi}_m) \quad (21)$$

where  $\psi_m$  is a sub-vector of  $\psi$  containing the parameters used in  $l_m \times 1$  vector of parametric restrictions  $\mathbf{R}(\psi_m)$ ,  $\mathbf{V}(\hat{\psi}_m)$  is the asymptotic covariance matrix of  $\sqrt{T}(\hat{\psi}_m - \psi_m)$ , and  $\mathbf{T}(\hat{\psi}_m)$  is a  $m \times l_m$  matrix of derivatives

$$\mathbf{T}(\hat{\psi}_m) = \left. \frac{\partial \mathbf{R}(\psi_m)}{\partial \psi_m} \right|_{\psi_m = \hat{\psi}_m}.$$

Under the null hypothesis of Granger noncausality  $\mathbf{W}(\hat{\psi}_m)$  has asymptotic  $\chi^2(l_m)$  distribution. However, in equation (21)  $\mathbf{T}(\psi_m)$  must be of full rank. Boudjellaba et al. (1992) claims that this is not always the case under the true null hypothesis and derive sequential testing procedure. The procedure has however this drawback that it might not be conclusive (for details on the testing procedure see Boudjellaba et al., 1992).

Summarizing, in order to test Granger noncausality conditions for VARMA model, like restrictions (6), an investigator can use classical Wald test statistic that has well known properties. It has a limitation as well in the form of full rank condition for the matrix of first derivatives of the parametric restrictions  $\mathbf{R}(\psi_m)$  with respect to  $\psi_m$ .

Suppose that  $\psi_n$  contains the parameters that appear in the restrictions for second-order Granger noncausality for multivariate GARCH model derived from condition (16). In order to test such restrictions Wald test can also be used with test statistics  $\mathbf{W}(\psi_n)$ . Given that  $\sqrt{T}(\hat{\psi}_n - \psi_n)$  has asymptotic normal distribution the test statistic has  $\chi^2(l_n)$  asymptotic distribution. Still, the asymptotic covariance matrix of the parametric restrictions,  $[\mathbf{T}(\hat{\psi}_n)' \mathbf{V}(\hat{\psi}_n) \mathbf{T}(\hat{\psi}_n)]$  may under the null hypothesis be singular. Moreover, as Ling & McAleer (2003) proved,  $\sqrt{T}(\hat{\psi}_n - \psi_n)$  has asymptotic normal distribution under the condition of existence of 6th order moments of  $y_t$ . For many financial time series analyzed with multivariate GARCH models this may not be the case, as such data are often leptokurtic and the existence of higher order moments is uncertain.

The joint test of Granger noncausality and second-order Granger noncausality is a simple generalization of the two separate tests. Suppose that  $\psi_{mn}$  stacks the parameters from restrictions derived from conditions (4) and (16). The Wald test statistics for such a hypothesis is simply  $\mathbf{W}(\psi_{mn})$  and is asymptotically  $\chi^2(l_m + l_n)$  distributed. It also inherits the properties and limitations from both of separate tests.

We propose an alternative approach to testing. The Bayesian procedure presented in subsequent part overcomes the limitations of the Wald test. More specifically, singularities of the asymptotic covariance matrix of restrictions are excluded by construction and the assumptions of existence of higher order moments of time series are relaxed. Before, we explain in details solutions to this problem, a necessary introduction to Bayesian inference is first presented.

In Bayesian approach the model is fully specified by a prior distribution and a likelihood function. The prior distribution,  $p(\psi)$ , reflects the knowledge about the parameters that an investigator has before seeing the data,  $y$ . Of course, if one prefers not to perturb the inference with subjective beliefs she/he may choose to use uninformative prior distribution. Further, the prior beliefs are updated with information from the data that is represented by the likelihood function,  $L(\psi; y)$ . As a result of the update a posterior distribution of parameters of the model is obtained. The posterior distribution is proportional to the product of the prior distribution and the likelihood function

$$p(\psi|y) \propto L(\psi; y)p(\psi).$$

Suppose that the investigator is interested in some function of parameters,  $g(\psi)$ . Given the posterior distribution of parameters, a posterior distribution of function  $g(\cdot)$  is computed automatically,  $p(g(\psi)|y)$ . Moreover, every characteristics of this distribution is

available. For instance, a posterior mean of  $g(\psi)$  is calculated by definition of the expected value by integrating the product of the function and its distribution over the whole parameter space

$$E[g(\psi)|y] = \int_{\psi \in \Psi} g(\psi)p(g(\psi)|y)d\psi.$$

In order to evaluate such integral numerical methods need to be employed for GARCH models, as analytical forms are not known.

Let  $g(\psi) = \mathbf{R}(\psi) = \mathbf{R}$ , i.e. noncausality restrictions as in (6) or (18), be a  $l \times 1$  vector containing the values of function of interest. Then its posterior distribution is  $p(\mathbf{R}(\psi)|y)$  and posterior expected value and covariance matrix are respectively  $E[\mathbf{R}(\psi)|y]$  and  $V[\mathbf{R}(\psi)|y]$ . Now, let  $\{\mathbf{R}(\psi_i)\}_{i=1}^{S_1}$  be a sample of  $S_1$  draws from the stationary posterior distribution  $p(\mathbf{R}(\psi)|y)$  obtained with one of the numerical integration methods. Then one estimates the posterior mean and the covariance matrix with:

$$E[\widehat{\mathbf{R}(\psi)}|y] = S_1^{-1} \sum_{i=1}^{S_1} \mathbf{R}(\psi_i) \quad (22)$$

$$V[\widehat{\mathbf{R}(\psi)}|y] = S_1^{-1} \sum_{i=1}^{S_1} [\mathbf{R}(\psi_i) - E[\widehat{\mathbf{R}(\psi)}|y]] [\mathbf{R}(\psi_i) - E[\widehat{\mathbf{R}(\psi)}|y]]' \quad (23)$$

Define a scalar function  $\kappa : \mathbb{R}^l \rightarrow \mathbb{R}^+$  by

$$\kappa(\mathbf{R}) = [\mathbf{R} - E[\widehat{\mathbf{R}(\psi)}|y]]' V[\widehat{\mathbf{R}(\psi)}|y]^{-1} [\mathbf{R} - E[\widehat{\mathbf{R}(\psi)}|y]] \quad (24)$$

where  $\mathbf{R}$  is an argument of the function,  $E[\widehat{\mathbf{R}(\psi)}|y]$  and  $V[\widehat{\mathbf{R}(\psi)}|y]$  are defined in (22) and (23). The function  $\kappa$  is a positive semidefinite quadratic form of a real-valued vector. It gives a measure of the deviation of the value of the vector of restrictions from its posterior mean,  $\mathbf{R} - E[\widehat{\mathbf{R}(\psi)}|y]$ , rescaled by positive definite posterior covariance matrix,  $V[\widehat{\mathbf{R}(\psi)}|y]$ . Notice that the positive definite covariance matrix is a characteristic of the posterior distribution and by construction cannot be singular as long as the restrictions are linearly independent. I use function  $\kappa$  in order to test the noncausality conditions.

