THE UNIVERSITY OF MELBOURNE

Department of Economics

Working Paper Series

Testing Causality Between Two Vectors in Multivariate GARCH Models

Tomasz Woźniak

January 2012

Research Paper Number 1139

ISSN: 0819-2642
ISBN: 978 0 7340 4490 7

Department of Economics
The University of Melbourne
Parkville VIC 3010
www.economics.unimelb.edu.au

Testing Causality Between Two Vectors in Multivariate 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

Spillover and contagion effects have gained significant interest in the recent years of financial crisis. Attention has not only been directed to relations between returns of financial variables, but to spillovers in risk as well. I use the family of Constant Conditional Correlation GARCH models to model the risk associated with financial time series and to make inferences about Granger causal relations between second conditional moments. The restrictions for *second-order Granger noncausality* between two vectors of variables are derived. To assess the credibility of the noncausality hypotheses, I employ Bayes factors. Bayesian testing procedures have not yet been applied to the problem of testing Granger noncausality. Contrary to classical tests, Bayes factors make such testing possible, regardless of the form of the restrictions on the parameters of the model. Moreover, they relax the assumptions about the existence of higher-order moments of the processes required in classical tests.

Keywords: Second-Order Causality, Volatility Spillovers, Bayes Factors, GARCH Models
JEL classification: C11, C12, C32, C53,

1. Introduction

The concept of the *Granger causality* was introduced in econometrics by Granger (1969) and Sims (1972). One vector of variables does not Granger-cause the other vector of variables if past information about the former cannot improve the forecast of the latter. Thus, the Granger causality or noncausality refers to the forecast of the conditional mean

[☆]This paper was presented at BMRC-QASS Conference in London in May 2010, FindEcon Conference in Łódź in May 2010 under the title *Bayesian testing of second-order causality*. The author thanks Helmut Lütkepohl, Jacek Osiewalski, Helmut Herwartz, Massimiliano Marcellino, Timo Teräsvirta, Vance Martin, Herman van Dijk and Lennart Hoogerheide for useful comments and remarks on the study. Also, the author thanks Michał Markun and colleagues from the EUI Time Series Econometrics Working Group for multiple discussions and suggestions.

Email address: tomasz.wozniak@unimelb.edu.au (Tomasz Woźniak)
URL: http://bit.ly/twozniak (Tomasz Woźniak)

process. The basic definition is set for a forecast of one period ahead value. The conditions imposed on the parameters of the linear Vector Autoregressive Moving Average model for Granger noncausality were derived by Boudjellaba, Dufour & Roy (1992) and Boudjellaba, Dufour & Roy (1994). The forecast horizon in the definition may, however, be generalized to h or up to h periods ahead, and h may have its limit in infinity (see Lütkepohl, 1993; Dufour & Renault, 1998). Irrespective of the forecast horizon, the restrictions imposed on parameters assuring noncausality may be nonlinear. This fact motivated development of nonstandard testing procedures that allow for the empirical verification of hypotheses (see Boudjellaba et al., 1992; Lütkepohl & Burda, 1997; Dufour, Pelletier & Renault, 2006).

This work examines the Granger causality for conditional variances. Consequently, I refer to the concept of the *second-order Granger causality*, introduced by Robins, Granger & Engle (1986), and formally distinguished from the *Granger causality in variance* by Comte & Lieberman (2000). One vector of variables does not second-order Granger-cause the other vector of variables if past information about the variability of the former cannot improve the forecast of conditional variances of the latter. The definition of the *second-order noncausality* assumes that Granger causal relations might exist in the conditional mean process, however, they should be modeled and filtered out. Otherwise, such relations may impact on the parameters responsible for causal relations in conditional variances (see the empirical illustration of the problem in Karolyi, 1995).

Restrictions already exist for the one-period-ahead second-order noncausality for the family of BEKK-GARCH models, delivered by Comte & Lieberman (2000). They have a form of several nonlinear functions of original parameters of the model. However, no good test has been established for such restrictions. The problem is that the matrix of the first partial derivatives of the restrictions, with respect to the parameters of the model, may not be of full rank. This fact translates to the unknown asymptotic properties of classical tests, even if the asymptotic distribution of the estimator is normal. As a consequence, the testing strategy developed by Comte & Lieberman (2000) and Hafner & Herwartz (2008b) is to derive linear (zero) restrictions on the parameters, which would be a sufficient condition for the original restrictions, and then to apply to them the Wald test.

I derive the conditions for the one-period-ahead second-order noncausality for the family of Extended Constant Correlation GARCH models of Jeantheau (1998). In this setting, all the considered variables are split in two vectors, between which we investigate causal relations in conditional variances. Then, the conditions for the one-period-ahead second-order noncausality appear to be the same as those for second-order noncausality in all future periods. When compared with the work of Comte & Lieberman (2000), these conditions result in a smaller number of restrictions. This has a practical meaning in computing the restricted models and may also potentially have a significant impact on the properties of tests applied to the problem.

In order to assess the credibility of the noncausality hypotheses, I employ Bayes factors, a standard Bayesian procedure. Surprisingly, Bayesian inference has not yet been proposed as a solution of testing the nonlinear restrictions for the Granger noncausality (Jarociński & Maćkowak, 2011, consider Bayes factors in order to test the linear restrictions for block-exogeneity in VAR models). Since the inference is performed using the Bayes factors, it is based on the exact finite sample results. Therefore, referring to the asymptotic results becomes pointless. This finding enables a relaxing of the assumptions required in the classical inference about the existence of the higher-order moments. For instance, in order to test the second-order noncausality hypothesis, the existence of the fourth order unconditional moments is required in Bayesian inference, whereas in classical testing the currently existing solutions require the existence of eighth-order moments. Notice that this assumption for testing such a hypothesis cannot be further relaxed in the context of the causal inference on second-conditional moments modeled with GARCH models. I justify this finding with the fact that this assumption does not come from the properties of the test, but from the derivation of the restrictions for the second-order noncausality.

