European University Institute

EUI Working Papers

GRANGER-CAUSAL ANALYSIS OF VARMA-GARCH MODELS

Tomasz Woźniak

EUROPEAN UNIVERSITY INSTITUTE, FLORENCE

DEPARTMENT OF ECONOMICS

Granger-Causal Analysis of VARMA-GARCH Models

TOMASZ WOŹNIAK

Granger-causal analysis of VARMA-GARCH models [☆]

Tomasz Woźniak[a,b]

[a]*Department of Economics, University of Melbourne, Arts West Building, Victoria 3010, Australia*
[b]*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 their volatility. I investigate Granger causality in conditional mean and conditional variances of time series. For this purpose a VARMA-GARCH model 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 large systems of economic variables. Such systems should be considered in order to avoid the problem of omitted variable bias. Further, I propose a Bayesian Lindley-type testing procedure in order to evaluate hypotheses of noncausality. It avoids the singularity problem that may appear in the Wald test. Also, it relaxes the assumption of the existence of higher-order moments of the residuals required for the derivation of asymptotic results of the classical tests. In the empirical example, I find that the dollar-to-Euro exchange rate does not second-order cause the pound-to-Euro exchange rate, in the system of variables containing also the Swiss frank-to-Euro exchange rate, which confirms the *meteor shower* hypothesis of Engle, Ito & Lin (1990).

Keywords: Granger causality, second-order noncausality, VARMA-GARCH models, 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. One variable does not Granger-cause the other, if adding past observations of the former to the information set with which we forecast the latter does not improve this forecast. In this study, I look at the Granger noncausality concept for conditional variances of the time series. For this purpose two concepts of *second-order Granger noncausality* and *Granger noncausality in variance* are discussed (see also Comte & Lieberman, 2000; Robins, Granger & Engle, 1986). If one variable does not second-order Granger-cause the other, then past information about the variability of the former is dispensable for conditional variance the forecasting of the conditional variances of the latter. I investigate Granger causality in conditional mean and conditional variances of time series. *Granger noncausality in variance* is established when both *Granger noncausality* and *second-order noncausality* hold.

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

Why is information about Granger-causal relations between time series important? First of all, it gives an understanding of the 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. Therefore, modeling transmissions in volatility may have a significant impact on volatility forecasting. If there are Granger-causal relations in conditional variances, then such modeling is potentially important in all applications based on volatility forecasting such as portfolio selection, Value at Risk estimation and option pricing.

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

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

The contribution of this study is twofold. Firstly, I derive conditions for second-order Granger noncausality for a family of GARCH models. The conditions are applicable when the system of time series consists of a potentially large number of variables. Their novelty is that the second-order noncausality between two groups of variables is analyzed when there are other variables in the system as well. So far, such conditions have been derived when all the variables in the system were divided in two groups (e.g. Comte & Lieberman, 2000; Hafner & Herwartz, 2008; Woźniak, 2012). The introduced conditions reduce the dimensionality of the problem. They also allow the formation and testing of some hypotheses that could not be tested in the previous settings.

Secondly, I propose a Bayesian Lindley-type testing procedure of the conditions for Granger noncausality in conditional mean and noncausality in conditional variance processes. It is easily applicable and solves some of the drawbacks of the classical testing. In comparison with the Wald test of Boudjellaba, Dufour & Roy (1992), adapted to testing noncausality relations in the VARMA-GARCH model, the Bayesian test does not have the problem of singularities. In the Wald test considered so far 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; thus, a well defined covariance matrix is available as well. Additionally, in this study the existence only of fourth-order moments of time series is assumed, which is an improvement in comparison with the assumptions of available classical tests.

2

The reminder of this paper is organized as follows: the notation and the parameter restrictions for Granger noncausality in VARMA models are presented in Section 2. The GARCH model used in the analysis is set in Section 3. Alsoin this section, I present the main theoretical findings of the paper, deriving the conditions for Granger noncausality in the conditional variance process. In Section 4, I start by discussing of classical testing for noncausality in the VARMA-GARCH models, and then propose Bayesian testing with appealing properties. Section 5 presents an empirical illustration, with the example of daily exchange rates of the Swiss franc, the British pound and the US dollar all denominated in Euro. Section 6 concludes.

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:

$$y_t = (y_{1t}^{'}, y_{2t}^{'}, y_{3t}^{'})^{'}, \tag{1}$$

where y_{it} is a $N_i \times 1$ vector such that $y_{1t} = (y_{1t}, \ldots, y_{N_1.t})^{'}$, $y_{2t} = (y_{N_1+1.t}, \ldots, y_{N_1+N_2.t})^{'}$ and $y_{3t} = (y_{N_1+N_2+1.t}, \ldots, y_{N_1+N_2+N_3.t})^{'}$ ($N_1, N_2 \geq 1, N_3 \geq 0$ 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 relations. Vector y_3 (which for $N_3 = 0$ is empty) contains auxiliary variables that are also used for forecasting and modeling purposes. Further, let $I(t)$ be the Hilbert space generated by the components of y_τ, for $\tau \leq t$, i.e. an information set generated by the past realizations of y_t. Then, $\epsilon_{t+h} = y_{t+h} - P(y_{t+h}|I(t))$ is an error component.

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

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

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

Definition 1, proposed by Boudjellaba et al. (1992), states simply that the forecast of y_2 cannot be improved by adding to the information set past realizations of y_1.

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

$$\alpha(L)y_t = \beta(L)\epsilon_t, \tag{2}$$

for all $t = 1, \ldots, T$, where L is a lag operator such that $L^i y_t = y_{t-i}$, $\alpha(z) = I_N - \alpha_1 z - \cdots - \alpha_p z^p$, $\beta(z) = I_N + \beta_1 z + \cdots + \beta_q z^q$ are matrix polynomials. I_N denotes the identity matrix of order N, and $\{\epsilon_t : t \in \mathbb{Z}\}$ is a white noise process with nonsingular unconditional covariance matrix $\mathbf{V}$. Comte & Lieberman (2000) mention that all the results in this section hold also if $E[\epsilon_t \epsilon_t^{'}|I^2(t-1)] = H_t$, i.e. if the conditional covariance matrix of ϵ_t is time-varying, provided that unconditional covariance matrix, $E[H_t] = \mathbf{V}$, is constant and nonsingular. Without the loss of generality, I assumed in (2) that $E[y_t] = 0$, however any deterministic terms, such as a vector of constants, a time trend or seasonal dummies may be considered for modeling. 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 complex 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:

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

Given Assumptions 1–2 and the VARMA(p,q) process in the form as in (3), I repeat after Theorem 4 of Boudjellaba et al. (1994) the conditions for Granger noncausality. Therefore, y_1 does not Granger-cause y_2 given y_3 ($y_1 \overset{G}{\nrightarrow} y_2|y_3$)[1] if and only if:

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

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

In general, condition (4) leads to $N_1 N_2$ determinant conditions. Each of them can be represented in the form of a 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 $i = 1, \ldots, p + q(N_1 + N_3)$, 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 let one be interested in whether y_1 Granger-causes y_2. The restriction for such a case is:

$$\mathbf{R}^I(\theta) = \alpha_{21} = 0, \tag{5}$$

where θ is a vector containing all the parameters of the model, $\theta \in \Theta \subset \mathbb{R}^k$, and k denotes the dimension of θ.

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

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

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

$$\mathbf{R}_3^{II}(\theta) = \alpha_{21} - \beta_{21} = 0, \tag{6c}$$

and let $\mathbf{R}^{II}(\theta) = (\mathbf{R}_1^{II}(\theta), \mathbf{R}_2^{II}(\theta), \mathbf{R}_3^{II}(\theta))'$ be a vector collecting the values of the restrictions on the LHS.

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

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

---

[1] The word *given* denoted by | in descriptions of noncausality relations does not mean the proper probabilistic conditioning. Here it should be read *when there are other variables in the system grouped in . . . .*

GARCH(r,s) model and its properties. The conditional mean part of the model is described with the VARMA process (2) and a residual term ϵ_t following a conditional variance process:

$$\epsilon_t = D_t r_t, \tag{7a}$$

$$r_t \sim i.i.d.(0, \mathbf{C}), \tag{7b}$$

for all $t = 1, \dots, T$, where $D_{ii.t} = [\sqrt{h_{i.t}}]$ for $i = 1, \dots, N$ is a $N \times N$ diagonal matrix with conditional standard deviations on the diagonal, r_t is a vector of standardized residuals that follows *i.i.d.* with zero mean and a correlation matrix $\mathbf{C}$.

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

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

for all $t = 1, \dots, T$, where h_t is a $N \times 1$ vector of conditional variances of ϵ_t, ω 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 effects, 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. $\mathbf{C}$ is a positive definite constant conditional correlation matrix with ones on the diagonal.

The conditional covariance matrix of the residual term ϵ_t is decomposed into $E[\epsilon_t \epsilon_t' | I_y^2(t-1)] = H_t = D_t \mathbf{C} D_t$. For the matrix H_t to be a well defined positive definite covariance matrix, h_t must be positive for all t, and $\mathbf{C}$ positive definite (see Bollerslev, 1990). Given the normality of r_t, the vector of conditional variances is $E[\epsilon_t^{(2)} | I_y^2(t-1)] = h_t$. When r_t follows a t distribution with v degrees of freedom, the conditional variances are $E[\epsilon_t^{(2)} | I^2(t-1)] = \frac{v}{v-2} h_t$. In both cases the best linear predictor of $\epsilon_t^{(2)}$ is $h_t = P(\epsilon_t^{(2)} | I_y^2(t-1))$.

The VARMA(p,q)-GARCH(r,s) model described by (2), (7) and (8), which is object of the analysis in this study, has its origins in the constant conditional correlation GARCH (CCC-GARCH) model proposed by Bollerslev (1990). That model consists of N univariate GARCH equations describing the vector of conditional variances h_t. The 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) called this model extended CCC-GARCH (ECCC-GARCH).

Jeantheau (1998) proves that the GARCH(r,s) model, as in (8), has a unique, ergodic, weakly and strictly stationary solution when $\det[I_N - A(z) - B(z)] = 0$ has its unit roots outside the complex unit circle. He & Teräsvirta (2004) give sufficient conditions for the existence of the fourth moments and derive complete structure of fourth moments. For instance, they give the conditions for existence and analytical form of $E[\epsilon_t^{(2)} \epsilon_t^{(2)'}]$, as well as for the nth 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 $D_{ii.N} = [\sqrt{\sigma_i^2}]$ for $i = 1, \dots, N$.

The VARMA-ECCC-GARCH model 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 complex 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 Assumptions 3.1 and 4, the GARCH(r,s) model is stationary and identifiable. Jeantheau (1998) showed that the minimum contrast estimator for the multivariate GARCH model is strongly consistent under, among others, stationarity and identifiability conditions. Ling & McAleer (2003) proved the strong consistency of the QMLE for the VARMA-GARCH model under Assumptions 1–4. Moreover, they have set the 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

$i = 1, \ldots, r$, $l = 1, \ldots, s$ and $j, k = 1, \ldots, N$. Conrad & Karanasos (2009) derived conditions such that some elements of A_i, B_l ($i = 1, \ldots, r$; $l = 1, \ldots, s$) and even ω 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) have found that some parameters of the model responsible for volatility transmissions are negative.

Classical estimation with the maximum likelihood method has been presented in Bollerslev (1990). The maximum likelihood estimator is the argument maximizing the likelihood function, $\hat{\theta} = \arg\max_{\theta \in \Theta} L(\theta; \mathbf{y})$. The likelihood functions for Normal and t-distributed ϵs are, respectively:

$$L_N(\theta; \mathbf{y}) = (2\pi)^{-TN/2} \prod_{t=1}^{T} |H_t|^{-1/2} \exp\left(\epsilon_t' H_t^{-1} \epsilon_t\right)^{-1/2}, \text{ and} \tag{9a}$$

$$L_{St}(\theta; \mathbf{y}) = \prod_{t=1}^{T} \frac{\Gamma\left(\frac{\nu+N}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} ((\nu-2)\pi)^{-\frac{N}{2}} |H_t|^{-\frac{1}{2}} \left(1 + \frac{1}{\nu-2} \epsilon_t' H_t^{-1} \epsilon_t\right)^{-\frac{\nu+N}{2}}, \tag{9b}$$

where ϵ_t is defined in equations (2) and (7). $\Gamma(.)$ is Euler's gamma and $|.|$ a matrix determinant. Algorithms maximizing the likelihood function, such as the 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 t conditional distributions of residuals. In the Bayesian estimation of the 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 the model.

A broad family of GARCH models has already been used in the volatility spillovers literature. More specifically, the empirical works of Worthington & Higgs (2004) and Caporale, Pittis & Spagnolo (2006) use 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) has been treated by Comte & Lieberman (2000), who derived the conditions on parameters of the model for second-order noncausality between two vectors of variables. No testing procedure, however, was 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 eight for ϵ_t. Hafner & Herwartz (2008) use the results of these two papers and propose a Wald statistics for sufficient conditions for noncausality in variance hypothesis. As a consequence of using asymptotic derivations of Comte & Lieberman (2003), the test also requires finiteness of eighth-order moments of the error term. Hafner (2009) presents the conditions under which temporal aggregation in GARCH models does not influence testing of the causality in conditional variances.