Suppose  $\{\mathbf{R}_i\}_{i=S_1+1}^{S_2}$  is a sample of  $S_2$  draws from the stationary posterior distribution  $p(\mathbf{R}(\psi)|y)$ . Using the transformation  $\kappa$  of the restrictions  $\mathbf{R}$  one obtain a sample of  $S_2$  draws of  $\kappa(\mathbf{R}_i)$  from the posterior distribution  $p(\kappa(\mathbf{R})|y)$ . The testing procedure for the null hypothesis of noncausality and the alternative hypothesis of causality is as follows:

**Step 1** Draw  $\{\kappa(\mathbf{R}_i)\}_{i=S_1+1}^{S_2}$  from  $p(\kappa(\mathbf{R})|y)$ .

**Step 2** Construct the Highest Posterior Density region of probability  $\pi$  ( $\text{HPD}_\pi$ ) for  $\kappa(\mathbf{R})$ .

**Step 3** Compute  $\kappa(0)$ .

**Step 4** If  $\kappa(\mathbf{0}) \notin \text{HPD}_{\pi}$ , then reject the null hypothesis, whereas if  $\kappa(\mathbf{0}) \in \text{HPD}_{\pi}$ , do not reject it.

The test based on the Highest Posterior Density region of  $\kappa(\mathbf{R})$  has several significant differences in comparison to classical tests. First of all, in Bayesian inference the observed data are treated as given, and given the data one proceeds with posterior inference. On the contrary, in classical approach dataset is an output of the data generating process conditioned on the parameters. This difference comes from the fact that classical estimation is based on the likelihood function which is identical to the conditional distribution of data given parameters. As a consequence, crucial for classical tests resampling properties are senseless in Bayesian approach.

Further, the posterior distribution is the distribution of the parameters of the model given the data. This means that it is the finite sample distribution of the parameters. Therefore, also the test is based on the exact finite sample distribution. In contrast, in classical inference for GARCH models only the asymptotic distribution of the estimator of the parameters is available. Since there is no need to refer to asymptotic theory in this study, there is also no need to keep its assumptions. As a result, the Bayesian test relaxes the assumptions of existence of higher order moments. Only the existence of fourth order moments is assumed (see Assumption 5) in comparison to the assumption of existence of sixth order moments in a classical derivation of the asymptotic distribution of QMLE.

## 5. Granger causality analysis of exchange rates

So far the methods of analysis and testing of causality in VARMA-GARCH models were presented. In the following section I use the derived conditions and Bayesian testing procedure in order to analyze the relations between series of exchange rates. This way we learn about the structure of integration of the series. Moreover, testing the dependencies I check the economic theory of efficiency of exchange rates markets. For instance, if the null hypothesis of second-order noncausality between time series is rejected, then we also reject efficiency of the exchange rates market. In the same manner, we prove that the market dynamics exhibits volatility persistence due to private information or heterogeneous beliefs.

I illustrate the use of the methods and testing of the economic theory with three time series of daily exchange rates. The series, all denominated in Euro, are Swiss franc (CHF/EUR), British pound (GBP/EUR) and United States dollar (USD/EUR). I analyze a series of logarithmic rates of return expressed in percentage points,  $y_t = 100(\ln x_t - \ln x_{t-1})$ , where  $x_t$  are levels of the exchange rates. The time span of the series begins in September 16, 2008 and ends in October 22, 2010, which gives  $T = 539$  observations for each of the series. The analyzed period starts a day after Lehman Brothers filed for Chapter 11 bankruptcy protection. Rationales behind this choice is that the series and its volatility in that period do not contain a structural break. Indeed, from Figure B.1 we see that the series exhibit a usual pattern of financial time series. The periods of higher and smaller volatility are driven by clustering of inflow of news modeled by GARCH processes.

I estimate a trivariate VAR(1)-GARCH(1,1) model in order to perform analysis and testing. The model with constant conditional correlations is obtained by setting the order of VARMA process (3) to  $(p, q) = (1, 0)$  and the order of GARCH process (8) to  $(r, s) = (1, 1)$ . I also choose the Student-t likelihood function (9b) that allows to model outliers that may appear in the data and are a common feature of the financial time series. I use the prior specification that I claim to be uninformative, but at the same time consists of proper probability distributions. For the constant conditional correlations,  $\rho$ , I assume a uniform distribution on interval  $(-1, 1)$ ,  $p(\rho) = \mathcal{U}(-1, 1)$ . For the degrees of freedom parameter,  $\nu$ , of the Student-t likelihood function a truncated normal distribution with mean 2 and standard deviation 50 is assumed,  $p(\nu) = \mathcal{N}(2, 50)\mathbf{1}(\nu \geq 2)$ , where  $\mathbf{1}(\cdot)$  is an indicator function equal to 1 when a condition in the brackets holds, and 0 otherwise. The choice of such a prior distribution means that *a priori* with probability equal to 0.57 the values of the parameter above 30 will appear. Consequently, the same prior probability is given to a normal likelihood function which is approximated with Student-t likelihood function with the degrees of freedom parameter value above 30. For the remaining  $k-4$  parameters,  $\psi_{k-4}$ , I assume a truncated to the parameter space,  $\psi \in \Psi$ , multivariate normal distribution with zero mean and covariance matrix containing 100 on a diagonal and zeros elsewhere,  $p(\psi_{k-4}) = \mathcal{N}^{k-4}(\mathbf{0}, 100I_{k-4})\mathbf{1}(\psi_{k-4} \in \Psi_{k-4})$ . Such a prior is practically flat on the parameter space, and thus, uninformative. I also assume that the three groups of parameters are *a priori* independent:

$$p(\psi) = p(\psi_{k-4})p(\nu)p(\rho).$$

For the Bayesian estimation I use the Metropolis-Hastings algorithm adopted for GARCH models by Vrontos et al. (2003). Having calibrated the covariance matrix of the sampling distribution of the algorithm the simulation is initiated by any starting values for parameters from the parameter space. After  $S^{out}$  steps of the algorithm, samples  $S_1$  and  $S_2$  of parameters from the ergodic posterior distribution are drawn. The convergence of the algorithm to the stationary posterior distribution is checked with cum-sum plots. From the samples  $\{\psi_i\}_{i=1}^{S_1}$  and  $\{\psi_i\}_{i=S_1+1}^{S_1+S_2}$  samples of  $\mathbf{R}(\psi)$  from  $p(\mathbf{R}(\psi)|y)$  and of  $\kappa(\mathbf{R})$  from  $p(\kappa(\mathbf{R})|y)$  are computed and the testing procedure described in Section 4 is performed.

I start the empirical investigation with analysis of the interaction between the returns of the time series. The results are presented in Table C.1 where the null hypotheses of Granger noncausality in conditional means are tested. Clearly, interactions in returns are apparent and very strong. The following relations are established by rejecting the null hypothesis of noncausality in favor of the alternative hypothesis of Granger causality between series. The Swiss franc exchange rate has strong impact on the British pound as well as on British pound and U.S. dollar taken jointly. Further, British pound influences the U.S. dollar and U.S. dollar and Swiss franc taken jointly. Finally, The U.S. dollar strongly influences Swiss franc and also has an impact on British pound. These effects are significant no matter if franc and pound are taken jointly or separately. I conclude the analysis of interactions between the returns of the exchange rates with the finding that each of the series is strongly influenced by two remaining series from the system.

The pattern for the network of interactions between returns of the series is changed



for volatilities. In Table C.2 the results of testing second-order Granger noncausality are presented. Second-order noncausality is only tested when before interactions in mean were filtered out. Two basic causal relations in conditional variances are established. Namely, Swiss franc's volatility affects British pound's volatility and is influenced by U.S. dollar's volatility. This very clearly sets a network of connections in risk. Further, each of the exchange rates influences both of the remaining taken jointly. However, only the dollar is significantly affected by franc and pound taken jointly.