The structure of the paper is as follows. Section 2 introduces the considered model and the main theoretical finding of this work: namely, the restrictions for the second-order Granger noncausality. The assumptions behind the causal analysis are discussed. In Section 3, I present and discuss the existing classical approaches to testing the Granger noncausality. Since they have limited use in the considered context, I further present the Bayes factors as the solution. In Section 4, the empirical illustration of the methodology for two main exchange rates of the Eurozone is presented, and Section 5 concludes. The proofs are presented in Appendix A, while Appendix B and Appendix C report figures and tables.

2. Second-order noncausality for multivariate GARCH models

Model. First, we set the notation, following Boudjellaba et al. (1994). Let $\{y_t : t \in \mathbb{Z}\}$ be a $N \times 1$ multivariate square integrable stochastic process on the integers $\mathbb{Z}$. Let $\mathbf{y} = (y_1, \dots, y_T)'$ denote a time series of T observations. Write

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

for all $t = 1, \dots, T$, where y_{it} is a $N_i \times 1$ vector such that $y_{1t} = (\mathrm{y}_{1t}, \dots, \mathrm{y}_{N_1.t})'$ and $y_{2t} = (\mathrm{y}_{N_1+1.t}, \dots, \mathrm{y}_{N_1+N_2.t})'$ ($N_1, N_2 \geq 1$ and $N_1 + N_2 = N$). y_1 and y_2 contain the variables of interest between which we want to study causal relations. 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 $I^2(t)$ be the Hilbert space generated by the product of variables $\epsilon_{i\tau}\epsilon_{j\tau}$, $1 \leq i, j \leq N$ for $\tau \leq t$. $I_{-1}(t)$ is the closed subspace of $I(t)$ generated by the components of $y_{2\tau}'$ and $I^2_{-1}(t)$ is the closed subspace of $I^2(t)$ generated by the variables $\epsilon_{i\tau}\epsilon_{j\tau}$, $N_1 + 1 \leq i, j \leq N$ for $\tau \leq t$. For any subspace I_t of $I(t)$ and for $N_1 + 1 \leq i \leq N_1 + N_2$, we denote by $P(y_{it+h}|I_t)$ the affine projection of y_{it+h} on I_t, i.e. the best linear prediction of y_{it+h} based on the variables in I_t and a constant term.

The model under consideration is the Vector Autoregressive process of Sims (1980) for conditional mean, and the Extended Constant Conditional Correlation Generalized Autoregressive Conditional Heteroskedasticity process of Jeantheau (1998) for conditional variances. The conditional mean part models linear relations between current and lagged observations of the considered variables:

$$y_t = \alpha_0 + \alpha(L)y_t + \epsilon_t \tag{2a}$$

$$\epsilon_t = D_t r_t \tag{2b}$$

$$r_t \sim i.i.St^N(\mathbf{0}, \mathbf{C}, \nu), \tag{2c}$$

for all $t = 1, \ldots, T$, where y_t is a $N \times 1$ vector of data at time t, $\alpha(L) = \sum_{i=1}^{p} \alpha_i L^i$ is a lag polynomial of order p, ϵ_t and r_t are $N \times 1$ vectors of residuals and standardized residuals respectively, $D_t = diag(\sqrt{h_{1t}}, \ldots, \sqrt{h_{Nt}})$ is a $N \times N$ diagonal matrix with conditional standard deviations on the diagonal. The standardized residuals follow a N-variate Student t distribution with a vector of zeros as a location parameter, a matrix $\mathbf{C}$ as a scale matrix and $\nu > 2$ a degrees of freedom parameter. The choice of the distribution is motivated, on the one hand, by its ability to model potential outlying observations in the sample (for $\nu < 30$). On the other hand, it is a good approximation of the normal distribution when the value of degrees of freedom parameter exceeds 30.

The conditional covariance matrix of the residual term ϵ_t is decomposed into:

$$H_t = D_t \mathbf{C} D_t \quad \forall t = 1, \ldots, T. \tag{3}$$

For the matrix H_t to be a positive definite covariance matrix, h_t must be positive for all t and $\mathbf{C}$ positive definite (see Bollerslev, 1990). A $N \times 1$ vector of current conditional variances is modeled with lagged squared residuals, $\epsilon_t^{(2)} = (\epsilon_{1t}^2, \ldots, \epsilon_{Nt}^2)'$, and lagged conditional variances:

$$h_t = \omega + A(L)\epsilon_t^{(2)} + B(L)h_t, \tag{4}$$

for all $t = 1, \ldots, T$, where ω is a $N \times 1$ vector of constants, $A(L) = \sum_{i=1}^{q} A_i L^i$ and $B(L) = \sum_{i=1}^{r} B_i L^i$ are lag polynomials of orders q and r of ARCH and GARCH effects respectively. The vector of conditional variances is given by $E[\epsilon_{t+1}^{(2)} | I^2(t)] = \frac{\nu}{\nu-2} h_{t+1}$, and the best linear predictor of $\epsilon_{t+1}^{(2)}$ in terms of a constant and $\epsilon_{t+1-i}^{(2)}$ for $i = 1, 2, \ldots$ is $P(\epsilon_{t+1}^{(2)} | I^2(t)) = h_{t+1}$. Equation (4) has a form respecting the partitioning of the vector of data (1):