Karolyi (1995) uses the VARMA-ECCC-GARCH model to show the necessity of modeling the volatility spillovers for the inference about transmissions in returns of stock exchange indexes. The assumption of constant conditional correlation may be too strong for such data. The ECCC-GARCH model, however, proved its usefulness in modeling the volatility of the exchange rates. In a recent study, Omrane & Hafner (2009) use the trivariate model for volatility spillovers between exchange rates. Conrad & Karanasos (2009) and Nakatani & Teräsvirta (2008) show the important case that volatility transmissions may be negative, the former for the system containing inflation rate and output growth, and the latter for Japanese stock returns. A formal test for the volatility transmissions has been 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 sixth-order moments of the residual term, $E|\epsilon_t^6| < \infty$. Woźniak (2012) introduces the notion of Granger second-order causality and causality in variance for ECCC-GARCH models for the setting similar to that of Comte & Lieberman (2000), in which the vector of variables is partitioned in two parts. In this paper, I extend the analysis such that an inference about causality between two (vectors of) variables is performed when there are also other variables in the system used for forecasting.

Before I introduce the notion of Granger noncausality for conditional variances, I 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:

$$\phi(L)\epsilon_t^{(2)} = \omega + \psi(L)v_t, \tag{10}$$

for all $t = 1, \ldots, T$, where $\phi(L) = I_N - A(L) - B(L)$ and $\psi(L) = I_N - B(L)$ are matrix polynomials of the VARMA representation of the 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 the form:

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

Given Assumption 3.2, the VARMA process (10) is invertible and can be written in the VAR form:

$$\Pi(L)\epsilon_t^{(2)} - \omega^* = v_t, \tag{12}$$

for all $t = 1, \ldots, T$, where $\Pi(L) = \psi(L)^{-1}\phi(L) = [I_N - B(L)]^{-1}[I_N - A(L) - B(L)]$ is a matrix polynomial of potentially infinite order of the VAR representation of the GARCH(r,s) process and $\omega^* = \psi(1)^{-1}\omega$ is a constant term. Again, partitioning the vectors, I rewrite (12) in the form:

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

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

Noncausality restrictions. In this paragraph, I present the main theoretical findings of the paper, that is the derivation of the conditions for second-order Granger noncausality for the ECCC-GARCH model. I start by defining two concepts: Granger noncausality in variance and second-order Granger noncausality. Further, I 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) call this concept *second-order Granger causality* and distinguish it from *Granger causality in variance*. I define these noncausalities slightly differently than Comte & Lieberman (2000) do. In the definition for second-order noncausality below, the Hilbert space $I_\epsilon^2(t)$ is used, whereas Comte & Lieberman use $I_y^2(t)$. I formally 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 $y_1 \overset{so}{\nrightarrow} y_2|y_3$, if:

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

Definition 3. y_1 does not Granger-cause y_2 in variance given y_3, denoted by $y_1 \overset{v}{\nrightarrow} y_2|y_3$, if:

$$P\Big([y_{2t+1} - P(y_{2t+1}|I(t))]^{(2)}|I_y^2(t)\Big) = P\Big([y_{2t+1} - P(y_{2t+1}|I_{-1}(t))]^{(2)}|I_{y.-1}^2(t)\Big) \quad \forall t \in \mathbb{Z}, \tag{14}$$

where $[.]^{(2)}$ means that we square every element of a vector. Another difference between the two definitions is in the Hilbert spaces on which y_{2t+1} is projected. On the right-hand side of Definition 2 we take the affine projection of y_{2t+1} on $I(t)$, whereas on the right-hand side of Definition 3 we take the affine projection of y_{2t+1} on $I_{-1}(t)$. In other words, before considering whether there is *second-order Granger noncausality*, one first needs to model and to filter out the Granger causality in mean. Further, an implicit assumption in the definition of *Granger noncausality in variance* is that y_1 does not Granger-cause in mean y_2, $y_1 \overset{G}{\nrightarrow} y_2|y_3$. The

relation between Granger noncausality in mean, noncausality in variance and second-order noncausality have been established by Comte & Lieberman (2000) and are as follows:

$$y_1 \overset{v}{\not\to} y_2|y_3 \Leftrightarrow (y_1 \overset{G}{\not\to} y_2|y_3 \text{ and } y_1 \overset{so}{\not\to} y_2|y_3). \tag{15}$$

One implication of this statement is that Definitions 2 and 3 are equivalent when y_1 does not Granger-cause y_2. And conversely, if y_1 Granger-causes y_2, then the Granger noncausality in variance is excluded, but still y_1 may not second-order cause y_2.

Under Assumptions 1–4, the VARMA-ECCC-GARCH model is stationary, identifiable and invertible in both of its parts: VARMA processes for y_t and for $\epsilon_t^{(2)}$. One more assumption is needed in order to state noncausality relations in the conditional variances process:

Assumption 5. The process ν_t is covariance stationary with covariance matrix $\mathbf{V}_\nu$.

I 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 3–5). Then, y_1 does not second-order Granger-cause y_2 given y_3 (denoted by $y_1 \overset{so}{\not\to} y_2|y_3$) if and only if:*

$$\Pi_{21}(z) \equiv 0 \quad \forall z \in \mathbb{C}. \tag{16}$$

Proof. Theorem 1 may be proved by applying Proposition 1 of Boudjellaba et al. (1992). However, since that proof is derived for the VAR models, several modifications are required to make it applicable to the GARCH model of Jeantheau (1998) in the VAR form, as in (12) and (13). Here I project the squared elements of the residual term, $\epsilon^{(2)}(y_{2t+1}|I(t))$, on the Hilbert spaces $I_\epsilon^2(t)$ or $I_{\epsilon,-1}^2(t)$, both defined in Section 2. □

Theorem 1 is an adaptation of Proposition 1 of Boudjellaba et al. (1992) to the ECCC-GARCH model in the VAR representation for $\epsilon_t^{(2)}$. It sets the conditions for the second-order noncausality between two vectors of variables when in the system there are other auxiliary variables collected in vector y_{3t}. The parametric condition (16), 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 infinite order, when $s > 0$. Therefore, evaluation of the matrix polynomial $\Pi(z)$ is further presented in Theorem 2.

Theorem 2. *Let $\epsilon_t^{(2)}$ follow a stationary vector autoregressive moving average process, as in (10), partitioned, as in (11), which is identifiable and invertible (Assumptions 3–5). Then y_1 does not second-order Granger-cause y_2 given y_3 (denoted by $y_1 \overset{so}{\not\to} y_2|y_3$), if and only if:*

$$\Gamma_{ij}^{so}(z) = \det \begin{bmatrix} \phi_{11}^{\cdot j}(z) & \psi_{11}(z) & \psi_{13}(z) \\ \varphi_{n_1+i,j}(z) & \psi_{21}^{i\cdot}(z) & \psi_{23}^{i\cdot}(z) \\ \phi_{31}^{\cdot j}(z) & \psi_{31}(z) & \psi_{33}(z) \end{bmatrix} = 0 \quad \forall z \in \mathbb{C}, \tag{17}$$

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

Proof. In order to prove the simplified conditions for second-order Granger noncausality, (17), apply to equation (16) from Theorem 1 the matrix transformations of Theorem 3 and then of Theorem 4 of Boudjellaba et al. (1994). □

As was the case for restriction (4), condition (17) leads to $N_1 N_2$ determinant conditions. Each of them can be represented in a form of polynomial in z of degree $\max(r,s) + (N_1 + N_3)s$: $\Gamma_{ij}^{so}(z) = \sum_{i=1}^{\max(r,s)+(N_1+N_3)s} b_i z^i$, where b_i are nonlinear functions of parameters of the GARCH process. We obtain parameter restrictions for the hypothesis of second-order Granger noncausality by setting $b_i = 0$ for $i = 1, \ldots, \max(r,s) + (N_1 + N_3)s$. Such restrictions are ready to be tested.

The innovation of condition (17) 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}. Such a setting has not been considered so far in the problem of testing the second-order noncausality. The restrictions can even be used for large systems of variables. In the Granger-causality analysis, it is particularly important to consider a sufficiently large set of variables. Sims (1980), on the example of the vector moving average model, shows that the Granger causal relation may appear in the model due to the omitted variables problem. Further, Lütkepohl (1982) shows that because of the omitted variables problem a noncausality relation may arrive. The conclusions of these two papers are maintained for the second-order causality analysis in multivariate GARCH models: one should consider a sufficiently large set of relevant variables in order to avoid the omitted variables bias problem.

Condition (17) generalizes results from other studies. Comte & Lieberman (2000) derive similar restriction for the BEKK-GARCH model, with the difference that vector y_t is partitioned only into two sub-vectors y_{1t}^* and y_{2t}^*. Woźniak (2012) does the same for the 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 $y_{1t}^* \overset{so}{\nrightarrow} y_{2t}^*$ (which can be written as $y_{1t} \overset{so}{\nrightarrow} (y_{2t}, y_{3t})$) does not imply that $y_{1t} \overset{so}{\nrightarrow} y_{2t}|y_{3t}$ or that $y_{1t} \overset{so}{\nrightarrow} y_{3t}|y_{2t}$. Moreover, the results of Woźniak (2012) are nested in condition (17) by setting $N_3 = 0$.

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

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

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

$$
= \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} + \begin{bmatrix} 1 - B_{11}L & -B_{12}L & -B_{13}L \\ -B_{21}L & 1 - B_{22}L & -B_{23}L \\ -B_{31}L & -B_{32}L & 1 - B_{33}L \end{bmatrix} \begin{bmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{bmatrix}. \quad (18)
$$

If one is interested in testing the hypothesis $y_1 \overset{so}{\nrightarrow} y_2|y_3$, then by applying Theorem 2 one obtains the following set of restrictions:

$$
\mathbf{R}_1^{III}(\theta) = A_{11}(B_{23}B_{31} - B_{21}B_{33}) + A_{31}(B_{13}B_{21} - B_{11}B_{23}) = 0, \quad (19a)
$$

$$
\mathbf{R}_2^{III}(\theta) = A_{11}B_{21} + A_{31}B_{23} = 0, \quad (19b)
$$

$$
\mathbf{R}_3^{III}(\theta) = A_{21} = 0. \quad (19c)
$$

If one is interested in testing the hypothesis $y_1 \overset{so}{\nrightarrow} (y_2, y_3)$, 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 the restrictions:

$$
\mathbf{R}_1^{IV}(\theta) = A_{11}B_{21} = 0 \quad \text{and} \quad \mathbf{R}_2^{IV}(\theta) = A_{21} = 0, \quad (20a)
$$

$$
\mathbf{R}_3^{IV}(\theta) = A_{11}B_{31} = 0 \quad \text{and} \quad \mathbf{R}_4^{IV}(\theta) = A_{31} = 0. \quad (20b)
$$

Example 4. Let $\epsilon_t^{(2)}$ follow a $N = 3$ dimensional ARCH(r) process, and let one be interested whether y_1 second-order Granger-causes y_2 (given y_3). The restrictions for this case are:

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

4. Bayesian testing of noncausality in VARMA-GARCH models

In the following section, the problem of testing restrictions imposed on the original parameters of the VARMA-GARCH model is considered. Apart from deriving separate tests for the Granger causality and second-order Granger causality hypotheses, I propose a joint test of the parametric restrictions from conditions (4) and (17). Thus, not only do I emphasize the role of joint modeling of the transmissions in conditional mean and conditional variance processes, but I also present a complete set of tools for the underlying analysis. Moreover, a Bayesian testing procedure is proposed as a solution for some of the drawbacks of classical tests.

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

$$\mathbf{W}(\hat{\theta}_m) = T\mathbf{R}(\hat{\theta}_m)'[\mathbf{T}(\hat{\theta}_m)'\mathbf{V}(\hat{\theta}_m)\mathbf{T}(\hat{\theta}_m)]^{-1}\mathbf{R}(\hat{\theta}_m), \tag{22}$$

where θ_m is a sub-vector of θ, containing the parameters used in $l_m \times 1$ vector of parametric restrictions $\mathbf{R}(\theta_m)$, $\mathbf{V}(\hat{\theta}_m)$ is the asymptotic covariance matrix of $\sqrt{T}(\hat{\theta}_m - \theta_m)$, and $\mathbf{T}(\hat{\theta}_m)$ is a $m \times l_m$ matrix of partial derivatives of the restrictions with respect to the parameters collected in θ_m:

$$\mathbf{T}(\hat{\theta}_m) = \frac{\partial \mathbf{R}(\theta_m)}{\partial \theta_m}\bigg|_{\theta_m = \hat{\theta}_m}. \tag{23}$$

Under the null hypothesis of Granger noncausality $\mathbf{W}(\hat{\theta}_m)$ has asymptotic $\chi^2(l_m)$ distribution. However, in equation (22) $\mathbf{T}(\theta_m)$ must be of full rank. Otherwise, the asymptotic covariance matrix is singular and the asymptotic distribution is no longer $\chi^2(l_m)$. Boudjellaba et al. (1992), testing the nonlinear restrictions, as in Example (2), show that there are cases when $\mathbf{T}(\theta_m)$ is not of full rank under the null hypothesis. Several works coping with this problem have appeared (Dufour, 1989; Boudjellaba et al., 1992; Lütkepohl & Burda, 1997; Dufour, Pelletier & Renault, 2006), in the context of testing Granger noncausality for conditional mean processes.

Suppose that a $l_n \times 1$ vector θ_n contains the parameters that appear in the restrictions for second-order Granger noncausality for the multivariate GARCH model derived from condition (17). In order to test such restrictions the Wald test can also be used with test statistics $\mathbf{W}(\theta_n)$. Given that $\sqrt{T}(\hat{\theta}_n - \theta_n)$ has asymptotic normal distribution, the test statistic has asymptotic $\chi^2(l_n)$ distribution with l_n degrees of freedom. However, the determinant condition (17) results in several nonlinear restrictions on the parameters. The testing of the nonlinear restrictions leads in the problem with the asymptotic distribution of the Wald statistic. The matrix of partial derivatives of the restrictions with respect to the parameters of the model, (23), may not be of full rank, and thus the asymptotic covariance matrix of the parametric restrictions under the null hypothesis may be singular. The asymptotic distribution of the test statistics in this case is unknown.