The results of noncausality analysis in conditional mean and conditional variance processes imply the noncausality in variance relations. From (14) one sees that the noncausality in variance relation is the strongest concept as it implies and is implied by both: Granger noncausality and second-order noncausality. Therefore, only noncausality relations established for the first and the second conditional moments imply noncausality in variance. Such implied noncausality in variance relations are that British pound does not cause in variance Swiss franc and Swiss franc does not cause in variance U.S. dollar.

However, testing does not confirm these implications fully. In Table C.3 I present the results of testing of noncausality in variance hypotheses. Two noncausality relations cannot be rejected. British pound does not cause in variance Swiss franc, a relation that is actually implied. Further, U.S. dollar does not cause in variance British pound. This finding is not implied by previous results. The relation that is implied and that does not find support in joined testing is that franc does not cause dollar in variance. In general, all other null hypotheses are rejected, which proves that the integration of the exchange rates is strong. The integration process is present not only in returns, but in risk as well.

The finding that the implied structure of noncausalities in variance is not supported by joint testing is however possible. Notice that Granger noncausality and second-order noncausality relations are not independent. The rationale behind this statement is that the parameters of VARMA and GARCH processes are not independent and are in fact correlated. Therefore, the importance of joint testing is even more legitimate.

## 6. Conclusions

I presented a framework in which an investigator may analyze causal relations between time series. The conditions for noncausality between conditional variances of the time series were derived. Taken together with conditions for noncausality in conditional mean process they give a complete toolbox for testing economic theory and analysis of integration of the markets or assets. In order to test the conditions I propose a Bayesian test which avoids some of the problems of classical testing and as well as relaxes its assumptions.

## References

- Berndt, E. K., Hall, B. H., Hall, R. E., & Hausman, A. (1974). Estimation and Inference in Nonlinear Structural Models. *Annals of Economic and Social Measurement*, 3, 653–665.
- Bollerslev, T. (1990). Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized Arch Model. *The Review of Economics and Statistics*, 72, 498–505.

- Boudjellaba, H., Dufour, J.-M., & Roy, R. (1992). Testing Causality Between Two Vectors in Multivariate Autoregressive Moving Average Models. *Journal of the American Statistical Association*, 87, 1082–1090.
- Boudjellaba, H., Dufour, J.-M., & Roy, R. (1994). Simplified conditions for noncausality between vectors in multivariate ARMA models. *Journal of Econometrics*, 63, 271–287.
- Caporale, G. M., Pittis, N., & Spagnolo, N. (2006). Volatility transmission and financial crises. *Journal of Economics and Finance*, 30, 376–390.
- Chib, S., & Greenberg, E. (1995). Understanding the Metropolis-Hastings Algorithm. *The American Statistician*, 49, 327.
- Comte, F., & Lieberman, O. (2000). Second-Order Noncausality in Multivariate GARCH Processes. *Journal of Time Series Analysis*, 21, 535–557.
- Comte, F., & Lieberman, O. (2003). Asymptotic theory for multivariate GARCH processes. *Journal of Multivariate Analysis*, 84, 61–84.
- Conrad, C., & Karanasos, M. (2009). Negative Volatility Spillovers in the Unrestricted Eccc-Garch Model. *Econometric Theory*, (pp. 1–25).
- Engle, R. F., Ito, T., & Lin, W.-l. (1990). Meteor Showers or Heat Waves? Heteroskedastic Intra-Daily Volatility in the Foreign Exchange Market. *Econometrica*, 58, 525–542.
- Engle, R. F., & Kroner, K. F. (1995). Multivariate Simultaneous Generalized ARCH. *Econometric Theory*, 11, 122–150.
- Fiorentini, G., Sentana, E., & Calzolari, G. (2003). Maximum Likelihood Estimation and Inference in Multivariate Conditionally Heteroscedastic Dynamic Regression Models With Student t Innovations. *Journal of Business and Economic Statistics*, 21, 532–546.
- Florens, J. P., & Mouchart, M. (1985). A Linear Theory for Noncausality. *Econometrica*, 53, 157–176.
- Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37, 424–438.
- Hafner, C. M. (2009). Causality and forecasting in temporally aggregated multivariate GARCH processes. *Econometrics Journal*, 12, 127–146.
- Hafner, C. M., & Herwartz, H. (2008). Testing for causality in variance using multivariate GARCH models. *Annales d'Économie et de Statistique*, 89, 215 – 241.
- He, C., & Teräsvirta, T. (2004). An Extended Constant Conditional Correlation Garch Model and Its Fourth-Moment Structure. *Econometric Theory*, 20, 904–926.
- Hong, Y. (2001). A test for volatility spillover with application to exchange rates. *Journal of Econometrics*, 103, 183–224.
- Jeantheau, T. (1998). Strong consistency of estimators for multivariate arch models. *Econometric Theory*, 14, 70–86.
- Karolyi, G. A. (1995). A Multivariate GARCH Model of International Transmissions of Stock Returns and Volatility: The Case of the United States and Canada. *Journal of Business & Economic Statistics*, 13, 11–25.
- Koutmos, G., & Booth, G. (1995). Asymmetric volatility transmission in international stock markets. *Journal of International Money and Finance*, 14, 747–762.
- Lin, W.-L., Engle, R. F., & Ito, T. (1994). Do Bulls and Bears Move across Borders? International Transmission of Stock Returns and Volatility. *The Review of Financial Studies*, 7, 507–538.
- Ling, S., & McAleer, M. (2003). Asymptotic Theory for a Vector {Arma-Garch} Model. *Econometric Theory*, 19, 280–310.
- Lütkepohl, H. (1982). Non-causality due to omitted variables. *Journal of Econometrics*, 19, 367–378.
- Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer.
- Nakatani, T., & Teräsvirta, T. (2008). Positivity constraints on the conditional variances in the family of conditional correlation GARCH models. *Finance Research Letters*, 5, 88–95.
- Nakatani, T., & Teräsvirta, T. (2009). Testing for volatility interactions in the Constant Conditional Correlation GARCH model. *Econometrics Journal*, 12, 147–163.
- Omrane, W. B., & Hafner, C. M. (2009). Information Spillover, Volatility and the Currency Markets. *International Econometric Review*, 1, 47–59.
- Robins, R. P., Granger, C. W. J., & Engle, R. F. (1986). Wholesale and Retail Prices: Bivariate Time-Series

Modeling with forecastable Error Variances. In *Model Reliability* (pp. 1–17). The MIT Press.

Sims, C. A. (1972). Money, Income, and Causality. *The American Economic Review*, 62, 540 – 552.

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica*, 48, 1–48.

Taylor, M. P. (1995). The Economics of Exchange Rates. *Journal of Economic Literature*, 33, 13–47.

Vrontos, I. D., Dellaportas, P., & Politis, D. N. (2003). Inference for some multivariate ARCH and GARCH models. *Journal of Forecasting*, 22, 427–446.

Worthington, A., & Higgs, H. (2004). Transmission of equity returns and volatility in Asian developed and emerging markets: a multivariate GARCH analysis. *International Journal of Finance & Economics*, 9, 71–80.

Woźniak, T. (2010). Bayesian testing of second-order causality.