$$\begin{bmatrix} h_{1t} \\ h_{2t} \end{bmatrix} = \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} + \begin{bmatrix} A_{11}(L) & A_{12}(L) \\ A_{21}(L) & A_{22}(L) \end{bmatrix} \begin{bmatrix} \epsilon_{1t}^{(2)} \\ \epsilon_{2t}^{(2)} \end{bmatrix} + \begin{bmatrix} B_{11}(L) & B_{12}(L) \\ B_{21}(L) & B_{22}(L) \end{bmatrix} \begin{bmatrix} h_{1t} \\ h_{2t} \end{bmatrix}. \tag{5}$$

Assumptions and properties. Let $\theta \in \Theta \subset \mathbb{R}^k$ be a vector of size k, collecting all the parameters of the model described with equations (2)–(4). Then the likelihood function has the following form:

$$p(\mathbf{y}|\theta) = \prod_{t=1}^{T} \frac{\Gamma\left(\frac{\nu+N}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} ((\nu-2)\pi)^{-\frac{N}{2}} |H_t|^{-\frac{1}{2}} \left(1 + \frac{1}{\nu-2} \epsilon_t' H_t^{-1} \epsilon_t\right)^{-\frac{\nu+N}{2}}. \tag{6}$$

This model has its origins in the Constant Conditional Correlation GARCH (CCC-GARCH) model proposed by Bollerslev (1990). That model consisted of N univariate GARCH equations describing the vector of conditional variances, h_t. The CCC-GARCH model is equivalent to equation (4) with diagonal matrices $A(L)$ and $B(L)$. Its extended version, with non-diagonal matrices $A(L)$ and $B(L)$, was analyzed by Jeantheau (1998). He & Teräsvirta (2004) call this model the Extended CCC-GARCH (ECCC-GARCH). Such a formulation of the GARCH process allows for the modeling of volatility spillovers, as matrices of the lag polynomials $A(L)$ and $B(L)$ are not diagonal. Therefore, causality between variables in second-conditional moments may be analyzed.

For the purpose of deriving the restrictions for second-order Granger noncausality, I impose four assumptions on the parameters of the conditional variance process.

Assumption 1. Parameters $\omega, A = \left(\text{vec}(A_1)', \ldots, \text{vec}(A_q)'\right)'$ and $B = (\text{vec}(B_1)', \ldots, \text{vec}(B_r)')'$ are such that the conditional variances, h_t, are positive for all t (see Conrad & Karanasos, 2010, for the detailed restrictions).

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

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

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

Define a process $v_t = \epsilon_t^{(2)} - h_t$. Then $\epsilon_t^{(2)}$ follows a VARMA process given by:

$$\phi(L)\epsilon_t^{(2)} = \omega + \psi(L)v_t, \tag{7}$$

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(q,r) process. Suppose that $\epsilon_t^{(2)}$ and v_t are partitioned as y_t in (1). Then (7) can be written in the following form:

$$\begin{bmatrix} \phi_{11}(L) & \phi_{12}(L) \\ \phi_{21}(L) & \phi_{22}(L) \end{bmatrix} \begin{bmatrix} \epsilon_{1t}^{(2)} \\ \epsilon_{2t}^{(2)} \end{bmatrix} = \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} + \begin{bmatrix} \psi_{11}(L) & \psi_{12}(L) \\ \psi_{21}(L) & \psi_{22}(L) \end{bmatrix} \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix}. \tag{8}$$

Given Assumption 3, the VARMA process (7) is invertible and can be written in a VAR form:

$$\Pi(L)\epsilon_t^{(2)} - \sigma^* = v_t, \tag{9}$$

where $\Pi(L) = \psi(L)^{-1}\phi(L) = [I_N - B(L)]^{-1}[I_N - A(L) - B(L)]$ is a matrix polynomial of the VAR representation of the GARCH(q,r) process and $\sigma^* = \psi(L)^{-1}\omega$ is a constant term. Again, partitioning the vectors, we can rewrite (9) in the form:

$$\begin{bmatrix} \Pi_{11}(L) & \Pi_{12}(L) \\ \Pi_{21}(L) & \Pi_{22}(L) \end{bmatrix} \begin{bmatrix} \epsilon_{1t}^{(2)} \\ \epsilon_{2t}^{(2)} \end{bmatrix} - \begin{bmatrix} \sigma_1^* \\ \sigma_2^* \end{bmatrix} = \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix}. \tag{10}$$

Under Assumptions 2 and 3, processes (7) and (9) are both stationary. One more assumption is required for the inference about second-order noncausality in the GARCH model:

Assumption 5. The process v_t is covariance stationary.

The GARCH model has well-established properties under Assumptions 1–5. Under Assumption 1, conditional variances are positive. This result does not require that all the parameters of the model are positive (see Conrad & Karanasos, 2010). Further, Jeantheau (1998) proves that the GARCH(r,s) model, as in (4), has a unique, ergodic, weakly and strictly stationary solution when Assumption 2 holds. Under Assumptions 2–4 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 conditions of, among others, stationarity and identifiability. Ling & McAleer (2003) proved strong consistency of the Quasi Maximum Likelihood Estimator (QMLE) for the VARMA-GARCH model under Assumptions 2–4, and when all the parameters of the GARCH process are positive. Moreover, they have set asymptotic normality of QMLE, provided that $E\|y_t\|^6 < \infty$. The extension of the asymptotic results under the conditions of (Conrad & Karanasos, 2010) has not yet been established. Finally, He & Teräsvirta (2004) give sufficient conditions for the existence of the fourth moments and derive complete fourth-moment structure.