In fact, the Wald test was applied to test the restrictions for the second-order noncausality in the BEKK-GARCH models by Comte & Lieberman (2000) and Hafner & Herwartz (2008). The Wald statistics, proposed by Comte & Lieberman and Hafner & Herwartz, is χ^2-distributed, given the asymptotic normality of the QMLE of the parameters of the BEKK-GARCH model – the result established by Comte & Lieberman (2003). The asymptotic distribution of the test statistic, however, could only be obtained due to the simplifying approach taken. The strategy of Comte & Lieberman (2000) and Hafner & Herwartz (2008) is to derive linear zero restrictions on the original parameters of the model, which are a sufficient condition for the restrictions obtained from the determinant condition (corresponding to determinant condition (17) but for the BEKK-GARCH models and with $N_3 = 0$). Among the classical solutions proposed for the problem of testing the Granger noncausality in conditional means, only the modified Wald test of Lütkepohl & Burda

([1997](#)) seems applicable for testing second-order noncausality in the GARCH models. Nevertheless, further research of this topic is required.

For the VARMA-ECCC-GARCH models, Ling & McAleer ([2003](#)) proved that $\sqrt{T}(\hat{\theta}_n - \theta_n)$ has asymptotic normal distribution. For this model, the application of the Wald test meets the same obstacles as for the BEKK-GARCH model, if one is interested in the testing of the original restrictions for the second-order noncausality and not only those representing the sufficient condition.

Moreover, the asymptotic normality of the QMLE of the parameters for the VARMA-ECCC-GARCH models was derived by Ling & McAleer ([2003](#)), under the assumption of the existence of sixth-order moments of y_t. Similar result was obtained by Comte & Lieberman ([2003](#)) for the BEKK-GARCH models, under the assumption of the existence of eighth-order moments. For many financial time series analyzed with multivariate GARCH models, these assumptions may not hold, as such data are often leptokurtic and the existence of higher-order moments is uncertain.

Finally, the joint test of Granger noncausality and second-order Granger noncausality is a simple generalization of the two separate tests. Suppose that θ_{m+n} stacks the parameters from restrictions derived from conditions ([4](#)) and ([17](#)). The Wald test statistics for such a hypothesis is simply $\mathbf{W}(\theta_{m+n})$ and is asymptotically $\chi^2(l_m + l_n)$ distributed, given that matrix $\mathbf{T}(\theta)$ is of full rank. It also inherits the properties and limitations of both of the separate tests.

Bayesian testing. An alternative approach to testing is proposed in this study. First of all, I propose the method of testing the original restrictions on the parameters for the Granger noncausality and the second-order noncausality presented in Sections [2](#) and [3](#). Secondly, the Bayesian procedure presented in the 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 the existence of higher-order moments of time series are relaxed.

In the context of Granger causality testing in time series models, Bayesian methods have been used in several works. Woźniak ([2012](#)) uses Bayes factors and Posteriors Odds Ratios to infer second-order noncausality between two vectors in GARCH models. Droumaguet & Woźniak ([2012](#)) use these tools to make an inference about Granger noncausality in mean and the independence of the hidden Markov process in Markov-switching VARs. Bayesian methods have also been used also in the context of testing exogeneity, a concept related to Granger noncausality . Jarociński & Maćkowiak ([2011](#)) use Savage-Dickey Ratios to test block-exogeneity in Bayesian VARs. Finally, Pajor ([2011](#)) uses Bayes factors to infer exogeneity in models with latent variables, in particular, in multivariate Stochastic Volatility models.

Consider the following set of hypotheses. The null hypothesis, $\mathcal{H}_0$, states that the $l \times 1$ vector of possibly nonlinear functions of parameters, $\mathbf{R}(\theta)$, is set to a vector of zeros. The alternative hypothesis, $\mathcal{H}_1$, states that it is different from a vector of zeros. The considered set of hypotheses is represented by:

$$\mathcal{H}_0 : \mathbf{R}(\theta) = 0,$$
$$\mathcal{H}_1 : \mathbf{R}(\theta) \neq 0.$$

In the context of Granger-causality, the null hypothesis states that y_1 does not cause y_2 (given that there are also other variables in the system collected in y_3). Then the alternative hypothesis states that y_1 causes y_2. The formulation of the hypotheses is general and encompasses Granger noncausality, second-order noncausality and noncausality in variance. In the following part a Bayesian procedure of evaluation of the credibility of the null hypothesis is described.

In the Bayesian approach, a complete model is specified by a prior distribution of the parameters and a likelihood function. The prior distribution, $p(\theta)$, formalizes the knowledge about the parameters that one has before seeing the data, $\mathbf{y}$. The prior beliefs are updated with information from the data that is represented by the likelihood function, $L(\theta; \mathbf{y})$. As a result of the update of the prior beliefs, a posterior distribution of the parameters of the model is obtained. The posterior distribution is proportional to the product of the likelihood function and the prior distribution:

$$p(\theta|\mathbf{y}) \propto L(\theta; \mathbf{y})p(\theta). \tag{24}$$

Given the posterior distribution of the parameters, the posterior distribution of the function $\mathbf{R}(\theta)$ is available, $p(\mathbf{R}(\theta)|\mathbf{y})$. Moreover, every characteristic of this distribution is available as well. For instance, the posterior mean of $\mathbf{R}(\theta)$ is calculated by definition of the expected value by integrating the product of the function and its posterior distribution over the whole parameter space:

$$E[\mathbf{R}(\theta)|\mathbf{y}] = \int_{\theta \in \Theta} \mathbf{R}(\theta)p(\mathbf{R}(\theta)|\mathbf{y})d\theta.$$

In order to compute such an integral, numerical methods need to be employed for the VARMA-GARCH models, as analytical forms are not known.

Let $\{\theta^{(i)}\}_{i=1}^{S_1}$ be a sample of S_1 draws from the posterior distribution $p(\theta|\mathbf{y})$. Then, $\{\mathbf{R}(\theta^{(i)})\}_{i=1}^{S_1}$ appears a sample drawn from the posterior distribution $p(\mathbf{R}(\theta)|\mathbf{y})$. The posterior mean and the posterior covariance matrix of the restrictions are estimated with:

$$\hat{E}[\mathbf{R}(\theta)|\mathbf{y}] = S_1^{-1} \sum_{i=1}^{S_1} \mathbf{R}(\theta^{(i)}), \tag{25}$$

$$\hat{V}[\mathbf{R}(\theta)|\mathbf{y}] = S_1^{-1} \sum_{i=1}^{S_1} \left[\mathbf{R}(\theta^{(i)}) - \hat{E}[\mathbf{R}(\theta)|\mathbf{y}]\right]\left[\mathbf{R}(\theta^{(i)}) - \hat{E}[\mathbf{R}(\theta)|\mathbf{y}]\right]'. \tag{26}$$

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

$$\kappa(\mathbf{R}) = \left[\mathbf{R} - E[\mathbf{R}(\theta)|\mathbf{y}]\right]' V[\mathbf{R}(\theta)|\mathbf{y}]^{-1}\left[\mathbf{R} - E[\mathbf{R}(\theta)|\mathbf{y}]\right], \tag{27}$$

where $\mathbf{R}$ is the argument of the function. In order to distinguish the argument of the function $\mathbf{R} = \mathbf{R}(\theta)$, I use the simplified notation, neglecting the dependence on the vector of parameters. In place of the expected value and the covariance matrix of the vector of restrictions, $E[\mathbf{R}(\theta)|\mathbf{y}]$ and $V[\mathbf{R}(\theta)|\mathbf{y}]$, one should use their estimators, defined in equations (25) and (26).

The function κ 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, $\mathbf{R} - E[\mathbf{R}(\theta)|\mathbf{y}]$, rescaled by the positive definite posterior covariance matrix, $V[\mathbf{R}(\theta)|\mathbf{y}]$. 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. Drawing an analogy to a Wald test, the main problem of the singularity of the asymptotic covariance matrix of the restrictions is resolved by using the posterior covariance matrix. It does not need to be constructed with the delta method and, thus, avoids the potential singularity of the asymptotic covariance matrix. Notice, however, that the function κ is not a test statistic, but a scalar function that summarizes multiple restrictions on the parameters of the model.

Moreover, if $\mathbf{R} - E[\mathbf{R}(\theta)|\mathbf{y}]$ follows a normal density function, then $\kappa(\mathbf{R})$ would have a $\chi^2(l)$ distribution with l degrees of freedom (see e.g. Proposition B.3 (2) of Lütkepohl, 2005, pp. 678). Consider testing only the Granger noncausality in mean in the VAR model, when the covariance matrix of the innovations is assumed to be constant over time and known. Then, assuming a normal likelihood function and a normal conjugate prior distribution leads to a normal posterior distribution of the parameters. This finding still does not guarantee the χ^2-distributed κ function, as the restrictions on the parameters of the model might be nonlinear and contain sums of products of the parameters. Further, in the general setting of this study, in which the VARMA-GARCH models with t-distributed likelihood function are analyzed, the posterior distribution of the parameters of neither the VARMA nor GARCH parts are in the form of some known distributions (see Bauwens & Lubrano, 1998). Therefore, the exact form of the distribution of $\kappa(\mathbf{R})$ is not known either. It is known up to a normalizing constant, as in equation (24). Luckily, using the Monte Carlo Markov Chain methods, the posterior distributions of the parameters of the model, θ, of the restrictions imposed on them, $\mathbf{R}$, as well as of the function $\kappa(\mathbf{R})$, may be easily simulated. I propose to use the posterior distribution of the function κ in order to evaluate the hypothesis of noncausality.

Let $\kappa(\mathbf{0})$ be the value of function κ, evaluated at the vector of zeros, representing the null hypothesis. Then, a negligible part of the posterior probability mass of $\kappa(\mathbf{R})$ attached to the values greater than $\kappa(\mathbf{0})$ is an argument against the null hypothesis. Therefore, the credibility of the null hypothesis can be assessed by computing the posterior probability of the condition $\kappa(\mathbf{R}) > \kappa(\mathbf{0})$:

$$p_0 = Pr\left(\kappa(\mathbf{R}) > \kappa(\mathbf{0})|\mathbf{y}\right) = \int_{\kappa(\mathbf{0})}^{\infty} p(\kappa(\mathbf{R})|\mathbf{y})d\kappa(\mathbf{R}). \tag{28}$$

Estimation of the probability, p_0, has to be performed using numerical integration methods. Let $\{\mathbf{R}^{(i)}\}_{i=S_1+1}^{S_2}$ be a sample of S_2 draws from the stationary posterior distribution $p(\mathbf{R}(\theta)|\mathbf{y})$, where $\mathbf{R}^{(i)} = \mathbf{R}(\theta^{(i)})$. Using the transformation κ of the restrictions $\mathbf{R}$, one obtains a sample of S_2 draws, $\{\kappa(\mathbf{R}^{(i)})\}_{i=1}^{S_2}$, from the posterior distribution, $p(\kappa(\mathbf{R})|\mathbf{y})$. Then the probability, p_0, is simply estimated by the fraction of the draws from the posterior distribution of $\kappa(\mathbf{R})$, for which the inequality $\kappa(\mathbf{R}) > \kappa(\mathbf{0})$ holds:

$$\hat{p}_0 = \frac{\#\{\kappa(\mathbf{R}^{(i)}) > \kappa(\mathbf{0})\}}{S_2}. \tag{29}$$

The probability, p_0, should be compared to a probability, π, that represents a confidence level of the test. The usual values used in many statistical works are 0.05 or 0.1.

The procedure is summarized in four steps:

Step 1 Draw $\{\mathbf{R}(\theta^{(i)})\}_{i=1}^{S_1}$ and compute the estimators of the posterior mean, $\hat{E}[\mathbf{R}(\theta)|\mathbf{y}]$, and the posterior covariance matrix, $\hat{V}[\mathbf{R}(\theta)|\mathbf{y}]$, for the vector of restrictions on the parameters.

Step 2 Draw $\{\kappa(\mathbf{R}^{(i)})\}_{i=S_1+1}^{S_2}$ from the posterior distribution $p(\kappa(\mathbf{R})|y)$, using the estimated posterior mean and covariance matrix from **Step 1** to compute $\kappa(.)$.

Step 3 Compute $\kappa(\mathbf{0})$ and $\hat{p}_0$.

Step 4 If $\hat{p}_0 < \pi_0$, then reject the null hypothesis, $\mathcal{H}_0$. Otherwise, do not reject the null hypothesis.

Discussion. The proposed Bayesian procedure allows testing of the noncausality restrictions resulting directly from the determinant condition (16). There is no need to derive the simplified zero restrictions on the parameters of the model in order to test the noncausality hypothesis, as proposed by Comte & Lieberman (2000) and Hafner & Herwartz (2008). Second, the procedure requires the estimation of only one unrestricted model for the purpose of testing the noncausality hypotheses. Given the time required to estimate the multivariate VARMA-GARCH models, this is a significant gain in comparison to the procedure proposed by Woźniak (2012). He used Bayes factors to test the second-order noncausality hypotheses between two vectors of variables in ECCC-GARCH models. Consequently, his method requires the estimation of multiple models: the unrestricted and the restricted models representing the hypotheses of interest.