## Appendix A. Lemmas and Proofs

### Proof of the Theorem 1

From (13) (a constant term dropped in this proof without the loss of generality) it follows that  $\sum_{i=1}^3 \Pi_{2i}(L)\epsilon_{it}^{(2)} = v_{2t}$ . Writing  $\Pi_{2i}(L) = \delta_{2i}I_{n_2} - \sum_{k=1}^{\infty} \Pi_{2i(k)}L^k$ ,  $i = 1, 2, 3$ , where  $\delta_{ij}$  denotes the Kronecker delta and  $I_n$  stands for the identity matrix, we have:

$$\epsilon_{2t}^{(2)} = \sum_{i=1}^3 Z_{it} + v_{2t} \quad (\text{A.1})$$

where

$$Z_{it} = \sum_{k=1}^{\infty} \Pi_{2i(k)}\epsilon_{i,t-k}^{(2)} \quad i = 1, 2, 3. \quad (\text{A.2})$$

From (A.1) we have

$$P(\epsilon_{2t}^{(2)}|I^2(t-1)) = \sum_{i=1}^3 Z_{it} \quad (\text{A.3})$$

because  $v_t$  is orthogonal to  $I^2(t-1)$ . Further,

$$P(\epsilon_{2t}^{(2)}|I_{-1}^2(t-1)) = P(Z_{1t}|I_{-1}^2(t-1)) + Z_{2t} + Z_{3t}. \quad (\text{A.4})$$

If  $\Pi_{21}(z) \equiv 0$ , then we have  $Z_{1t} = 0$ , so that  $P(Z_{1t}|I_{-1}^2(t-1)) = 0$ . From (A.3) and (A.4) it follows that  $P(\epsilon_{2t}^{(2)}|I^2(t-1)) = P(\epsilon_{2t}^{(2)}|I_{-1}^2(t-1)) = Z_{2t} + Z_{3t}$ . The *if* part of the proof is completed.

Conversely, if  $y_1$  does not second-order Granger cause  $y_2$  (given  $y_3$ ), then  $P(\epsilon_{2t}^{(2)}|I^2(t-1)) = P(\epsilon_{2t}^{(2)}|I_{-1}^2(t-1))$ . From A.3 and A.4 it follows that  $Z_{1t} = P(Z_{1t}|I_{-1}^2(t-1))$ . That is, the components of  $Z_{1t}$  are contained in the closed span  $I_{-1}^2(t-1)$ . Therefore we can find sequences of matrices

$$\{\Pi_{22(k)}^K : k = 1, \dots, K\}_{K=1}^{\infty} \quad \text{and} \quad \{\Pi_{23(k)}^K : k = 1, \dots, K\}_{K=1}^{\infty}$$

such that:

$$\sum_{k=1}^K \Pi_{22(k)}^{(K)} \epsilon_{2,t-k}^{(2)} + \sum_{k=1}^K \Pi_{23(k)}^{(K)} \epsilon_{3,t-k}^{(2)} \xrightarrow{q.m.} -Z_{1t} \quad (\text{A.5})$$

where  $\xrightarrow{q.m.}$  refers to convergence in quadratic mean (as  $K \rightarrow \infty$ ).

From (A.2) for  $i = 1$  and from (A.5) we get:

$$\sum_{k=1}^K D_k^{(K)} \epsilon_{t-k}^{(2)} \xrightarrow{q.m.} 0 \quad (\text{A.6})$$

where  $D_k^{(K)} = (\Pi_{21(k)}, \Pi_{22(k)}, \Pi_{23(k)})$ ,  $k = 1, \dots, K$  and  $K \geq 1$  is the  $N_2 \times N$  matrix.

Multiplying the left hand side of (A.6) by  $v'_{t-1}$ , we have:  $\sum_{k=1}^K D_k^{(K)} \epsilon_{t-k}^{(2)} v'_{t-1} \xrightarrow{L_1} 0$ , where  $\xrightarrow{L_1} 0$  means convergence in the  $L_1$  norm (as  $K \rightarrow \infty$ ). Hence  $\sum_{k=1}^K D_k^{(K)} E[\epsilon_{t-k}^{(2)} v'_{t-1}] = D_1^{(K)} V \rightarrow 0$ , because

$$E[\epsilon_{t-k}^{(2)} v'_{t-1}] = \begin{cases} V & k = 1 \\ 0 & k > 1. \end{cases}$$

With the matrix  $V$  being nonsingular, we have  $D_1^{(K)} \rightarrow 0$ . Hence  $\Pi_{21(1)} = 0$ . Similarly, multiplying the left hand side of (A.6) by  $v'_{t-2}$ , we find that  $D_1^{(K)} E[\epsilon_{t-1}^{(2)} v'_{t-2}] + D_2^{(K)} V \rightarrow 0$ , so that  $D_2^{(K)} V \rightarrow 0$ , and  $\Pi_{21(2)} = 0$ . And so on for  $k = 3, 4, \dots$ . Thus  $\Pi_{21(k)} = 0$ ,  $k \geq 1$  and  $\Pi_{21}(z) \equiv 0$ .

**Lemma 1.** Let  $\epsilon_t^{(2)}$  follow a stationary vector autoregressive moving average process as in (10) partitioned as in (11) that is identifiable and invertible (assumptions 1 - 5). Then  $y_1$  does not second-order Granger cause  $y_2$  given  $y_3$  (denoted by  $y_1 \xrightarrow{so} y_2|y_3$ ) if and only if

$$\bar{\phi}_{21}(z) - \bar{\theta}_{21}(z) \bar{\theta}_{11}(z)^{-1} \bar{\phi}_{11}(z) = 0 \quad \text{for } |z| < \delta \quad (\text{A.7})$$

where  $\delta$  is some positive constant and  $\bar{\phi}_{ij}(z)$  and  $\bar{\theta}_{ij}(z)$  are defined by:

$$\bar{\theta}_{ij}(z) = \theta_{ij}(z) - \theta_{i3}(z) \theta_{33}(z)^{-1} \theta_{3j}(z) \quad i, j = 1, 2 \quad (\text{A.8a})$$

$$\bar{\phi}_{ij}(z) = \phi_{ij}(z) - \theta_{i3}(z) \theta_{33}(z)^{-1} \theta_{3j}(z) \quad i, j = 1, 2 \quad (\text{A.8b})$$

### Proof of the Lemma 1

Process (10) or (11) being invertible (assumption 3.2) can be expressed as an infinite autoregressive process  $\Pi(L) \epsilon_t^{(2)} = v_t$  as in (12) and (13), where  $\Pi(L) = \theta(L)^{-1} \phi(L) = I_{N_2} - \sum_{i=1}^{\infty} AB^{i-1} L^i = [\Pi_{ij}(L)]_{i,j=1,2,3}$ . From Theorem 1 we know that  $y_1$  does not second-order Granger cause  $y_2$  if and only if  $\Pi_{21}(L) \equiv 0$ ,  $\forall z \in \mathbb{C}$ . We thus need to evaluate  $\Pi_{21}(z)$ .