Estimation. Classical estimation consists of maximizing the likelihood function (6). This is possible, using one of the available numerical optimization algorithms. Due to the complexity of the problem, the algorithms require derivatives of the likelihood function. Hafner & Herwartz (2008a) give analytical solutions for first and second partial derivatives of normal likelihood function, whereas Fiorentini, Sentana & Calzolari (2003) derive numerically reliable analytical expressions for the score, Hessian and information matrix for the models with conditional multivariate t-Student distribution. Bayesian estimation requires numerical methods in order to simulate the posterior density of the parameters. Unfortunately, neither the posterior distribution of the parameters nor full conditional distributions have the form of some known distribution. Therefore, the application of the Metropolis-Hastings algorithm (see Chib & Greenberg, 1995, and references therein) was proposed by Vrontos, Dellaportas & Politis (2003).

The posterior distribution of the parameters of the model is proportional to the product of the likelihood function (6) and the prior distribution of the parameters:

$$p(\theta|\mathbf{y}) \propto p(\mathbf{y}|\theta)p(\theta). \tag{11}$$

For the unrestricted VAR-GARCH model, I assume the following prior specification. All the parameters of the VAR process are *a priori* normally distributed with a vector of zeros as a mean and a diagonal covariance matrix with 100s on the diagonal. A similar prior distribution is assumed for the parameters of the GARCH process, with the difference that a truncation of the distribution to the parameter space respecting Assumptions 1–5 is imposed. All the correlation parameters of the correlation matrix $\mathbf{C}$ follow a uniform distribution on the interval $[-1, 1]$. Finally, for the degrees of freedom parameter, I assume the prior distribution proposed by Deschamps (2006). To summarize, the prior specification for the considered model has a detailed form of:

$$p(\theta) = p(\alpha)p(\omega, A, B)p(\nu) \prod_{i=1}^{N(N-1)/2} p(\rho_i), \tag{12}$$

where each of the prior distributions is assumed:

$$\begin{aligned}
\alpha &\sim \mathcal{N}^{N+pN^2}\left(\mathbf{0}, 100 \cdot I_{N+pN^2}\right) \\
(\omega', A', B')' &\sim \mathcal{N}^{N+N^2(q+r)}\left(\mathbf{0}, 100 \cdot I_{N+N^2(q+r)}\right)\mathcal{I}(\theta \in \Theta) \\
\nu &\sim .04 \exp\left[-.04(\nu - 2)\right]\mathcal{I}(\nu \geq 2) \\
\rho_i &\sim \mathcal{U}(-1, 1) \quad \text{ for } i = 1, \ldots, N(N-1)/2,
\end{aligned}$$

where $\alpha = (\alpha_0', \text{vec}(\alpha_1)', \ldots, \text{vec}(\alpha_p)')'$ stacks all the parameters of the VAR process in a vector of size $N + pN^2$. I_n is an identity matrix of order n. $\mathcal{I}(.)$ is an indicator function taking value equal to 1 if the condition in the brackets holds and 0 otherwise. Finally, ρ_i is an ith element of a vector stacking all the elements below the diagonal of the correlation matrix, $\rho = (\text{vecl}(\mathbf{C}))$.

Such prior assumptions, with only proper distributions, have serious consequences. First, together with the bounded likelihood function, the proposed prior distribution guarantees the existence of the posterior distribution (see Geweke, 1997). Second, the proper prior distribution for the degrees of freedom parameter of the t-Student distribution is required for the posterior distribution to be integrable, as proven by Bauwens & Lubrano (1998). Finally, it gives raise to subjective interpretation of the probability, which is a controversial feature of the Bayesian inference. However, note that prior distributions for all the parameters, except ν, do not in fact discriminate any of the values that these parameters may take. The prior distribution of the degrees of freedom parameter gives more than 32 percent a chance of that its value will be higher than 30, which makes the likelihood function a good approximation to the normally distributed function.

Second-Order Noncausality Conditions. We focus on the question of the causal relations between variables in conditional variances. Therefore, the proper concept to refer to is *second-order Granger noncausality*:

Definition 1. y_1 does not second-order Granger-cause h periods ahead y_2 if:

$$P\left[\epsilon^{(2)}(y_{2t+h}|I(t))|I^2(t)\right] = P\left[\epsilon^{(2)}(y_{2t+h}|I(t))|I^2_{-1}(t)\right], \tag{13}$$

for all $t \in \mathbb{Z}$, where $\epsilon_{2t+h} = \epsilon(y_{2t+h}|I(t)) = y_{2t+h} - P(y_{2t+h}|I(t))$ is an error component and $[.]^{(2)}$ means that we square each element of a vector and $h \in \mathbb{Z}$.