Further, the posterior distribution of function κ is a finite sample distribution. Therefore, the test is also based on the exact finite sample distribution. On the contrary, in the classical inference on VARMA-GARCH models only the asymptotic distribution of the QML 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 strict assumptions. As a result, the Bayesian test relaxes the assumptions of the existence of higher-order moments. Only the existence of fourth-order moments is assumed (see Assumption 5), in comparison to the assumption of the existence of sixth-order moments in a classical derivation of the asymptotic distribution of the QMLE (see Ling & McAleer, 2003). Moreover, this testing procedure could be employed for the restrictions of Comte & Lieberman (2000) for testing the noncausality in variance in the BEKK-GARCH models. The asymptotic normality of the QMLE established by Comte & Lieberman (2003) requires the existence of the eighth-order moments, an assumption that can now be relaxed.

These improvements are particularly important in the context of the analysis of financial high-frequency data. Many empirical studies have proved that the empirical distribution of such data is leptokurtic, and that the existence of higher-order moments is questionable. Therefore, the relaxed assumptions may give an advantage on the applicability of the proposed testing procedure over the applicability of classical tests.

Table 1: Data: summary statistics

	CHF/EUR	GBP/EUR	USD/EUR
Mean	-0.034	0.012	-0.006
Median	-0.033	0.011	0.016
Standard Deviation	0.704	0.707	0.819
Minimum	-3.250	-2.657	-4.735
Maximum	7.997	3.461	4.038
Excess kurtosis	25.557	2.430	2.683
Excess kurtosis (robust)	0.785	0.060	0.085
Skewness	2.220	0.344	-0.091
Skewness (robust)	-0.038	0.010	-0.016
LJB test	21784.921	206.525	234.063
LJB p-value	0.000	0.000	0.000
T	777.0	777.0	777.0
Correlations	.	.	.
GBP/EUR	0.079	.	.
USD/EUR	0.301	0.368	.

Note: The excess kurtosis (robust) and the skewness (robust) coefficients are outlier-robust versions of the excess kurtosis and the skewness coefficients as described in Kim & White (2004). LJB test and LJB p-values describe the test of normality by Lomnicki (1961) and Jarque & Bera (1980).

5. Granger causal analysis of exchange rates

Data. I illustrate the use of the methods with three time series of daily exchange rates. The series, all denominated in Euro, are the Swiss franc (CHF/EUR), the British pound (GBP/EUR) and the United States dollar (USD/EUR). I analyze the logarithmic rates of return expressed in percentage points, $y_{it} = 100(\ln x_{it} - \ln x_{it-1})$ for $i = 1, 2, 3$, where x_{it} are levels of the assets. The data spans the period from September 16, 2008 to September 22, 2011, which gives $T = 777$ observations. It was downloaded from the European Central Bank website (http://sdw.ecb.int/browse.do?node=2018794). The analyzed period starts the day after Lehman Brothers filed for Chapter 11 bankruptcy protection.

The motivation behind this choice of variables and the period of analysis is its usefulness for the institutions for which the forecast of the exchange rates is a crucial element of financial planning. For instance, suppose that the government of a country participating in the Eurozone is indebted in currencies, and therefore its future public debt depends on the exchange rates. Or, suppose that a financial institution settled in the Eurozone keeps assets bought on the New York or London stock exchanges, or simply keeps currencies. In these and many other examples, the performance of an institution depends on the forecast of the returns, but even more important is the forecast of the future volatility of exchange rates. The knowledge that the past information about one exchange rate has an impact on the forecast of the variability of some other exchange rate may be crucial for the analysis of the risk of a portfolio of assets. The two exchange rates, GBP/EUR and USD/EUR, were analyzed for the same period in Woźniak (2012).

Figure A.1 from Appendix A plots the three time series. The clustering of the volatility of the data is evident. Two of the exchange rates, GBP/EUR and USD/EUR, during the first year of the sample period were characterized by higher volatility than in the subsequent years. The Swiss franc is characterized by more periods of different volatility. The first year of high variability was followed by nearly a year of low volatility. After that period, again there was a period of high volatility. As the volatility clustering seems to be present in the data, the GARCH models that are capable of modeling this feature are chosen for the subsequent analysis.

Table 1 presents the summary statistics of the time series. While the sample means are very similar

to one another and close to zero, the variability measured with the sample standard deviation seems to be a bit higher for the US dollar than for other currencies. All the series are leptokurtic, as the kurtosis coefficients are high. This is especially the case for the Swiss franc. The Swiss franc and the British pound are positively skewed. None of the variables follows a normal distribution, as shown by the results of the Lomnicki-Jarque-Bera test. These findings support the use of the *t*-distributed likelihood function. Probably some asymmetric *t* distribution could capture the skewness of the distribution. For the sake of simplicity, however, I assume a symmetric distribution.

Estimation of the model. The Bayesian estimation of the VARMA-GARCH models consists of the numerical simulation of the posterior distribution of the parameters, which is proportional to the product of the likelihood function and the prior distribution of the parameters of the model, as in equation (24).

For the parameters of the VAR-GARCH model, I assume the following prior specification. For the parameters of the vector autoregressive process of order one and of the GARCH(1,1) model, I assume the prior distribution proportional to a constant and constrained to a parameter space bounded according to Assumptions 1–5. Each of the parameters of the correlation matrix, **C**, collected in a $N(N-1)/2 \times 1$ vector $\rho = \text{vecl}(\mathbf{C})$, follows a uniform distribution on the interval $[-1, 1]$, where a *vecl* operator stacks lower-diagonal elements of a matrix in a vector. Finally, for the degrees of freedom parameter, I assume the prior distribution proposed by Deschamps (2006). Such a prior specification, with diffuse distributions for all the parameters but the degrees of freedom parameter v, guarantees the existence of the posterior distribution (see Bauwens & Lubrano, 1998). It does not discriminate any of the values of the parameters from within the parameter space. The prior distribution for the parameter v is a proper density function, and it gives as much as a 32 percent chance that its value is greater than 30. For such values of this parameter, the likelihood function given by equation (9b), is a close approximation of the normal likelihood function.

Summarizing, the prior specification for the considered model has the detailed form of:

$$p(\theta) = p\left(\alpha_0', \text{vec}(\alpha_1)\right) p\left(\omega', \text{vec}(A)', \text{vec}(B)'\right) p(v) \prod_{i=1}^{N(N-1)/2} p(\rho_i), \tag{30}$$

where each prior distribution is specified by:

$$p\left(\alpha_0', \text{vec}(\alpha_1)'\right)' \propto \mathcal{I}(\theta \in \Theta)$$
$$p\left(\omega', \text{vec}(A)', \text{vec}(B)'\right)' \propto \mathcal{I}(\theta \in \Theta)$$
$$v \sim .04 \exp\left[-.04(v-2)\right] \mathcal{I}(v \geq 2)$$
$$\rho_i \sim \mathcal{U}(-1, 1) \quad \text{for } i = 1, \dots, N(N-1)/2,$$

where $\mathcal{I}(.)$ is an indictor function, taking a value equal to 1 if the condition in brackets holds and 0 otherwise.

The kernel of the posterior distribution of the parameters of the model, given by equation (24), is a complicated function of the parameters. It is not given by the kernel of any known distribution function. In consequence, the analytical forms are known neither for the posterior distribution nor for full conditional distributions. Therefore, numerical methods need to be employed in order to simulate the posterior distribution. I use the Metropolis-Hastings algorithm adapted for the GARCH models by Vrontos et al. (2003). At each sth step of the algorithm, a candidate draw, θ^*, is made from the candidate density. The candidate generating density is a multivariate *t* distribution with the location parameter set to the previous state of the Markov chain, $\theta^{(s-1)}$, the scale matrix $c\Omega$ and the degrees of freedom parameter set to five. The scale matrix, Ω, should be a close approximation of the posterior covariance matrix of the parameters, and a constant c is set in order to obtain the desirable acceptance rate of the candidate draws. A new candidate draw, θ^*, is accepted with the probability:

$$\alpha\left(\theta^{(s-1)}, \theta^*|\mathbf{y}\right) = \min\left[1, \frac{L(\theta^*; \mathbf{y})p(\theta^*)}{L(\theta^{(s-1)}; \mathbf{y})p(\theta^{(s-1)})}\right].$$

Table 2: Summary of the estimation of the VAR(1)-ECCC-GARCH(1,1) model

	VAR(1)			
	α_0		α_1	
CHF/EUR	-0.022	-0.068	0.003	0.041
	(0.011)	(0.037)	(0.019)	(0.018)
GBP/EUR	0.014	-0.016	0.077	-0.019
	(0.021)	(0.030)	(0.040)	(0.033)
USD/EUR	0.027	-0.027	0.050	0.006
	(0.025)	(0.041)	(0.045)	(0.041)

	GARCH(1,1)						
	ω		A			B	
CHF/EUR	0.001	0.117	0.002	0.002	0.873	0.001	0.001
	(0.001)	(0.029)	(0.002)	(0.002)	(0.030)	(0.001)	(0.001)
GBP/EUR	0.011	0.002	0.062	0.017	0.002	0.808	0.063
	(0.009)	(0.002)	(0.024)	(0.011)	(0.003)	(0.158)	(0.098)
USD/EUR	0.086	0.018	0.117	0.051	0.031	0.787	0.164
	(0.059)	(0.018)	(0.062)	(0.034)	(0.028)	(0.215)	(0.147)

Degrees of freedom and correlations			
ν	ρ_{12}	ρ_{13}	ρ_{23}
6.267	0.145	0.356	0.400
(0.743)	(0.038)	(0.035)	(0.034)

The table summarises the estimation of the VAR(1)-ECCC-GARCH(1,1) model described by the equations (2), (7), (8) and the likelihood function (9b). The prior distributions are specified in equation (30). The posterior means and the posterior standard deviations (in brackets) are reported. For graphs of the marginal posterior distributions of the parameters, as well as for the summary of characteristics of the MCMC simulation of the posterior distribution, refer to Appendix B.

I kept every 100th state of the Markov Chain in the final sample of draws from the posterior distribution of the parameters. The rationale behind this strategy is that, at the cost of decreasing the sample size, I obtain the sample of desirable properties according to several criteria (see Geweke, 1989, 1992; Plummer, Best, Cowles & Vines, 2006). The summary of the properties of the final sample of draws from the posterior distribution is presented in Table B.6 in Appendix B.

Estimation results. Table 2 presents the results of the posterior estimation of the VAR(1)- ECCC-GARCH(1,1) model chosen for the analysis of causality relations in the system of three exchange rates: CHF/EUR, GBP/EUR and USD/EUR. Plots of marginal posterior densities of the parameters are presented in Appendix B.

Considering posterior means and standard deviations of the parameters of the VAR(1) process, one sees that none of the parameters but $\alpha_{1.13}$ is significantly different from zero. The graphs, however, show that the 90 percent highest posterior density regions of parameters $\alpha_{0.1}$, $\alpha_{1.11}$, $\alpha_{1.22}$ and $\alpha_{1.13}$ do not contain the value zero. The parameter $\alpha_{1.13}$ is responsible for the interaction of the lagged value for US dollar on the current value of the Swiss frank. This finding has its consequences in testing the Granger causality in mean hypothesis.

All the parameters of the GARCH(1,1) process are constrained to be non-negative. However, a significant part of the posterior probability mass concentrated at the bound given by zero is an argument for a lack of the statistical significance of the parameter. For most of the parameters of the GARCH process reported in Table 2, this is the case (see also graphs in Appendix B). The posterior probability mass of several of the parameters, however, is distant from zero. All the diagonal parameters of matrices A and Bm, beside parameters A_{33} and B_{33}, are different from zero. This finding is common for multivariate GARCH models

and reflects the persistence of volatility.

Nevertheless, it is the value of the posterior mean of parameter B_{32} equal to 0.787 that is interesting in this model. This parameter models the impact of the lagged conditional variance of British pound on the current conditional variance of the US dollar. This effect is significant. Moreover, estimates of the parameters for the system of variables that would include only GBP/EUR and USD/EUR are very similar to the values of the parameters of the bivariate VAR-ECCC-GARCH model estimated by Woźniak (2012) for the same period. To conclude, the estimate of this parameter in particular may be considered robust to including an additional variable to the model, namely the CHF/EUR, as well as to the prior distribution specification. Woźniak (2012) estimates two models with a truncated-normally distributed priors with two different variance parameters: 100 and 0.1.

Finally, Figure B.7 proves that the parameter of the degrees of freedom, ν, of the t-distributed residuals cannot be considered greater than 6. This value lies in the high posterior probability mass of this parameter. Therefore, the existence of moments of order 6 and higher of the error term is questionable. In effect, classical testing of the VARMA-ECCC-GARCH model has limited use in this case. This statement is justified by the requirement of the existence of sixth-order moments for the asymptotic normality of the QML estimator (see Ling & McAleer, 2003).