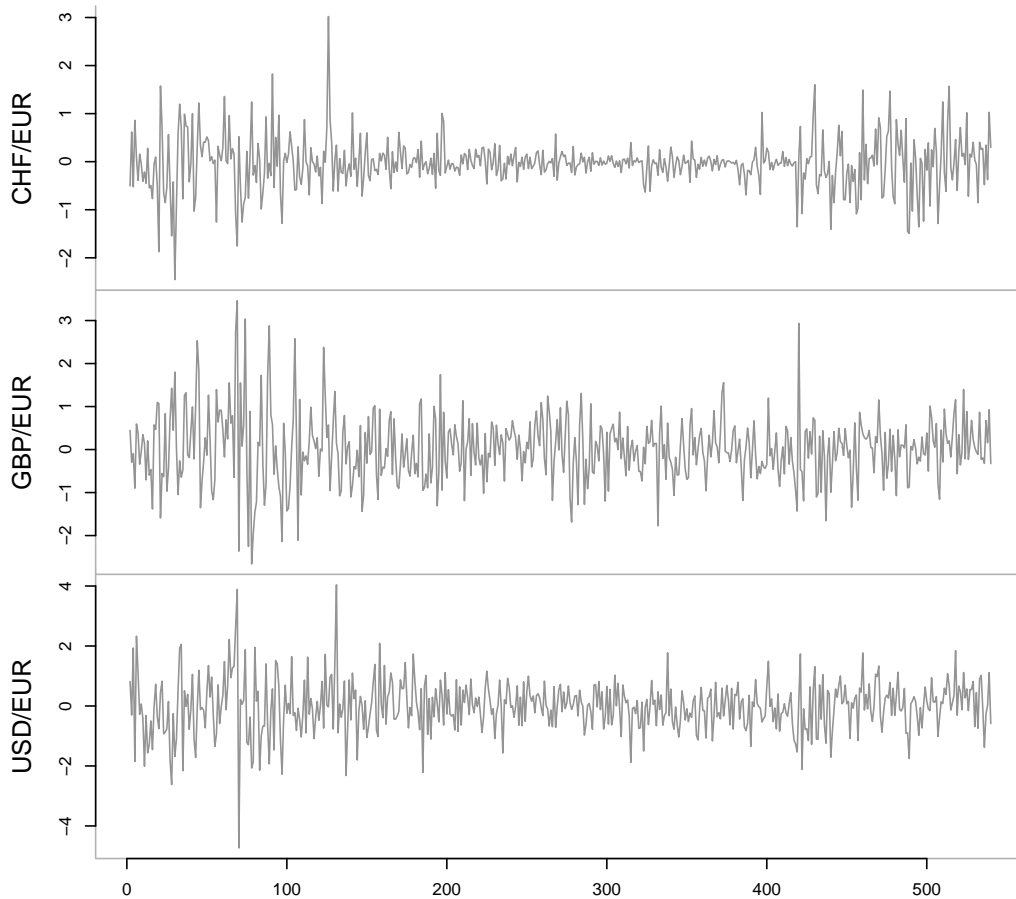
Notice, process (10) or (11) are exactly in the same form as process (2.5) of Boudjellaba et al. (1994) and  $\Pi(L) = \theta(L)^{-1} \phi(L)$  is the same subprocess as in Proposition 1 of Boudjellaba et al. (1992) that is used in the proof of Theorem 3 of Boudjellaba et al. (1994). The proof evaluates the matrix  $\Pi_{21}(L)$ , and therefore applies also to (15) which gives (A.7), and concludes this proof.

**Proof of the Theorem 2**

By Lemma 1  $y_1$  does not second-order Granger cause  $y_2$  if and only if condition (A.7) holds. In order to prove (16) apply Theorem 4 of Boudjellaba et al. (1994) and matrix transformations given in its proof.

## Appendix B. Data

Figure B.1: Data plot: (CHF/EUR,GBP/EUR,USD/EUR)



The graph presents logarithmic rates of return expressed in percentage points ( $y_t = 100(\ln x_t - \ln x_{t-1})$ ) of three exchange rates: Swiss franc, British pound and United States dollar all denominated in Euro. The data from September 16, 2008 to October 22, 2010 of length  $T = 539$  were downloaded from the European Central Bank website (<http://sdw.ecb.int/browse.do?node=2018794>).

## Appendix C. Results of hypotheses testing

Table C.1: Results of testing: Granger noncausality hypotheses

Hypothesis	$.HPD_{.95}$	$.HPD_{.99}$	$.\kappa(0)$	Figure
$y_1 \xrightarrow{G} y_2 y_3$	4.37	10.011	2137.382**	C.2.1.
$y_1 \xrightarrow{G} y_3 y_2$	3.599	9.094	2.029	C.2.2.
$y_2 \xrightarrow{G} y_1 y_3$	3.834	8.293	2.467	C.2.3.
$y_2 \xrightarrow{G} y_3 y_1$	3.611	7.554	558.389**	C.2.4.
$y_3 \xrightarrow{G} y_1 y_2$	3.816	5.957	7.291**	C.2.5.
$y_3 \xrightarrow{G} y_2 y_1$	3.515	6.0	4.395*	C.2.6.
$y_1 \xrightarrow{G} (y_2, y_3)$	7.287	11.751	2143.91**	C.3.1.
$y_2 \xrightarrow{G} (y_1, y_3)$	6.074	10.406	562.548**	C.3.2.
$y_3 \xrightarrow{G} (y_1, y_2)$	5.832	7.963	13.24**	C.3.3.
$(y_1, y_2) \xrightarrow{G} y_3$	6.369	10.636	588.395**	C.3.4.
$(y_1, y_3) \xrightarrow{G} y_2$	6.368	11.253	2244.26**	C.3.5.
$(y_2, y_3) \xrightarrow{G} y_1$	6.039	8.988	12.366**	C.3.6.

The table presents the results of testing of hypotheses of Granger noncausality. The hypotheses are stated in the first column, where  $\xrightarrow{G}$  is as in Definition 1 and  $y_1 = \text{CHF/EUR}$ ,  $y_2 = \text{GBP/EUR}$  and  $y_3 = \text{USD/EUR}$ . Second and third columns contain upper bounds of the one side highest posterior density region ( $HPD_\pi$ ) for  $\pi$  equal to .95 and .99 respectively. Fourth column presents the value of function  $\kappa$  defined in (24) that represents its values for the null hypothesis of noncausality from the first column. We use \*\* when the value of  $\kappa(0) \notin HPD_{.99}$  and \* when  $\kappa(0) \notin HPD_{.95}$  and  $\kappa(0) \in HPD_{.99}$ . The last column gives reference to figures with graphical representation of the results of testing.

Table C.2: Results of testing: second-order Granger noncausality hypotheses

Hypothesis	$.HPD_{.95}$	$.HPD_{.99}$	$.\kappa(0)$	Figure
$y_1 \xrightarrow{so} y_2 y_3$	9.565	17.307	18.146**	C.4.1.
$y_1 \xrightarrow{so} y_3 y_2$	8.951	15.439	4.417	C.4.2.
$y_2 \xrightarrow{so} y_1 y_3$	10.872	17.856	6.63	C.4.3.
$y_2 \xrightarrow{so} y_3 y_1$	12.89	22.223	3.955	C.4.4.
$y_3 \xrightarrow{so} y_1 y_2$	8.661	26.683	20.865*	C.4.5.
$y_3 \xrightarrow{so} y_2 y_1$	11.503	14.904	3.488	C.4.6.
$y_1 \xrightarrow{so} (y_2, y_3)$	10.842	13.486	20.936**	C.5.1.
$y_2 \xrightarrow{so} (y_1, y_3)$	11.491	19.174	20.736**	C.5.2.
$y_3 \xrightarrow{so} (y_1, y_2)$	11.491	19.174	20.736**	C.5.3.
$(y_1, y_2) \xrightarrow{so} y_3$	16.269	27.569	37.183**	C.5.4.
$(y_1, y_3) \xrightarrow{so} y_2$	17.853	28.243	7.773	C.5.5.
$(y_2, y_3) \xrightarrow{so} y_1$	15.318	24.175	23.096*	C.5.6.