A common part of both sides of (13) is that, in the first step, the potential Granger causal relations in the conditional mean process are filtered out. This is represented by a projection of the forecasted value, y_{2t+h}, on the Hilbert space generated by the full set of variables, $P(y_{2t+h}|I(t))$. In the second stage, the square of the error component, $\epsilon^{(2)}(y_{2t+h}|I(t))$, is projected on the Hilbert space generated by the full set of variables, $I^2(t)$ (on the LHS), and on the Hilbert space generated by the restricted set of variables, $I^2_{-1}(t)$ (on the RHS). If the two projections are equivalent, it means that $\epsilon^{(2)}(y_{2t+h}|I(t)) - P\left[\epsilon^{(2)}(y_{2t+h}|I(t))|I^2_{-1}(t)\right]$ is orthogonal to $I^2(t)$ for all t (see Florens & Mouchart, 1985; Comte & Lieberman, 2000). Note also the difference between this definition of second-order noncausality and the definition of Comte & Lieberman (2000). In Definition 1 the Hilbert space $I^2(t)$ is generated by the cross-products of the error components ϵ_τ, whereas in the definition of Comte & Lieberman it is generated by the cross-products of the variables $\mathsf{y}_{i\tau}$ and $\tau \leq T$.

The definition, in its original form, for one-period-ahead noncausality ($h = 1$), was proposed by Robins et al. (1986) and distinguished from *Granger noncausality in variance* by Comte & Lieberman (2000). The difference is that in the definition of *Granger noncausality in variance* there is another assumption of *Granger noncausality in mean*. On the contrary, in the definition of *second-order noncausality* there is no such assumption. However, any existing causal relation in conditional means needs to be modeled and filtered out before causality for the conditional variances process is analyzed.

The main theoretical contribution of this study is the theorem stating the restrictions for second-order Granger noncausality for the ECCC-GARCH model.

Theorem 1. *Let $\epsilon_t^{(2)}$ follow a stationary vector autoregressive moving average process as in (7) partitioned as in (8) that is identifiable and invertible (assumptions 1–5). Then y_1 does not second-order Granger-cause one period ahead y_2 if and only if:*

$$\Gamma^{so}_{ij}(z) = \det\begin{bmatrix} \phi^{\cdot j}_{11}(z) & \psi_{11}(z) \\ \varphi_{n_1+i,j}(z) & \psi^{i\cdot}_{21}(z) \end{bmatrix} = 0 \quad \forall z \in \mathbb{C} \tag{14}$$

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

Theorem 1 establishes the restrictions on parameters of the ECCC-GARCH model for the second-order noncausality one period ahead between two vectors of variables. Its proof, presented in Appendix A, is based of the theory introduced by Florens & Mouchart (1985) and applied by Boudjellaba et al. (1992) to VARMA models for conditional mean. It is applicable to any specification of the GARCH(q,r) process, irrespective of the order of the model, (q, r), and the size for the time series, N.

Due to the setting proposed in this study, in which the vector of variables is split in two parts, the establishment of one-period-ahead second-order Granger noncausality is equivalent to establishing the noncausality relation at all horizons up to infinity. This result is formalised in a corollary.

Corollary 1. *Suppose that the vector of observations is partitioned as in (1), and that y_1 does not second-order Granger-cause one period ahead y_2, such that the condition (14) holds. Then y_1 does not second-order Granger-cause h periods ahead y_2 for all $h = 1, 2, \ldots$.*

Corollary 1 is a direct application of Corollary 2.2.1 of (Lütkepohl, 2005, p. 45) to the GARCH process in the VAR form (10). For the proof of the restrictions for second-order Granger noncausality for the GARCH process in the VAR form, the reader is referred to Appendix A.

Corollary 1 shows the feature of the particular setting considered in this work, i.e. the setting in which all the variables are split between two vectors. If one is interested in the second-order causality relations at all the horizons at once, then one may use just one set of restrictions. The restrictions, however, imply the very strong result. If a more detailed analysis is required, then one must consider deriving other solutions.

The theorem has equivalent for other models from the GARCH family, namely the BEKK-GARCH models. The restrictions were introduced by Comte & Lieberman (2000). There are, however, serious differences between the approaches presented by Comte & Lieberman and by this study. First, in a bivariate model for the hypothesis that one variable does not second-order cause the other, the restrictions of Comte & Lieberman lead to six restrictions, whereas, in Example 1, I show that in order to test such a hypothesis, only two restrictions are required. The difference in the number of restrictions increases with the dimension of the time series. Secondly, due to the formulation of the BEKK-GARCH model, the noncausality conditions are much more complicated than the conditions for the ECCC-GARCH model considered here. They are simply much more complex functions of the original parameters of the model. Both these arguments have consequences in testing that require estimation of the restricted model or employment of the delta method. A high number of restrictions may have a strongly negative impact on the size and power properties of tests. However, the ECCC-GARCH model assumes that the correlations are time invariant, which is not the case for the BEKK-GARCH model.

Testing of the restrictions is, however, the main difference between the two approaches. Comte & Lieberman (2000), in order to test the restrictions, derive another set of zero restrictions that are only a sufficient condition for the determinant restriction (corresponding to (14) from Theorem 1). In Section 3 I propose a Bayesian test for the original restrictions derived for the determinant condition (14) and discuss the shortcomings of the existing testing procedures.

Nakatani & Teräsvirta (2009) propose the Lagrange Multiplier test for the hypothesis of no volatility spillovers in a bivariate ECCC-GARCH model. The restrictions they test are zero restrictions on the off-diagonal elements of matrix polynomials $A(L)$ and $B(L)$ from the GARCH equation (5). Consequently, the null hypothesis is represented by the CCC-GARCH model of Bollerslev (1990) and the alternative hypothesis by ECCC-GARCH of Jeantheau (1998). Note, that if all the parameters on the diagonal of the matrices of the lag polynomial $A_{11}(L)$ are assumed to be strictly greater than zero (which can be tested and which in fact is the case for numerous time series considered in applied studies), then the null hypothesis of Nakatani & Teräsvirta (2009) is equivalent to the second-order Granger noncausality condition, as in the Example 1. In a general case, for any dimension of the time series, the zero restrictions on the off-diagonal elements of matrix polynomials $A(L)$ and $B(L)$ represent a sufficient condition for the second-order noncausality.