Table 3: Results of testing: Granger causality hypothesis

$\mathcal{H}_0$:	$\kappa(\mathbf{0})$	$\hat{p}_0$	Figure Ref.
$y_1 \overset{G}{\nrightarrow} y_2 \vert y_3$	0.294	0.586	C.9.1
$y_1 \overset{G}{\nrightarrow} y_3 \vert y_2$	0.427	0.522	C.9.2
$y_2 \overset{G}{\nrightarrow} y_1 \vert y_3$	0.022	0.883	C.9.3
$y_2 \overset{G}{\nrightarrow} y_3 \vert y_1$	1.226	0.270	C.9.4
$y_3 \overset{G}{\nrightarrow} y_1 \vert y_2$	5.013	0.023	C.9.5
$y_3 \overset{G}{\nrightarrow} y_2 \vert y_1$	0.336	0.561	C.9.6
$(y_1, y_2) \overset{G}{\nrightarrow} y_3$	1.580	0.455	C.10.1
$(y_1, y_3) \overset{G}{\nrightarrow} y_2$	0.884	0.642	C.10.2
$(y_2, y_3) \overset{G}{\nrightarrow} y_1$	5.520	0.063	C.10.3
$y_1 \overset{G}{\nrightarrow} (y_2, y_3)$	0.530	0.765	C.10.4
$y_2 \overset{G}{\nrightarrow} (y_1, y_3)$	1.252	0.543	C.10.5
$y_3 \overset{G}{\nrightarrow} (y_1, y_2)$	5.776	0.059	C.10.6
$y_1 \overset{G}{\nrightarrow} y_2 \vert y_3$ & $y_2 \overset{G}{\nrightarrow} y_1 \vert y_3$	0.315	0.858	C.11.1
$y_1 \overset{G}{\nrightarrow} y_3 \vert y_2$ & $y_3 \overset{G}{\nrightarrow} y_1 \vert y_2$	5.249	0.072	C.11.2
$y_2 \overset{G}{\nrightarrow} y_3 \vert y_1$ & $y_3 \overset{G}{\nrightarrow} y_2 \vert y_1$	1.402	0.490	C.11.3

Note: The table presents the considered null hypotheses, $\mathcal{H}_0$, of Granger noncausality, as in Definition 1. The values of function κ associated with the null hypotheses, $\kappa(\mathbf{0})$, are reported in the second column. $\hat{p}_0$ is the posterior probability of the condition for not rejecting the null hypothesis, as defined in (28). For a graphical presentation of the posterior densities of $\kappa(\mathbf{R})$ and the values $\kappa(\mathbf{0})$, see the figure references given in the last column. The figures may be found in Appendix C.
Description of the variables: y_1 = CHF/EUR, y_2 = GBP/EUR, y_3 = USD/EUR.

Granger-causality testing results. Table 3 presents the results of the Granger noncausality in mean testing. The values of $\kappa(\mathbf{0})$ and of the estimate of the probability p_0 are reported. Appendix C presents plots of the posterior distribution of $\kappa(\mathbf{R})$ for each of the hypotheses.

Table 4: Results of testing: second-order Granger causality hypothesis

$\mathcal{H}_0:$	$\kappa(\mathbf{0})$	$\hat{p}_0$	Figure Ref.
$y_1 \overset{so}{\nrightarrow} y_2\|y_3$	1.521	0.279	C.12.1
$y_1 \overset{so}{\nrightarrow} y_3\|y_2$	3.120	0.221	C.12.2
$y_2 \overset{so}{\nrightarrow} y_1\|y_3$	2.550	0.247	C.12.3
$y_2 \overset{so}{\nrightarrow} y_3\|y_1$	9.811	0.039	C.12.4
$y_3 \overset{so}{\nrightarrow} y_1\|y_2$	2.379	0.231	C.12.5
$y_3 \overset{so}{\nrightarrow} y_2\|y_1$	2.354	0.193	C.12.6
$(y_1, y_2) \overset{so}{\nrightarrow} y_3$	11.539	0.113	C.13.1
$(y_1, y_3) \overset{so}{\nrightarrow} y_2$	2.733	0.372	C.13.2
$(y_2, y_3) \overset{so}{\nrightarrow} y_1$	3.926	0.386	C.13.3
$y_1 \overset{so}{\nrightarrow} (y_2, y_3)$	3.491	0.246	C.13.4
$y_2 \overset{so}{\nrightarrow} (y_1, y_3)$	10.714	0.061	C.13.5
$y_3 \overset{so}{\nrightarrow} (y_1, y_2)$	4.084	0.227	C.13.6
$y_1 \overset{so}{\nrightarrow} y_2\|y_3 \ \& \ y_2 \overset{so}{\nrightarrow} y_1\|y_3$	3.481	0.386	C.14.1
$y_1 \overset{so}{\nrightarrow} y_3\|y_2 \ \& \ y_3 \overset{so}{\nrightarrow} y_1\|y_2$	4.741	0.324	C.14.2
$y_2 \overset{so}{\nrightarrow} y_3\|y_1 \ \& \ y_3 \overset{so}{\nrightarrow} y_2\|y_1$	11.633	0.099	C.14.3

Note: The table presents the considered null hypotheses of second-order Granger causality, as in Definition 2. For a description of the notation, see the note to Table 3.

Description of the variables: y_1 = CHF/EUR, y_2 = GBP/EUR, y_3 = USD/EUR.

Only a few of the hypotheses of noncausality in mean are rejected at the confidence levels 0.05 or 0.1. All the rejected hypotheses relate to two of the exchange rates: CHF/EUR and USD/EUR. First, the US dollar has a significant effect on the Swiss frank, a result established at the level of confidence equal to 0.05. These two exchange rates impact on each other as well. Further, the US dollar has a significant effect on the Swiss frank and the British pound taken jointly. Finally, the frank is significantly affected by both the pound and the dollar, taken jointly. All the three last results are established at the level of confidence 0.1.

Consideration of the results of testing the second-order noncausality hypotheses, reported in Table 4, brings new findings. The pattern of connections between the exchange rates is different for second-order causality than for causality in mean. The rejected hypotheses of second-order noncausality relate to the British pound and the US dollar. Information about the history of volatility of GBP/EUR has a significant effect on the current conditional variance of variable USD/EUR at the level of confidence equal to 0.05. It has also a significant effect on CHF/EUR and USD/EUR, taken jointly at the level of confidence equal to 0.1. The same conclusions are found in Woźniak (2012). This finding is particularly interesting, as Woźniak uses Bayes factors and Posterior probabilities in order to assess the hypotheses. These conclusions are, therefore, robust to the choice of the testing procedure.

The following interpretation of the testing results of second-order noncausality hypothesis is proposed. The Swiss frank does not have any significant effect on the volatility of the British pound or the US dollar, which proves its minor role in modeling volatility in comparison to the other two exchange rates. The impact of the pound-to-Euro exchange rate on the volatility of the dollar-to-Euro exchange rate, is most probably related to the the *meteor showers* hypothesis of Engle et al. (1990). The proper conclusion seems to be that the spillovers in volatility are due to the activity of traders on the exchange rates market. Although the market is open 24 hours a day, there exist periods of higher activity of trading of particular currencies. Therefore, the behavior of traders in Europe, reflected in the exchange rate prices and their volatility, affects the decisions of traders in North America. Such a pattern can be captured by the dataset and the model

Table 5: Results of testing: Granger causality in variance hypothesis

$\mathcal{H}_0:$	$\kappa(\mathbf{0})$	$\hat{p}_0$	Figure Ref.
$y_1 \overset{V}{\nrightarrow} y_2\|y_3$	1.755	0.502	C.15.1
$y_1 \overset{V}{\nrightarrow} y_3\|y_2$	3.791	0.297	C.15.2
$y_2 \overset{V}{\nrightarrow} y_1\|y_3$	2.591	0.430	C.15.3
$y_2 \overset{V}{\nrightarrow} y_3\|y_1$	11.177	0.040	C.15.4
$y_3 \overset{V}{\nrightarrow} y_1\|y_2$	8.007	0.095	C.15.5
$y_3 \overset{V}{\nrightarrow} y_2\|y_1$	2.795	0.324	C.15.6
$(y_1, y_2) \overset{V}{\nrightarrow} y_3$	13.412	0.128	C.16.1
$(y_1, y_3) \overset{V}{\nrightarrow} y_2$	3.728	0.545	C.16.2
$(y_2, y_3) \overset{V}{\nrightarrow} y_1$	10.350	0.205	C.16.3
$y_1 \overset{V}{\nrightarrow} (y_2, y_3)$	4.229	0.440	C.16.4
$y_2 \overset{V}{\nrightarrow} (y_1, y_3)$	12.098	0.083	C.16.5
$y_3 \overset{V}{\nrightarrow} (y_1, y_2)$	10.599	0.118	C.16.6
$y_1 \overset{V}{\nrightarrow} y_2\|y_3$ & $y_2 \overset{V}{\nrightarrow} y_1\|y_3$	3.826	0.629	C.17.1
$y_1 \overset{V}{\nrightarrow} y_3\|y_2$ & $y_3 \overset{V}{\nrightarrow} y_1\|y_2$	10.900	0.184	C.17.2
$y_2 \overset{V}{\nrightarrow} y_3\|y_1$ & $y_3 \overset{V}{\nrightarrow} y_2\|y_1$	13.310	0.119	C.17.3

Note: The table presents the considered null hypotheses of Granger causality in variance, as in Definition 3. For a description of the notation, see the note to Table 3.

Description of the variables: y_1 = CHF/EUR, y_2 = GBP/EUR, y_3 = USD/EUR.

considered in this study.

One more hypothesis is rejected at the confidence level equal to 0.1: the pound is found to second-order cause the dollar, *and* the dollar second-order causes pound, which is mainly driven by parameter B_{32}.

Finally, the results of testing hypotheses of noncausality in variance are reported in Table 5. These results are not just a simple intersection of the results for Granger-causality in mean and second-order noncausality testing, as one could deduce from equation (15). The marginal distribution of the parameters of the VAR process is not independent of the marginal distribution of the parameters of the GARCH process. The posterior covariance matrix is not block-diagonal. Therefore, the results of noncausality in variance should be discussed separately. One of the hypotheses is rejected at the confidence level equal to 0.05: the hypothesis of noncausality in variance from the British pound to the US dollar. The other three hypotheses are rejected at the confidence level equal to 0.1. The following relations are found: the dollar causes the frank in variance; and the pound causes the frank and the dollar in variance, taken jointly.

6. Conclusions

This study first of all proposes the parameter restrictions for second-order noncausality between two vectors of variables, when there are also other variables in the considered system used for modeling and forecasting. The derivations are made within the framework of the popular VARMA-GARCH model. The novelty of these conditions is that, contrary to the developments of Comte & Lieberman (2000) and Woźniak (2012), they allow the finding of restrictions for a hypothesis of noncausality between chosen variables from the system. The two cited works use a setting in which all the variables are split into two vectors, which imposes a kind of a rigidity in forming hypotheses.

The conditions may result in several nonlinear restrictions on the parameters of the model, which results in a conclusion that the available classical tests have limited use. As a solution to this testing problem, I propose the Bayesian procedure based on the posterior distribution of a function summarizing all the restrictions. This procedure allows for testing of the hypotheses of Granger noncausality in mean and second-order noncausality jointly, forming a hypothesis of noncausality in variance as well as separately. The procedure requires the estimation of only one model, the unrestricted. This fact is an improvement, in comparison to the procedure proposed by Woźniak (2012), which required the estimation of several models representing different hypotheses. Further, the restrictions of the existence of the higher-order moments of the processes required in the classical tests are relaxed. Similarly to the test of Woźniak (2012), the existence of fourth-order moments is required in the proposed analysis, whereas the asymptotic derivations of Ling & McAleer (2003) require the existence of the sixth-order moments for the VARMA-GARCH models.

The main limitation of the noncausality analysis in this work, is that the conditions only for one-period-ahead noncausality are presented. In the works of Comte & Lieberman (2000) and Woźniak (2012), due to the specific setting of the vectors of variables from the system, these conditions imply noncausality at all the future horizons. In this work, however, when the third vector of variables, y_3, is non-empty, then the conditions from Theorem 2 are useful only for the analysis one period ahead.

This limitation forms a motivation for future research that would aim at derivation of the restrictions for h-period-ahead noncausality within the flexible framework of splitting the variables into three vectors, and where $h = 1, 2, 3, \ldots$. Such conditions would be informative of the non-direct causality that is, a situation in which, despite the fact that one variable does not Granger-cause the other one period ahead, it may still be causal several periods ahead through the channel of the third variable (see Dufour et al., 2006).

Another direction of possible research is a derivation of the conditions for second-order noncausality for GARCH models, when the data have specific features. Some financial data are proven to have persistent volatility that is modeled with integrated GARCH processes. Such processes are defined by the fact that the polynomial $|I_N - A(z) - B(z)| = 0$ has a unit root. This case is excluded from the analysis in this study. Further, the analysis of some financial time series conducted by Diebold & Yilmaz (2009), has proved that the values of financial assets as well as their volatility spillover at different rates in different periods. This finding might result in the parameters of the GARCH process changing values over time. Such nonlinearities may be modeled, e.g. with the GARCH processes with a regime change, or when the parameters change their values according to a latent hidden Markov process, as in the Markov-switching models. For such data, the analysis of Granger causality is of interest as well.