The table presents the results of testing of hypotheses of second-order Granger noncausality. The hypotheses are stated in the first column, where  $\xrightarrow{so}$  is as in Definition 2 and  $y_1 = \text{CHF/EUR}$ ,  $y_2 = \text{GBP/EUR}$  and  $y_3 = \text{USD/EUR}$ .

\*\* the value of  $\kappa(0) \notin HPD_{.99}$

\* the value of  $\kappa(0) \notin HPD_{.95}$  and  $\kappa(0) \in HPD_{.99}$



Table C.3: Results of testing: Granger noncausality in variance hypotheses

Hypothesis	$.HPD_{.95}$	$.HPD_{.99}$	$.\kappa(0)$	Figure
$y_1 \overset{v}{\nrightarrow} y_2 y_3$	10.186	17.468	308.149**	C.6.1.
$y_1 \overset{v}{\nrightarrow} y_3 y_2$	11.768	16.447	11.918*	C.6.2.
$y_2 \overset{v}{\nrightarrow} y_1 y_3$	11.481	18.57	7.86	C.6.3.
$y_2 \overset{v}{\nrightarrow} y_3 y_1$	13.634	23.336	15.314*	C.6.4.
$y_3 \overset{v}{\nrightarrow} y_1 y_2$	9.924	26.731	31.247**	C.6.5.
$y_3 \overset{v}{\nrightarrow} y_2 y_1$	12.072	15.659	8.194	C.6.6.
$y_1 \overset{v}{\nrightarrow} (y_2, y_3)$	13.152	15.438	338.67**	C.7.1.
$y_2 \overset{v}{\nrightarrow} (y_1, y_3)$	14.29	20.799	60.369**	C.7.2.
$y_3 \overset{v}{\nrightarrow} (y_1, y_2)$	13.498	20.058	30.826**	C.7.3.
$(y_1, y_2) \overset{v}{\nrightarrow} y_3$	19.075	29.685	112.201**	C.7.4.
$(y_1, y_3) \overset{v}{\nrightarrow} y_2$	19.424	29.257	494.053**	C.7.5.
$(y_2, y_3) \overset{v}{\nrightarrow} y_1$	17.127	25.855	96.782**	C.7.6.

The table presents the results of testing of hypotheses of Granger noncausality in variance.

The hypotheses are stated in the first column, where  $\overset{so}{\nrightarrow}$  is as in Definition 2 and  $y_1 = \text{CHF/EUR}$ ,  $y_2 = \text{GBP/EUR}$  and  $y_3 = \text{USD/EUR}$ .