To conclude, the condition (14) leads to the finite number of nonlinear restrictions on the original parameters of the model. Several examples will clarify how they are set.

Example 1. Suppose that y_t follows a bivariate GARCH(1,1) process, ($N = 2$, and $p = q = 1$). The VARMA process for $\epsilon_t^{(2)}$ is as follows:

$$\begin{bmatrix} 1-(A_{11}+B_{11})L & -(A_{12}+B_{12})L \\ -(A_{21}+B_{21})L & 1-(A_{22}+B_{22})L \end{bmatrix} \begin{bmatrix} \epsilon_{1t}^2 \\ \epsilon_{2t}^2 \end{bmatrix} = \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} + \begin{bmatrix} 1-B_{11}L & -B_{12}L \\ -B_{21}L & 1-B_{22}L \end{bmatrix} \begin{bmatrix} \nu_{1t} \\ \nu_{2t} \end{bmatrix}. \tag{15}$$

From Theorem 1, we see, that y_1 does not second-order Granger-cause y_2 if and only if:

$$\det \begin{bmatrix} 1-(A_{11}+B_{11})z & 1-B_{11}z \\ -(A_{21}+B_{21})z & -B_{21}z \end{bmatrix} \equiv \mathbf{0}, \tag{16}$$

which leads to the following set of restrictions:

$$\mathbf{R}_1^I(\theta) = A_{21} = 0, \quad \text{and} \quad \mathbf{R}_2^I(\theta) = B_{21}A_{11} = 0. \tag{17}$$

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

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

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

From Theorem 1, we see, that $y_1 = \mathsf{y}_1$ does not second-order Granger-cause $y_2 = (\mathsf{y}_2, \mathsf{y}_3)$ if and only if:

$$\det \begin{bmatrix} 1-(A_{11}+B_{11})z & 1-B_{11}z \\ -(A_{i1}+B_{i1})z & -B_{i1}z \end{bmatrix} = 0 \quad \text{for } i = 2, 3, \tag{19}$$

which results in the following restrictions:

$$\mathbf{R}_1^{II}(\psi) = A_{11}B_{21} = 0 \quad \text{and} \quad \mathbf{R}_2^{II}(\psi) = A_{21} = 0 \tag{20a}$$

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

However, $y_1 = (\mathsf{y}_1, \mathsf{y}_2)$ does not second-order Granger-cause $y_2 = \mathsf{y}_3$ if and only if:

$$\det \begin{bmatrix} 1-(A_{11}+B_{11})z & 1-B_{11}z & -B_{13}z \\ -(A_{21}+B_{21})z & -B_{21}z & -B_{23}z \\ -(A_{31}+B_{31})z & -B_{31}z & 1-B_{33}z \end{bmatrix} \equiv \mathbf{0}, \tag{21}$$

which leads to the following set of restrictions:

$$\mathbf{R}_1^{III}(\psi) = A_{11}(B_{23}B_{31} - B_{21}B_{33}) + A_{31}(B_{13}B_{21} - B_{11}B_{23}) = 0 \tag{22a}$$

$$\mathbf{R}_2^{III}(\psi) = A_{11}B_{21} + A_{31}B_{23} = 0 \tag{22b}$$

$$\mathbf{R}_3^{III}(\psi) = A_{21} = 0. \tag{22c}$$

3. Bayesian testing

The restrictions derived in Section 2 can be tested. I propose to use the Bayesian approach to testing the restrictions for second-order noncausality, which has not yet been done. Before the approach is presented, however, the classical tests proposed so far and their limitations are discussed.

Classical testing. Testing of second-order noncausality has been considered only for the family of BEKK-GARCH and vec-GARCH models. Comte & Lieberman (2000) did not propose any test because asymptotic normality of the maximum likelihood estimator had not been established at that time. The asymptotic result was presented in Comte & Lieberman (2003). This finding, however, does not solve the problem of testing the nonlinear restrictions imposed on the parameters of the model. In the easy case, when the restrictions are linear, the asymptotic normality of the estimator implies that the Wald, Lagrange Multiplier and Likelihood Ratio test statistics have asymptotic χ^2 distributions. Therefore, the Wald test statistic for the linear restrictions (which are the only sufficient condition for the original restrictions) proposed by Comte & Lieberman is χ^2-distributed. A similar procedure was presented in Hafner & Herwartz (2008b) for the Wald test, and in Hafner & Herwartz (2006) for the LM test. This approach is, in fact, an arbitrary choice of the restrictions and a simplification of the problem. For the ECCC-GARCH model, Nakatani & Teräsvirta (2009) proposed the Lagrange Multiplier test for the hypothesis of no volatility spillovers. The test statistic is shown to be asymptotically normally distributed. Again, Nakatani & Teräsvirta (2009) tested only the linear zero restrictions.