Appendix A. Data

Figure A.1: Data plot: (CHF/EUR, GBP/EUR, USD/EUR)
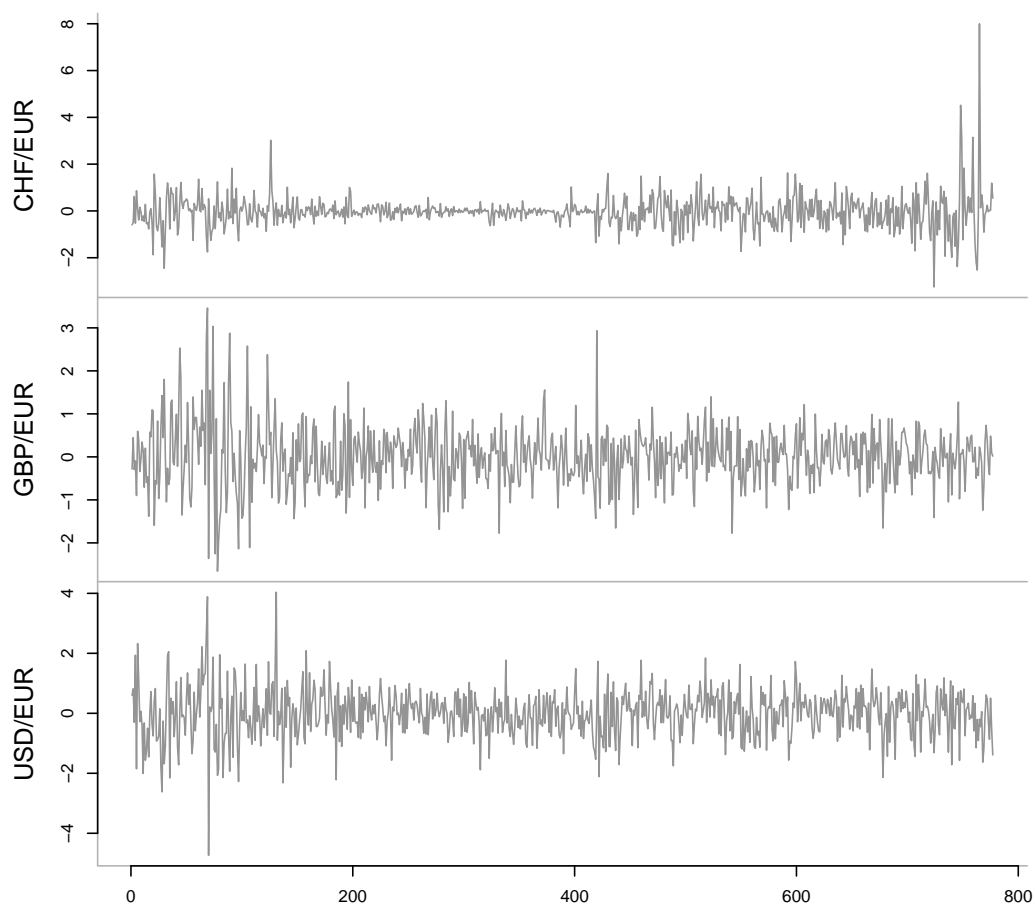
The graph presents daily logarithmic rates of return expressed in percentage points: $y_{it} = 100(\ln x_{it} - \ln x_{it-1})$, for $i = 1, 2, 3$, where x_{it} denotes the level of an asset of three exchange rates: the Swiss franc, the British pound and the US dollar, all denominated in Euro. The data spans the period from September 16, 2008 to September 22, 2011, which gives $T = 777$ observations. The data was downloaded from the European Central Bank website (http://sdw.ecb.int/browse.do?node=2018794).

Appendix B. Summary of the posterior density simulation

Note: Figures B.2–B.8 present the marginal posterior distribution of the parameters with the 95% and 90% highest posterior density regions represented by light-grey and dark-grey areas respectively.

Figure B.2: Summary of the posterior distribution: α_0
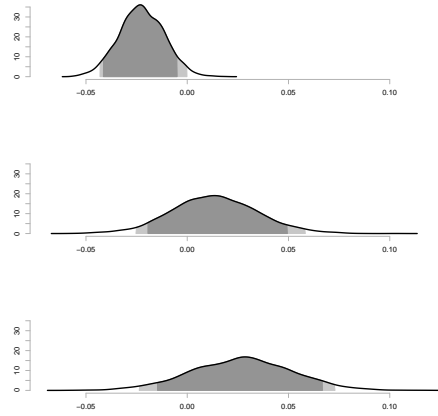
Figure B.3: Summary of the posterior distribution: α_1
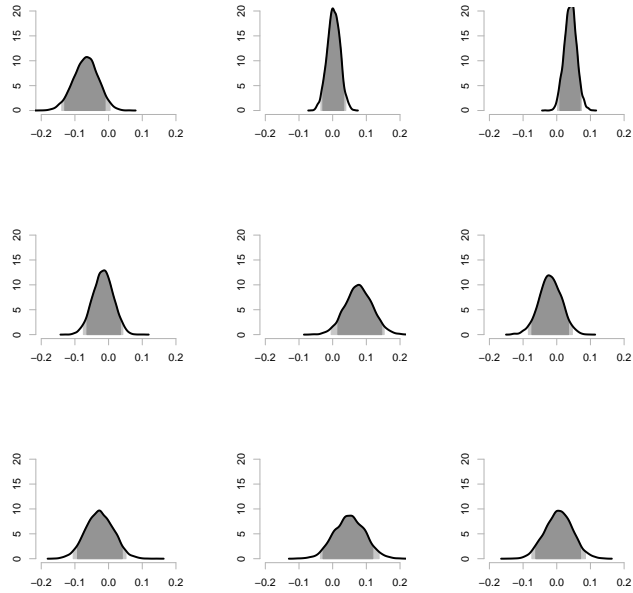
Figure B.4: Summary of the posterior distribution: ω
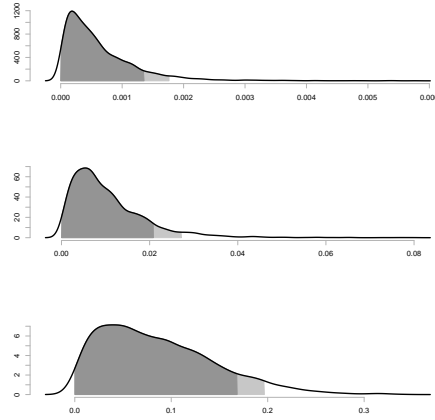
Figure B.5: Summary of the posterior distribution: A
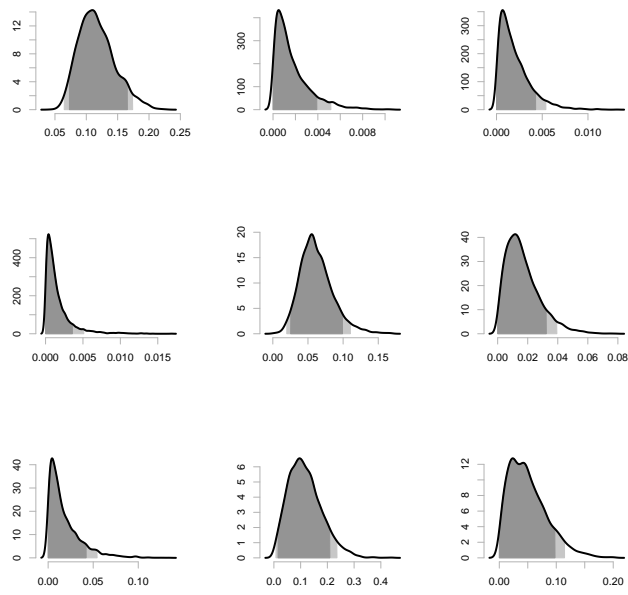
Figure B.6: Summary of the posterior distribution: B
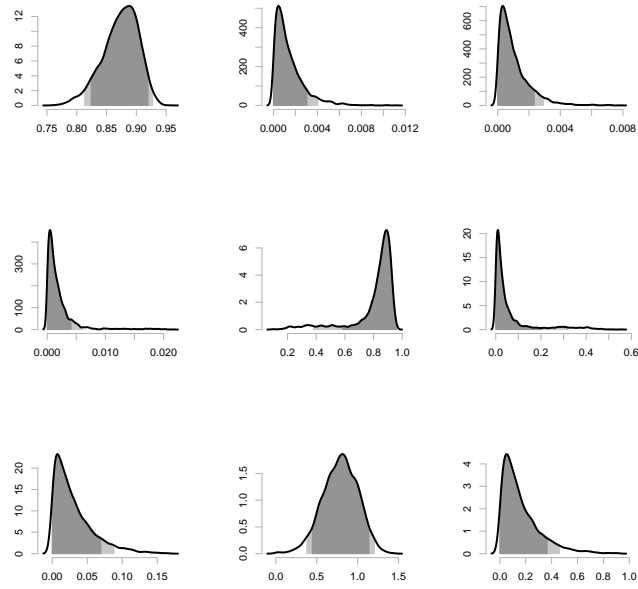
Figure B.7: Summary of the posterior distribution: ν
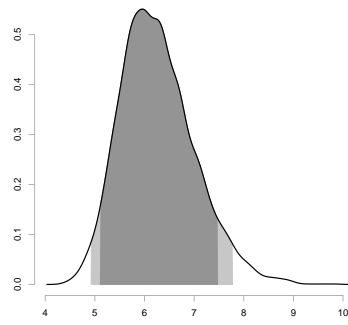
Figure B.8: Summary of the posterior distribution: ρ
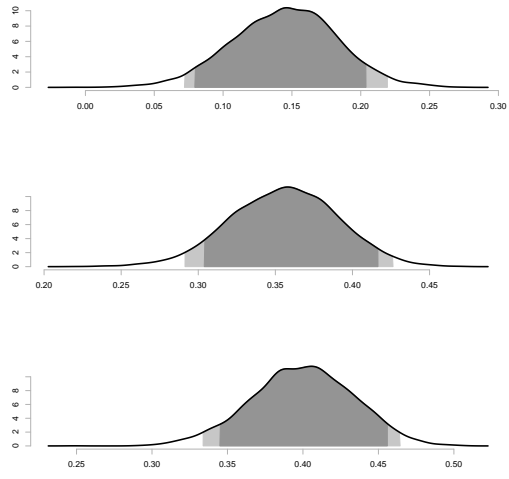
Table B.6: Summary of the posterior distribution simulation

		Mean	SD	Autocorrelations at lag 1	Autocorrelations at lag 50	RNE	Geweke's z
	Vector Autoregression						
α_0	$\alpha_{0.1}$	-0.022	0.011	0.316	-0.012	0.554	-0.475
	$\alpha_{0.2}$	0.014	0.021	0.341	-0.017	0.436	-0.089
	$\alpha_{0.3}$	0.027	0.025	0.335	-0.024	0.355	1.701
	$\alpha_{1.11}$	-0.068	0.037	0.338	0.001	0.429	0.568
	$\alpha_{1.21}$	-0.016	0.030	0.403	0.015	0.303	0.143
	$\alpha_{1.31}$	-0.027	0.041	0.378	-0.007	0.315	-0.658
	$\alpha_{1.12}$	0.003	0.019	0.290	0.002	0.322	1.350
α_1	$\alpha_{1.22}$	0.077	0.040	0.330	-0.012	0.435	0.166
	$\alpha_{1.32}$	0.050	0.045	0.233	0.022	0.511	-0.982
	$\alpha_{1.13}$	0.041	0.018	0.338	-0.008	0.415	-0.467
	$\alpha_{1.23}$	-0.019	0.033	0.371	-0.020	0.284	-2.032
	$\alpha_{1.33}$	0.006	0.041	0.347	0.013	0.404	0.471
	GARCH(1,1)						
	ω_1	0.001	0.001	0.677	0.032	0.108	-0.773
ω	ω_2	0.011	0.009	0.600	0.023	0.082	-1.154
	ω_3	0.086	0.059	0.392	0.058	0.145	1.818
	A_{11}	0.117	0.029	0.460	-0.019	0.186	0.176
	A_{21}	0.002	0.002	0.639	0.066	0.058	-1.143
	A_{31}	0.018	0.018	0.581	0.042	0.109	0.987
	A_{12}	0.002	0.002	0.630	-0.055	0.085	-0.364
A	A_{22}	0.062	0.024	0.575	0.005	0.134	0.606
	A_{32}	0.117	0.062	0.258	-0.001	0.468	-0.074
	A_{13}	0.002	0.002	0.533	-0.054	0.173	0.609
	A_{23}	0.017	0.011	0.483	0.058	0.121	-1.442
	A_{33}	0.051	0.034	0.425	-0.004	0.226	1.117
	B_{11}	0.873	0.030	0.480	-0.023	0.171	-0.021
	B_{21}	0.002	0.003	0.878	0.317	0.020	-0.731
	B_{31}	0.031	0.028	0.207	0.031	0.141	1.718
	B_{12}	0.001	0.001	0.619	-0.039	0.171	-0.354
B	B_{22}	0.808	0.158	0.956	0.569	0.015	2.401
	B_{32}	0.787	0.215	0.518	-0.006	0.122	-0.438
	B_{13}	0.001	0.001	0.575	-0.034	0.277	2.530
	B_{23}	0.063	0.098	0.968	0.585	0.015	-2.357
	B_{33}	0.164	0.147	0.659	0.056	0.064	-0.885
	Degrees of freedom and correlations						
	ν	6.267	0.743	0.415	0.017	0.447	0.123
	ρ_{12}	0.145	0.038	0.335	0.036	0.481	1.444
$\mathbf{C}$	ρ_{13}	0.356	0.035	0.359	-0.021	0.481	2.820
	ρ_{23}	0.400	0.034	0.328	-0.006	0.501	0.048

Note: The table reports posterior means and posterior standard deviations of the parameters of the model. Also, autocorrelations at lag 1 and 50 are given. The relative numerical efficiency coefficient (RNE) was introduced by Geweke (1989). Geweke's z scores test the stationarity of the draws from the posterior distribution, comparing the mean of the first 50% of the draws to the mean of the last 35% of the draws. z scores follow the standard normal distribution (see Geweke, 1992). The numbers presented in this table were computed using the package coda by Plummer et al. (2006).

Appendix C. Graphs summarising testing of the noncausality hypotheses

Graphs C.9–C.17 present simulated posterior distributions of function κ for different hypotheses. The shaded areas denote the 95% and 90% highest posterior density regions of the distributions. For more detailed results, refer to Tables 3–5. Description of the variables: y_1 = CHF/EUR, y_2 = GBP/EUR, y_3 = USD/EUR.