\*\* the value of  $\kappa(0) \notin HPD_{.99}$

\* the value of  $\kappa(0) \notin HPD_{.95}$  and  $\kappa(0) \in HPD_{.99}$

Figure C.2: Results of testing: Granger noncausality hypotheses I

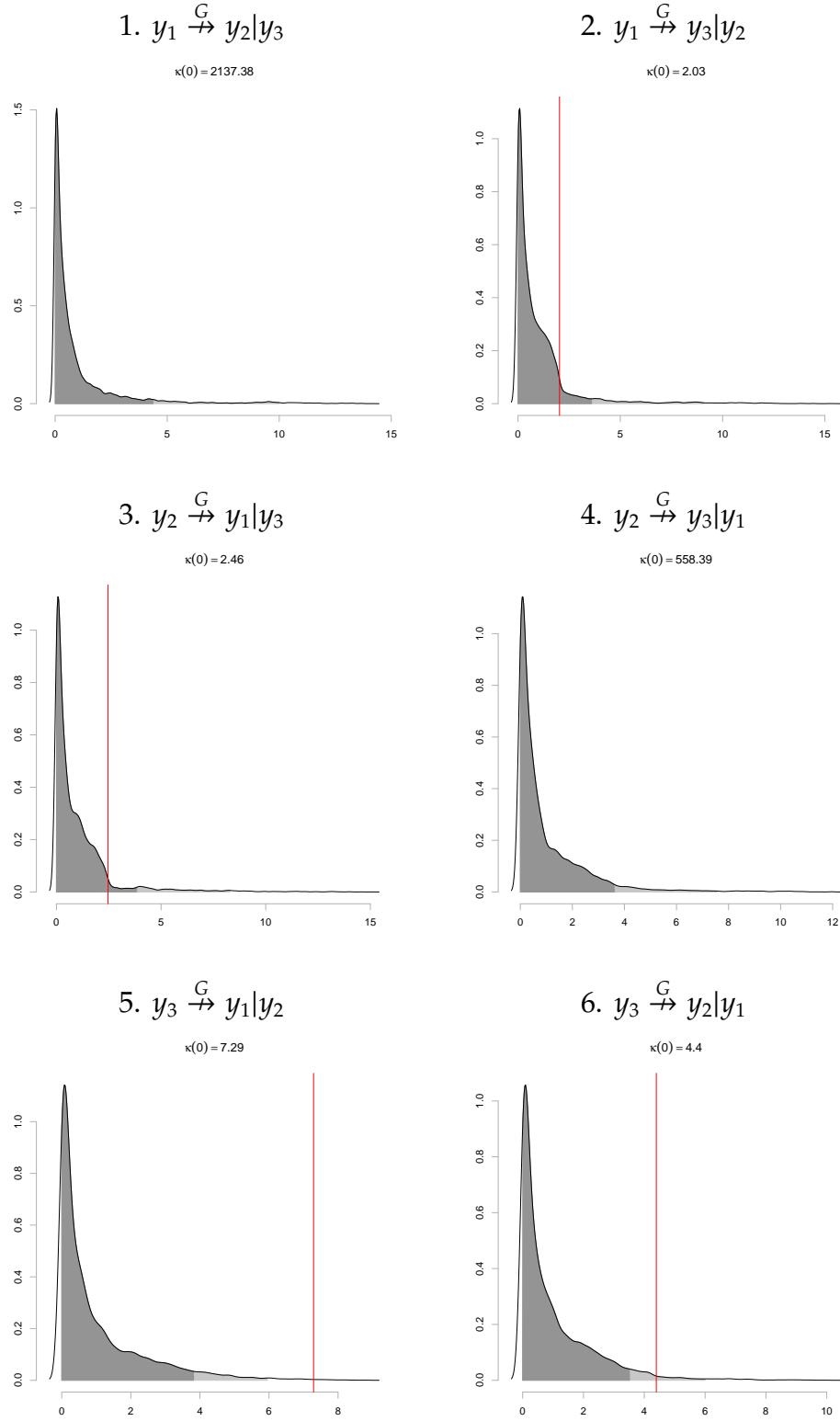


Figure C.3: Results of testing: Granger noncausality hypotheses II

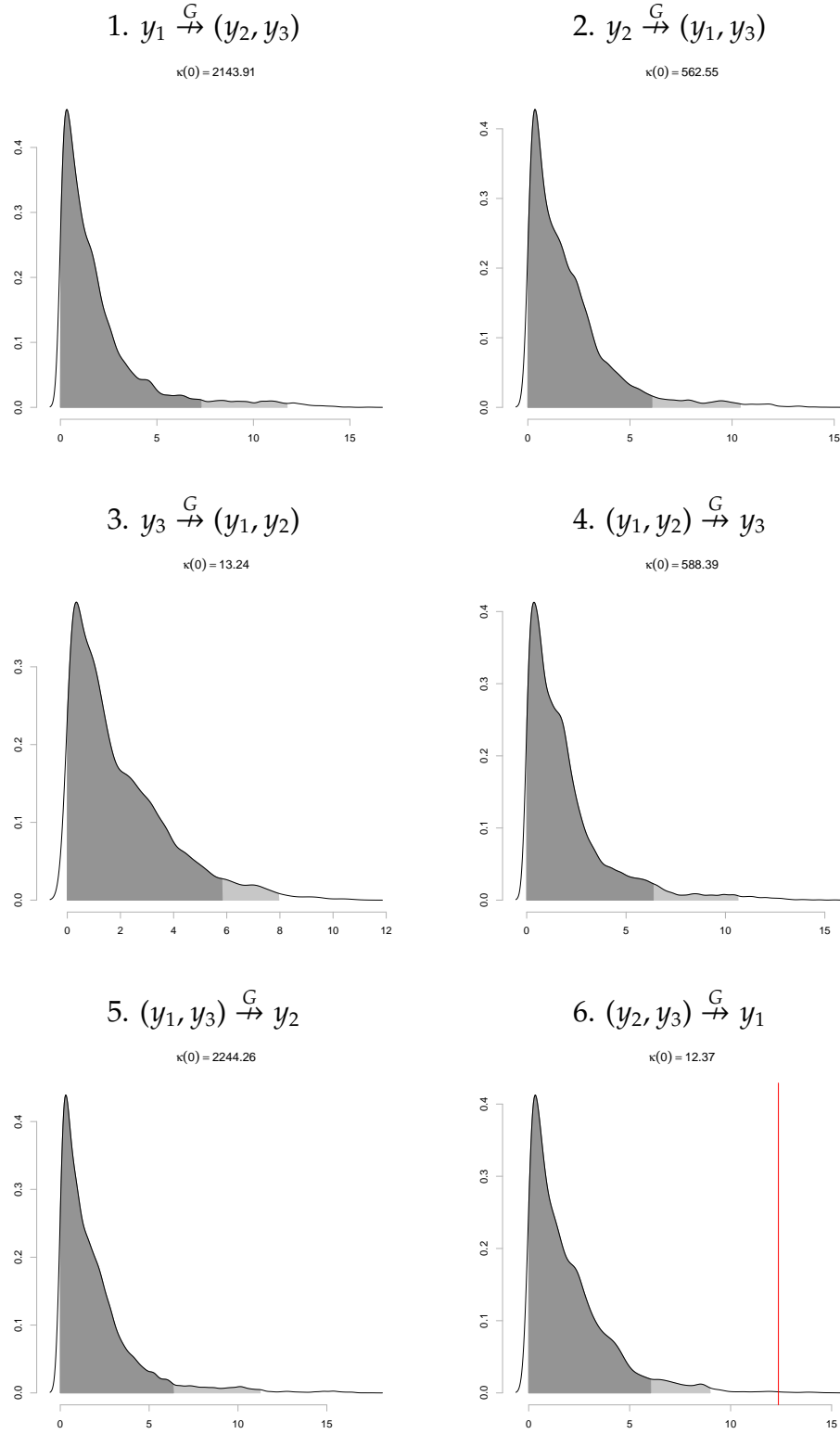


Figure C.4: Results of testing: second-order Granger noncausality I