In this study, the necessary and sufficient conditions for second-order noncausality between variables are tested. The restrictions, contrary to the conditions of Comte & Lieberman, Hafner & Herwartz and Nakatani & Teräsvirta, may be nonlinear (see Example 2). In such a case, a matrix of the first partial derivatives of the restrictions with respect to the parameters may not be of full rank. Thus, the asymptotic distribution of the Wald test statistic is no longer normal. In fact, for the time being it is unknown. Consequently, the Wald test statistic cannot be used to test the necessary and sufficient conditions for second-order noncausality in multivariate GARCH models.

This problem is well known in the studies on the testing of parameter conditions for Granger noncausality in multivariate models. Boudjellaba et al. (1992) derive conditions for Granger noncausality for VARMA models that result in multiple nonlinear restrictions on original parameters of the model. As a solution to the problem of testing the restrictions, they propose a sequential testing procedure. There are two main drawbacks in this method. First, despite being properly performed, the test may still appear inconclusive, and second, the confidence level is given in the form of inequalities. Dufour et al. (2006) propose solutions based on the linear regression techniques that are applied for h-step ahead Granger noncausality for VAR models. The proposed solutions, unfortunately, are only applicable to linear models for first conditional moments. Lütkepohl & Burda (1997) proposed a modified Wald test statistic as a solution to the problem of testing the nonlinear restrictions for the h-step ahead Granger noncausality for VAR models. This method could be applied to the problem of testing the nonlinear restrictions for the second-order noncausality in GARCH models. More studies are required, however.

Asymptotic results for the models and tests discussed here are established under the following moment conditions. For the BEKK-GARCH models, the Wald tests proposed by Hafner & Herwartz (2008b) and Comte & Lieberman (2000) require asymptotic normality of the Quasi Maximum Likelihood Estimator. This result is derived under the existence of bounded moments of order 8 by Comte & Lieberman (2003). For the ECCC-GARCH model considered in this study, the asymptotic normality of the Quasi Maximum Likelihood Estimator is derived in Ling & McAleer (2003) under the existence of moments of order 6. This assumption is, however, relaxed for the purpose of testing the existence of volatility spillovers by Nakatani & Teräsvirta (2009). Their Lagrange Multiplier test statistic requires the existence of fourth-order moments. The Bayesian test presented below further relaxes this assumption.

It is necessary to mention at this stage the approach to testing noncausality in variance presented by Cheung & Ng (1996). They propose a two-stage procedure. In the first stage, for each of the variables a univariate ARMA-GARCH model is fitted and estimated. In the second stage, a test based on the cross-correlation function between squared standardized residuals from the first stage for all the considered variables is performed. Such an approach allows for detection of the Granger causal relations, however, it does not model the spillovers and is proposed as a pre-estimation method of constructing multivariate models.

Bayesian testing. In order to test the restrictions derived in Theorem 1, I use Bayes factors. This is a well-known method for comparing econometric models (see Kass & Raftery, 1995; Geweke, 1995). Denote by $\mathcal{M}_i$, for $i = 1, \ldots, m$, the m models representing competing hypotheses. Let

$$p(\mathbf{y}|\mathcal{M}_i) = \int_{\theta\in\Theta} p(\mathbf{y}|\theta, \mathcal{M}_i)p(\theta|\mathcal{M}_i)d\theta \tag{23}$$

be marginal distributions of data corresponding to the models, for $i = 1, \ldots, m$. $p(\mathbf{y}|\theta, \mathcal{M}_i)$ and $p(\theta|\mathcal{M}_i)$ are the likelihood function (6) and the prior distribution (12) respectively. The extended notation respecting conditioning on one of the models is used here. A Bayes factor is a ratio of the marginal densities of data for the two selected models:

$$\mathcal{B}_{ij} = \frac{p(\mathbf{y}|\mathcal{M}_i)}{p(\mathbf{y}|\mathcal{M}_j)}, \tag{24}$$

where $i, j = 1, \ldots, m$ and $i \neq j$.

The Bayes factor takes positive values, and its value above 1 is interpreted as evidence for model $\mathcal{M}_i$, whereas its value below 1 is evidence for model $\mathcal{M}_j$. For further interpretation of the value of the Bayes factor, the reader is referred to the paper of Kass & Raftery (1995). Further, for the interpretation of the Bayes factor itself, it is useful to look at the Posterior Odds Ratio being a ratio of the posterior probabilities of models:

$$\frac{Pr(\mathcal{M}_i|\mathbf{y})}{Pr(\mathcal{M}_j|\mathbf{y})} = \frac{Pr(\mathcal{M}_i)}{Pr(\mathcal{M}_j)} \times \mathcal{B}_{ij}, \tag{25}$$

where $Pr(\mathcal{M}_i|\mathbf{y})$ denotes the posterior probability of model $\mathcal{M}_i$ and $Pr(\mathcal{M}_i)$ is its prior probability. The Posterior Odds Ratio (25) is a product of the Prior Odds Ratio and the Bayes factor. When an investigator does not want to discriminate *a priori* between any of the hypotheses or models, setting equal prior probabilities for all the models, then the Bayes factor is equal to the Posterior Odds Ratio. Consequently, it informs the amount by which (when its value is above 1) model $\mathcal{M}_i$ is more probable *a posteriori* than model $\mathcal{M}_j$.

This approach to testing requires an estimation of models representing all considered hypotheses as well as of the marginal densities of data (23).