Figure C.9: Results of testing: Granger causality hypotheses I
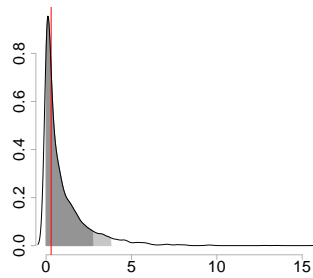
1. $\mathcal{H}_0 : y_1 \overset{G}{\nrightarrow} y_2|y_3$ 2. $\mathcal{H}_0 : y_1 \overset{G}{\nrightarrow} y_3|y_2$
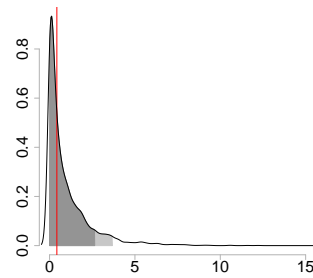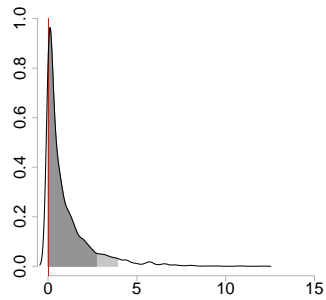
3. $\mathcal{H}_0 : y_2 \overset{G}{\nrightarrow} y_1|y_3$ 4. $\mathcal{H}_0 : y_2 \overset{G}{\nrightarrow} y_3|y_1$
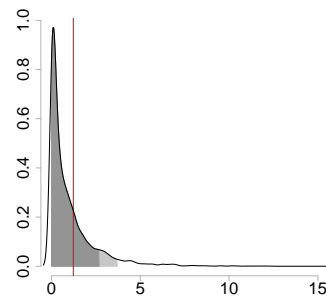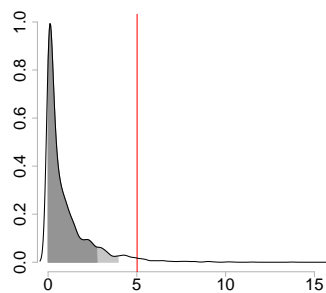
5. $\mathcal{H}_0 : y_3 \overset{G}{\nrightarrow} y_1|y_2$ 6. $\mathcal{H}_0 : y_3 \overset{G}{\nrightarrow} y_2|y_1$
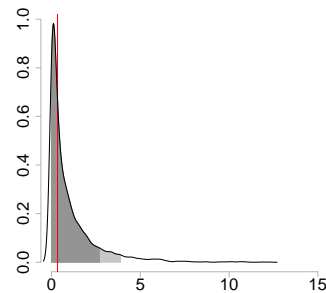
Visual representation of results from Table 3.

Figure C.10: Results of testing: Granger causality hypotheses II

1. $\mathcal{H}_0 : (y_1, y_2) \overset{G}{\nrightarrow} y_3$
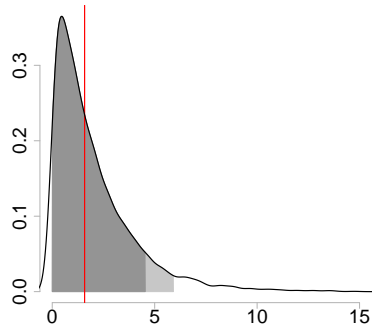
2. $\mathcal{H}_0 : (y_1, y_3) \overset{G}{\nrightarrow} y_2$
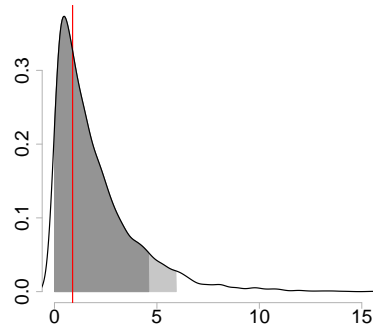
3. $\mathcal{H}_0 : (y_2, y_3) \overset{G}{\nrightarrow} y_1$
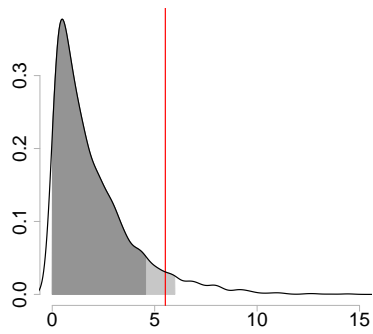
4. $\mathcal{H}_0 : y_1 \overset{G}{\nrightarrow} (y_2, y_3)$
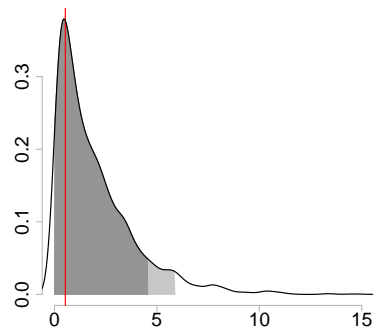
5. $\mathcal{H}_0 : y_2 \overset{G}{\nrightarrow} (y_1, y_3)$
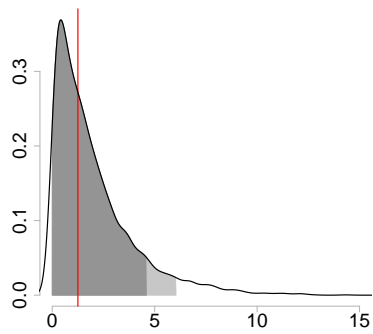
6. $\mathcal{H}_0 : y_3 \overset{G}{\nrightarrow} (y_1, y_2)$
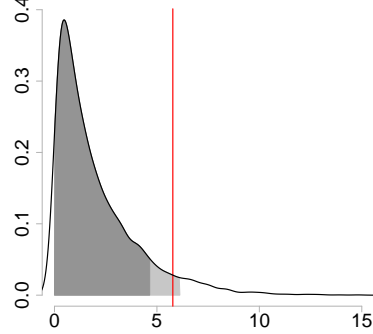
Visual representation of results from Table 3.

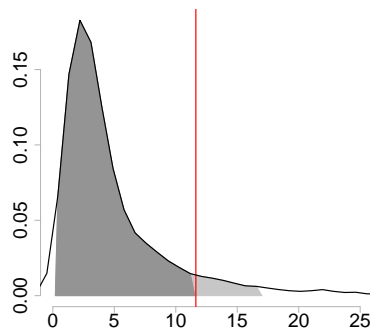Figure C.11: Results of testing: Granger causality hypotheses III

1. $\mathcal{H}_0 : y_1 \overset{G}{\nrightarrow} y_2 | y_3$ and $y_2 \overset{G}{\nrightarrow} y_1 | y_3$
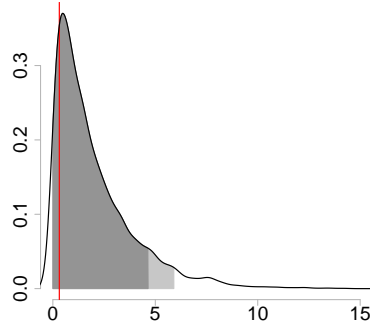
2. $\mathcal{H}_0 : y_1 \overset{G}{\nrightarrow} y_3 | y_2$ and $y_3 \overset{G}{\nrightarrow} y_1 | y_2$
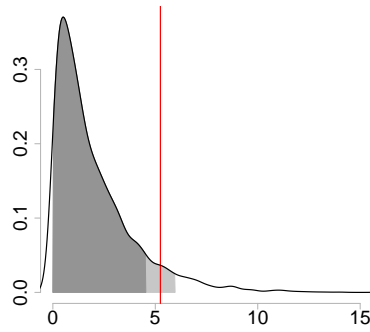
3. $\mathcal{H}_0 : y_2 \overset{G}{\nrightarrow} y_3 | y_1$ and $y_3 \overset{G}{\nrightarrow} y_2 | y_1$
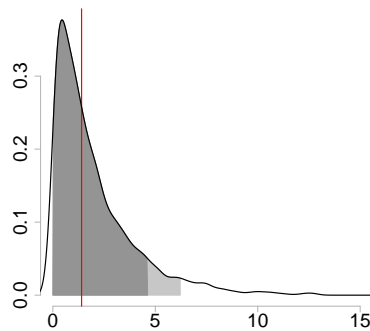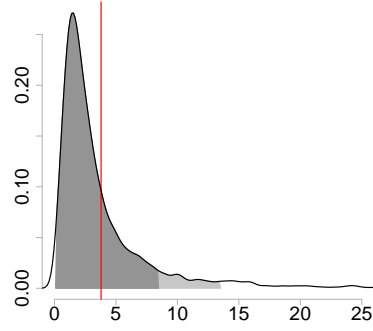

Visual representation of results from Table 3.

Figure C.12: Results of testing: second-order Granger causality hypotheses I

1. $\mathcal{H}_0 : y_1 \overset{so}{\nrightarrow} y_2 | y_3$
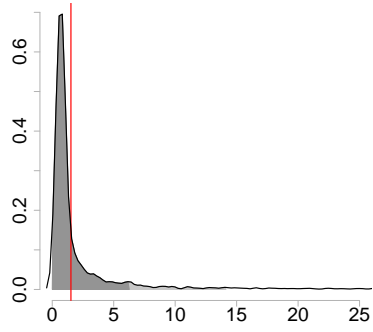

2. $\mathcal{H}_0 : y_1 \overset{so}{\nrightarrow} y_3 | y_2$
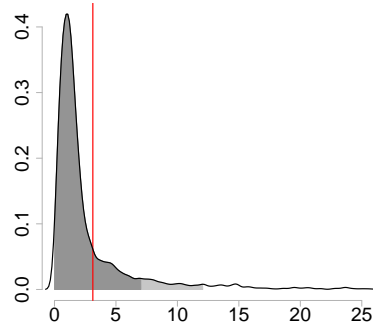

3. $\mathcal{H}_0 : y_2 \overset{so}{\nrightarrow} y_1 | y_3$
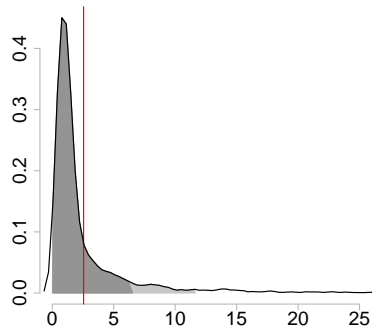
4. $\mathcal{H}_0 : y_2 \overset{so}{\nrightarrow} y_3 | y_1$
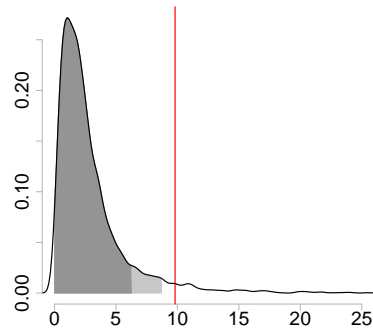
5. $\mathcal{H}_0 : y_3 \overset{so}{\nrightarrow} y_1 | y_2$
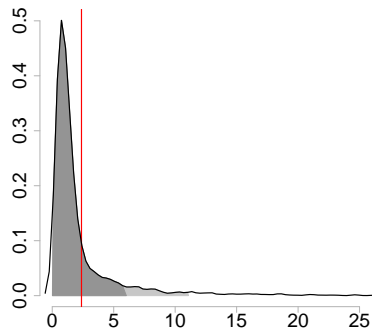
6. $\mathcal{H}_0 : y_3 \overset{so}{\nrightarrow} y_2 | y_1$
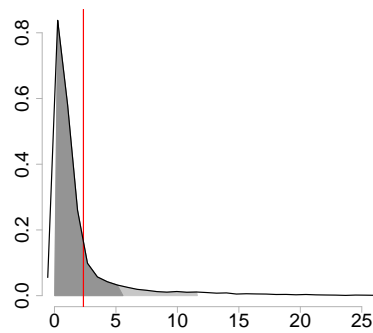
Visual representation of results from Table 4.

Figure C.13: Results of testing: second-order Granger causality hypotheses II

1. $\mathcal{H}_0 : (y_1, y_2) \overset{so}{\nrightarrow} y_3$
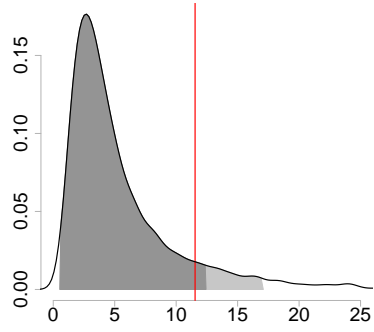
2. $\mathcal{H}_0 : (y_1, y_3) \overset{so}{\nrightarrow} y_2$
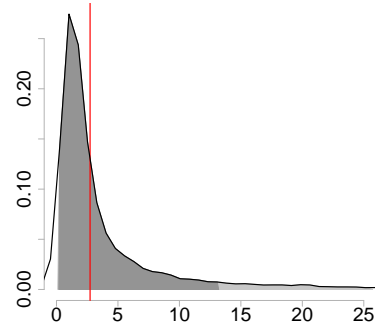
3. $\mathcal{H}_0 : (y_2, y_3) \overset{so}{\nrightarrow} y_1$
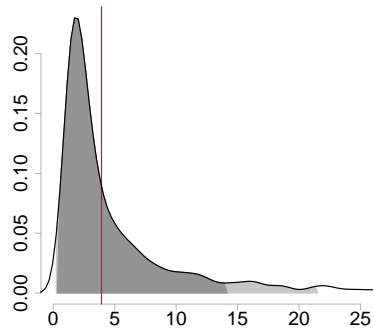
4. $\mathcal{H}_0 : y_1 \overset{so}{\nrightarrow} (y_2, y_3)$
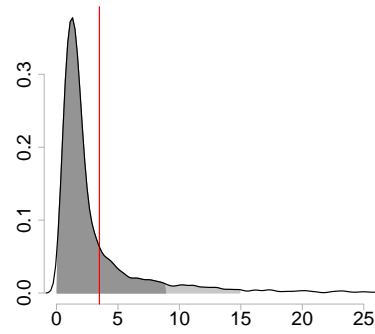
5. $\mathcal{H}_0 : y_2 \overset{so}{\nrightarrow} (y_1, y_3)$
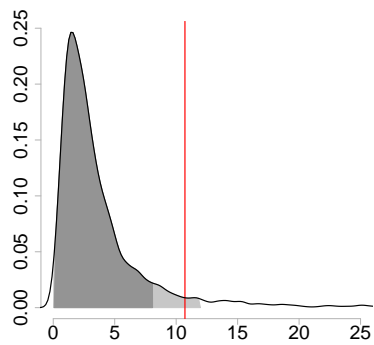
6. $\mathcal{H}_0 : y_3 \overset{so}{\nrightarrow} (y_1, y_2)$

Visual representation of results from Table 4.