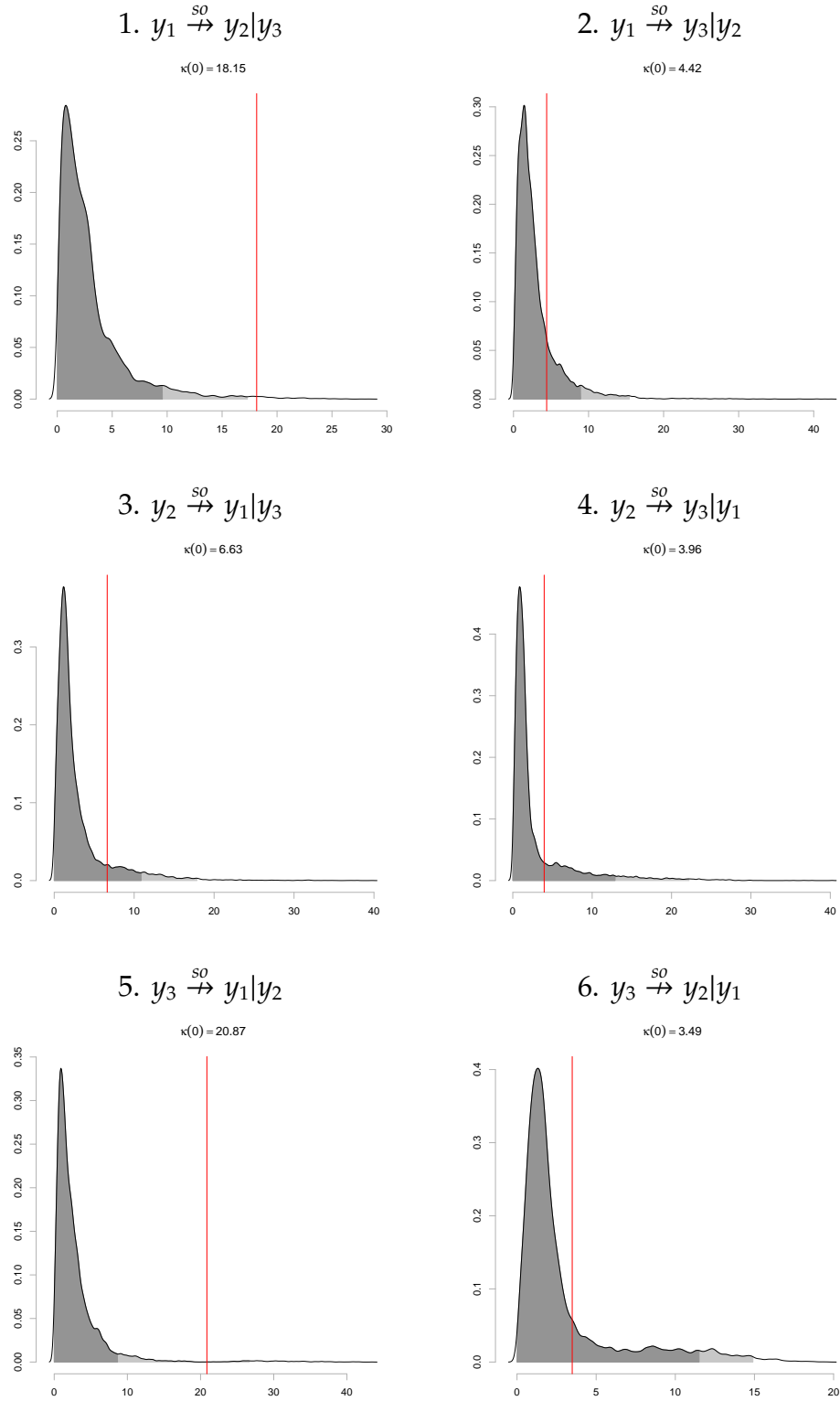


Figure C.5: Results of testing: second-order Granger noncausality II

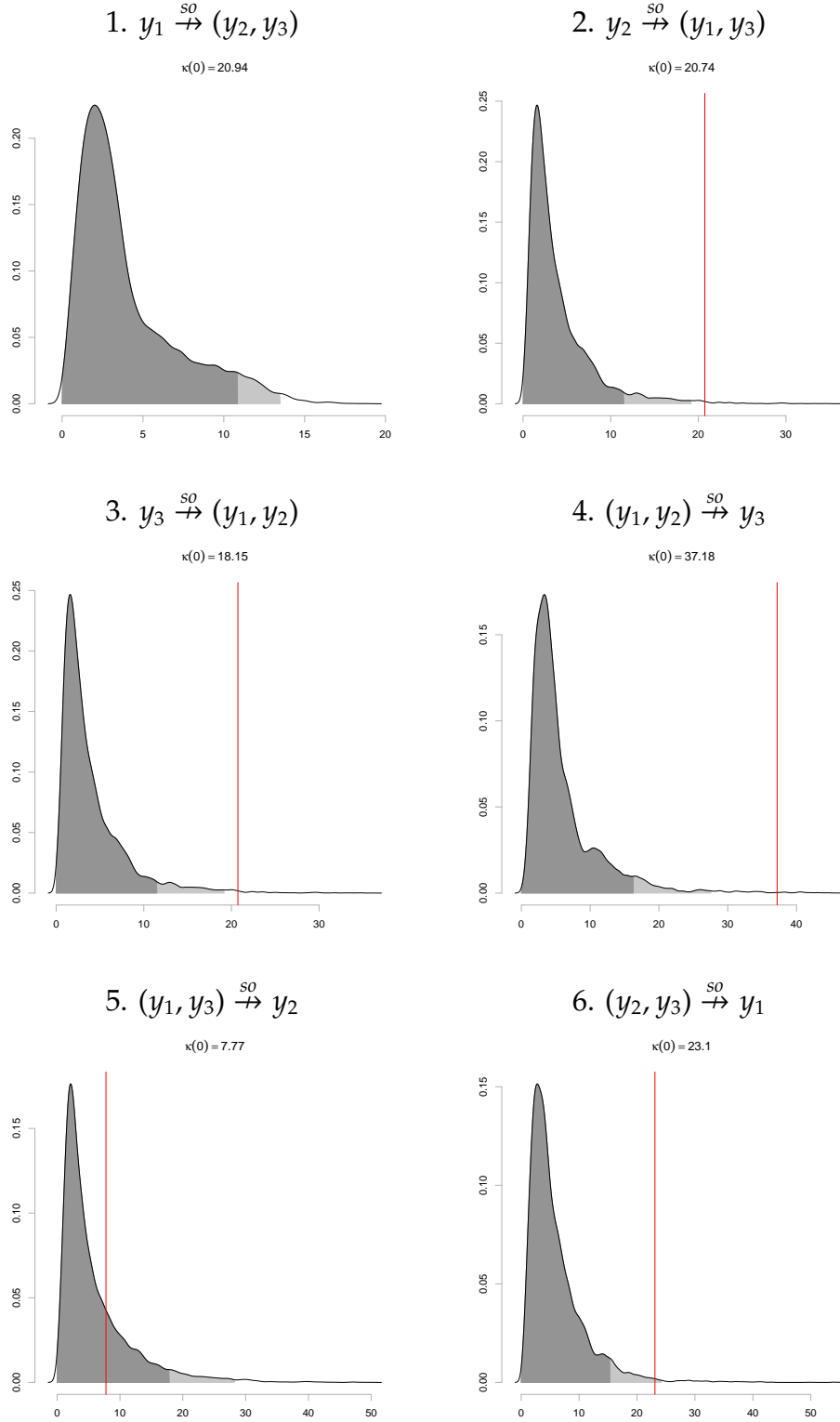


Figure C.6: Results of testing: Granger noncausality in variance hypotheses I

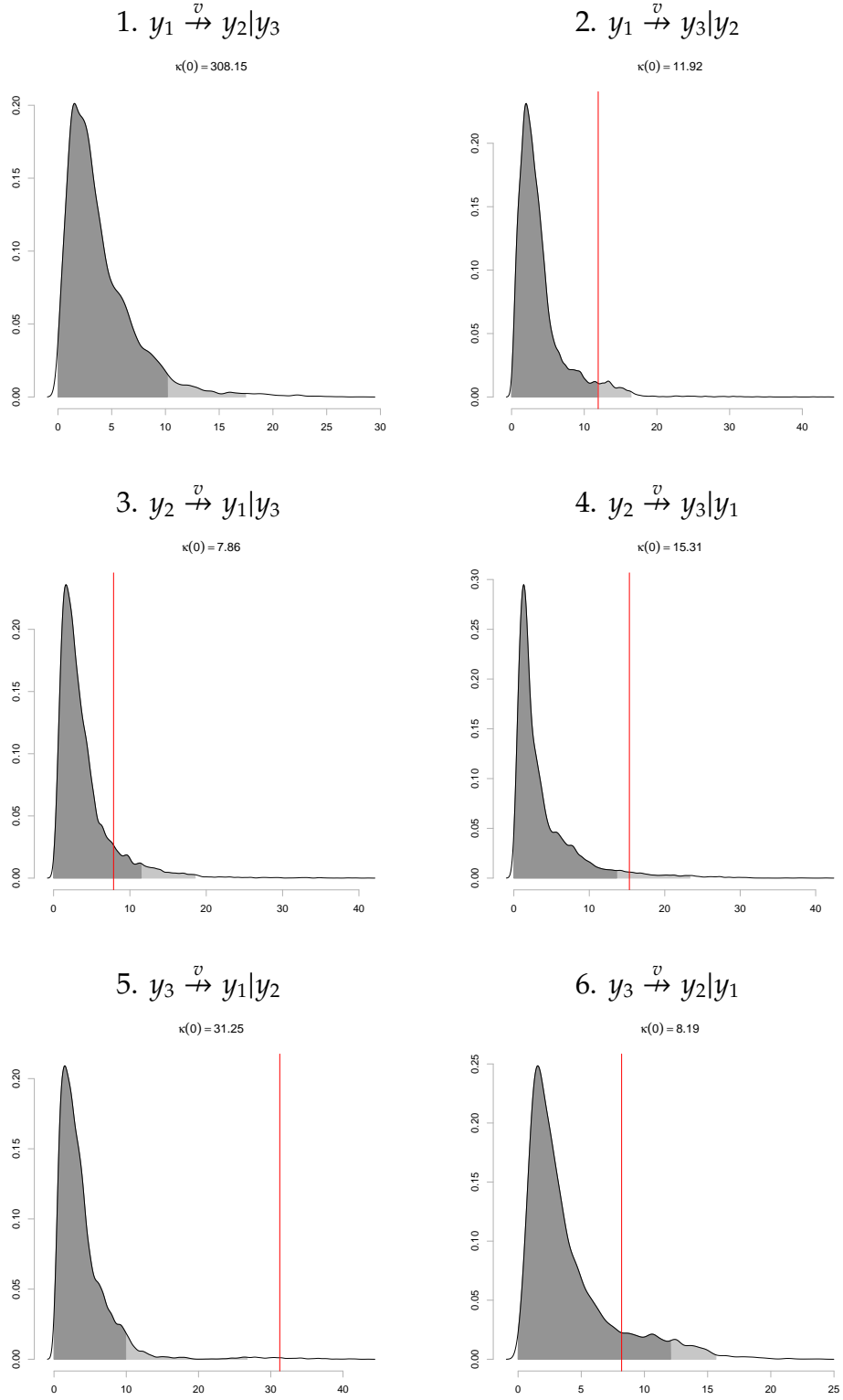


Figure C.7: Results of testing: Granger noncausality in variance hypotheses II