31
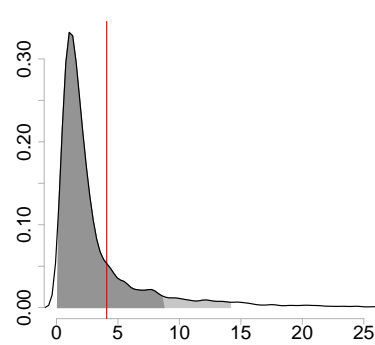
Figure C.14: Results of testing: second-order Granger causality hypotheses III

1. $\mathcal{H}_0 : y_1 \overset{so}{\nrightarrow} y_2|y_3$ and $y_2 \overset{so}{\nrightarrow} y_1|y_3$
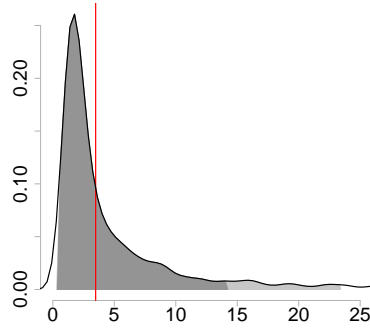
2. $\mathcal{H}_0 : y_1 \overset{so}{\nrightarrow} y_3|y_2$ and $y_3 \overset{so}{\nrightarrow} y_1|y_2$
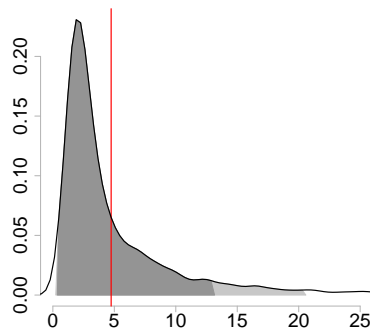
3. $\mathcal{H}_0 : y_2 \overset{so}{\nrightarrow} y_3|y_1$ and $y_3 \overset{so}{\nrightarrow} y_2|y_1$

Visual representation of results from Table 4.

Figure C.15: Results of testing: Granger causality in variance hypotheses I

1. $\mathcal{H}_0 : y_1 \overset{V}{\nrightarrow} y_2|y_3$
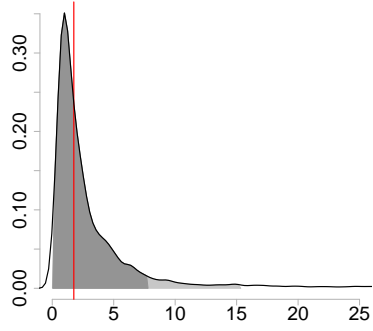
2. $\mathcal{H}_0 : y_1 \overset{V}{\nrightarrow} y_3|y_2$

3. $\mathcal{H}_0 : y_2 \overset{V}{\nrightarrow} y_1|y_3$
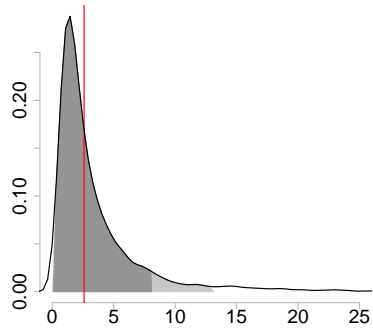
4. $\mathcal{H}_0 : y_2 \overset{V}{\nrightarrow} y_3|y_1$
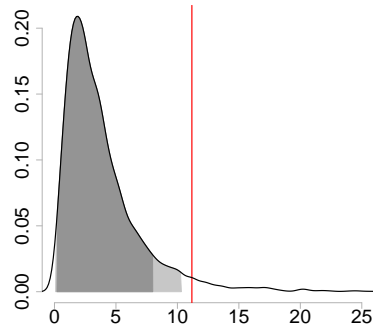
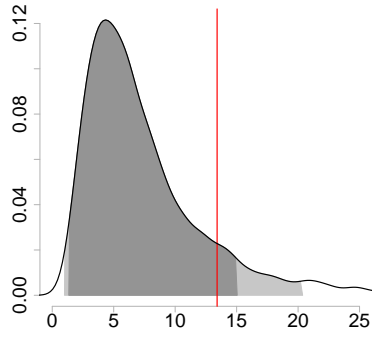5. $\mathcal{H}_0 : y_3 \overset{V}{\nrightarrow} y_1|y_2$
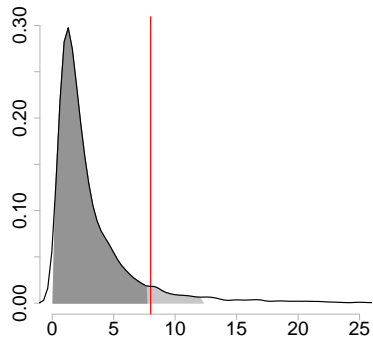
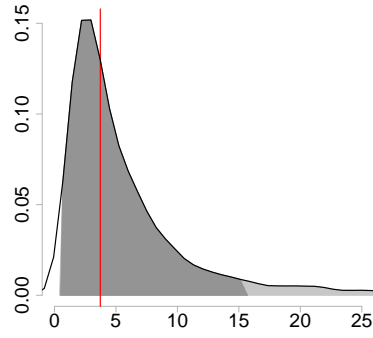
6. $\mathcal{H}_0 : y_3 \overset{V}{\nrightarrow} y_2|y_1$
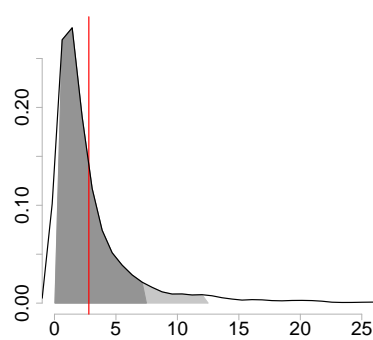
Visual representation of results from Table 5.

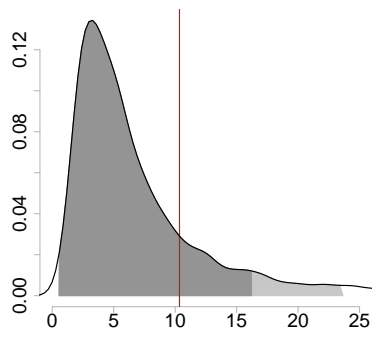Figure C.16: Results of testing: Granger causality in variance hypotheses II

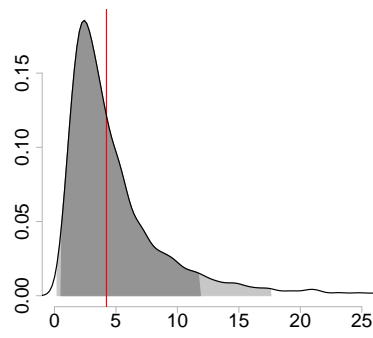1. $\mathcal{H}_0 : (y_1, y_2) \overset{V}{\nrightarrow} y_3$

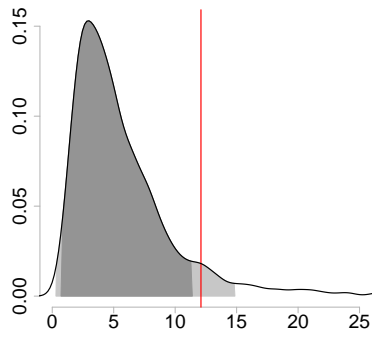2. $\mathcal{H}_0 : (y_1, y_3) \overset{V}{\nrightarrow} y_2$

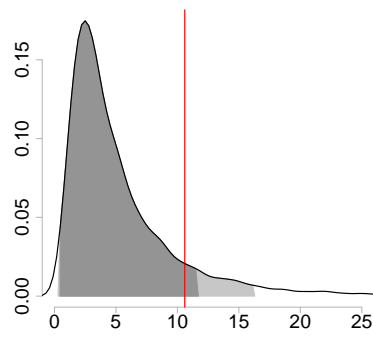3. $\mathcal{H}_0 : (y_2, y_3) \overset{V}{\nrightarrow} y_1$

4. $\mathcal{H}_0 : y_1 \overset{V}{\nrightarrow} (y_2, y_3)$

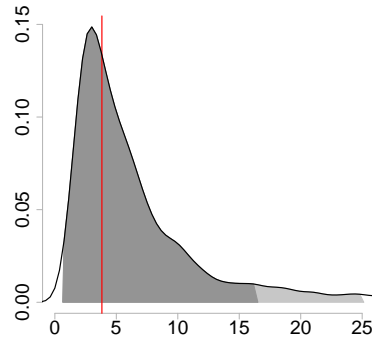5. $\mathcal{H}_0 : y_2 \overset{V}{\nrightarrow} (y_1, y_3)$

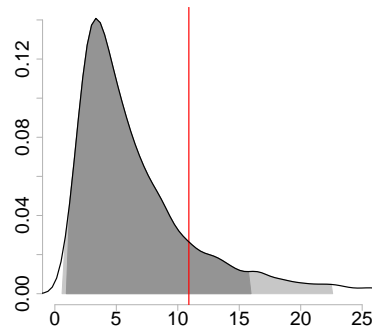6. $\mathcal{H}_0 : y_3 \overset{V}{\nrightarrow} (y_1, y_2)$

Visual representation of results from Table 5.

Figure C.17: Results of testing: Granger causality in variance hypotheses III

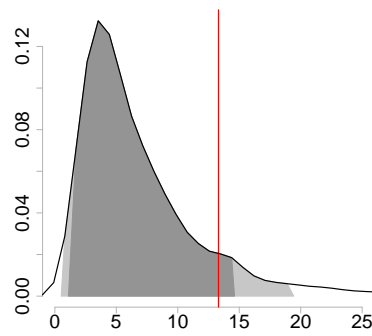### 1. $\mathcal{H}_0 : y_1 \overset{V}{\nrightarrow} y_2 | y_3$ and $y_2 \overset{V}{\nrightarrow} y_1 | y_3$

2. $\mathcal{H}_0 : y_1 \overset{V}{\nrightarrow} y_3 | y_2$ and $y_3 \overset{V}{\nrightarrow} y_1 | y_2$

3. $\mathcal{H}_0 : y_2 \overset{V}{\nrightarrow} y_3 | y_1$ and $y_3 \overset{V}{\nrightarrow} y_2 | y_1$

Visual representation of results from Table 5.

References

Bauwens, L., & Lubrano, M. (1998). Bayesian inference on GARCH models using the Gibbs sampler. *Econometrics Journal*, *1*, C23–C46.

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).

Deschamps, P. J. (2006). A flexible prior distribution for Markov switching autoregressions with Student-t errors. *Journal of Econometrics*, *133*, 153–190.

Diebold, F. X., & Yilmaz, K. (2009). Measuring Financial Asset Return and Volatility Spillovers, with Application to Global Equity Markets. *The Economic Journal*, *119*, 158–171.

Droumaguet, M., & Woźniak, T. (2012). *Bayesian Testing of Granger Causality in Markov-Switching VARs*. Working Paper Series European University Institute Florence, Italy. Download at: http://cadmus.eui.eu/bitstream/handle/1814/20815/ECO_2012_06.pdf?sequence=1.

Dufour, J.-M. (1989). Nonlinear hypotheses, inequality restrictions, and non-nested hypotheses: Exact simultaneous tests in linear regressions. *Econometrica*, *54*, 335–355.

Dufour, J.-M., Pelletier, D., & Renault, E. (2006). Short run and long run causality in time series: inference. *Journal of Econometrics*, *132*, 337–362.

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.

Geweke, J. (1989). Bayesian Inference in Econometric Models Using Monte Carlo Integration. *Econometrica*, *57*, 1317–1339.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In J. M. Bernardo, J. O. Berger, A. Dawid, & A. F. M. Smith (Eds.), *Bayesian Statistics 4*. Clarendon Press, Oxford volume 148.

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.

Jarociński, M., & Maćkowiak, B. (2011). Choice of Variables in Vector Autoregressions.

Jarque, C. M., & Bera, A. K. (1980). Efficient Tests for Normality, Homoscedasticity and Serial Indipendence of Regression Residuals. *Economics Letters*, *6*, 255–259.

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.

Kim, T., & White, H. (2004). On more robust estimation of skewness and kurtosis. *Finance Research Letters*, *1*, 56–73.

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.

Lomnicki, Z. (1961). Test for departure from normality in the case of linear stochastic processes. *Metrika*, *4*, 37–62.

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.

Lütkepohl, H., & Burda, M. M. (1997). Modified Wald tests under nonregular conditions. *Journal of Econometrics*, *78*, 315–332.

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.

Pajor, A. (2011). A Bayesian Analysis of Exogeneity in Models with Latent Variables. *Central European Journal of Economic Modelling and Econometrics*, *3*, 49–73.

Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). Coda: Convergence diagnosis and output analysis for mcmc. *R News*, *6*, 7–11.

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. (2012). *Testing Causality Between Two Vectors in Multivariate GARCH Models*. Working Paper Series 1139 Department of Economics, The University of Melbourne Melbourne. Download at: http://www.economics.unimelb.edu.au/research/1139.pdf.

European
University
Institute

European University Press