Estimation of models. The form of the posterior distribution (11) for all the parameters, θ, for the GARCH models, even with the prior distribution set to proper distribution functions as in (12), is not of any known form. Moreover, none of the full conditional densities for any sub-group of the parameter vector has a form of some standard distribution. Still, the posterior distribution, although it is known only up to a normalising constant, exists; this is ensured by the bounded likelihood function and the proper prior distribution. Therefore, the posterior distribution may be simulated with a Monte Carlo Markov Chain (MCMC) algorithm. Due to the above mentioned problems with the form of the posterior and full conditional densities, a proper algorithm to sample the posterior distribution (11) is the Metropolis-Hastings algorithm (see Chib & Greenberg, 1995, and references therein). The algorithm was adapted for multivariate GARCH models by Vrontos et al. (2003).

Suppose the starting point of the Markov Chain is some value $\theta_0 \in \Theta$. Let $q(\theta^{(s)}, \theta'|\mathbf{y}, \mathcal{M}_i)$ denote the proposal density (candidate-generating density) for the transition from the current state of the Markov chain $\theta^{(s)}$ to a candidate draw θ'. The candidate density for model $\mathcal{M}_i$ depends on the data $\mathbf{y}$. In this study, I use a multivariate Student's t distribution with the location vector set to the current state of the Markov chain, $\theta^{(s)}$, the scale matrix Ω_q and the degrees of freedom parameter set to five. The scale matrix, Ω_q, should be determined by preliminary runs of the MCMC algorithm, such that it is close to the covariance matrix of the posterior distribution. Such a candidate-generating density should enable the algorithm to draw relatively efficiently from the posterior density. A new candidate θ' is accepted with the probability:

$$\alpha(\theta^{(s)}, \theta'|\mathbf{y}, \mathcal{M}_i) = \min\left[1, \frac{p(\mathbf{y}|\theta', \mathcal{M}_i)p(\theta'|\mathcal{M}_i)}{p(\mathbf{y}|\theta^{(s)}, \mathcal{M}_i)p(\theta^{(s)}|\mathcal{M}_i)}\right],$$

and if it is rejected, then $\theta^{(s+1)} = \theta^{(s)}$. The sample drawn from the posterior distribution with the Metropolis-Hastings algorithm, $\{\theta^{(s)}\}_{s=1}^{S}$, should be diagnosed to ensure that it is a good sample from the stationary posterior distribution (see e.g. Geweke, 1999; Plummer, Best, Cowles & Vines, 2006).

Estimation of the marginal distribution of data. Having estimated the models, the marginal densities of data may be computed using one of the available methods. Since the estimation of the models is performed using the Metropolis-Hastings algorithm, the suitable estimators of the marginal density of data are presented by Newton & Raftery

(1994), Chib & Jeliazkov (2001) and Geweke (1997). However, any estimator of the marginal density of data applicable to the problem might be used (see Miazhynskaia & Dorffner, 2006, who review the estimators of marginal density of data for univariate GARCH models). The Bayesian comparison of bivariate GARCH models using Bayes factors was presented by Osiewalski & Pipień (2004).

The Harmonic Mean estimator was presented by Newton & Raftery (1994), and it estimates the marginal distribution of data by a harmonic mean of the values of the likelihood function evaluated at the draws from the posterior distribution of the parameters:

$$\hat{p}(\mathbf{y}|\mathcal{M}_i) = \left[S^{-1} \sum_{s=1}^{S} \frac{1}{p(\mathbf{y}|\theta^{(s)}, \mathcal{M}_i)} \right]^{-1}. \tag{26}$$

The Modified Harmonic Mean estimator of the marginal density of data presented by Geweke (1997) is based on the Harmonic Mean estimator. The variance of the estimator of Newton & Raftery is, however, unbounded for the case in which the prior distribution of parameters has thicker tails than the posterior distribution (see Frühwirth-Schnatter, 2004). The modification proposed by Geweke (1997) bounds this variance. The estimator is then computed, using the very simple formula:

$$\hat{p}(\mathbf{y}|\mathcal{M}_i) = \left[S^{-1} \sum_{s=1}^{S} \frac{f(\theta^{(s)})}{p(\mathbf{y}|\theta^{(s)}, \mathcal{M}_i) p(\theta^{(s)}|\mathcal{M}_i)} \right]^{-1}, \tag{27}$$

where $f(\theta^{(s)})$ is a multivariate truncated normal distribution, with the mean vector set to posterior mean and covariance matrix set to posterior covariance matrix. The truncation is set such that $f(\theta^{(s)})$ have thinner tails than the posterior distribution.

The estimator of Chib & Jeliazkov (2001) is based on the logarithm of the so-called basic marginal likelihood identity:

$$\log \hat{p}(\mathbf{y}|\mathcal{M}_i) = \log \hat{p}(\mathbf{y}|\theta^*, \mathcal{M}_i) + \log \hat{p}(\theta^*|\mathcal{M}_i) - \log \hat{p}(\theta^*|\mathbf{y}, \mathcal{M}_i), \tag{28}$$

where θ^* is some value of parameter vector, e.g. a posterior mean. The first two elements on the right-hand side of the equation (28) are the logarithms of the ordinates of the likelihood function and the prior distribution, and are easy to evaluate. The crucial third element is the logarithm of the posterior ordinate which, due to the unknown form of the posterior density (it is known up to a normalising constant), requires a special procedure. In order to evaluate the posterior density at θ^*, Chib & Jeliazkov (2001) propose the following estimator:

$$\hat{p}(\theta^*|\mathbf{y}, \mathcal{M}_i) = \frac{S^{-1} \sum_{s=1}^{S} \alpha(\theta^{(s)}, \theta^*|\mathbf{y}, \mathcal{M}_i) q(\theta^{(s)}, \theta^*|\mathbf{y}, \mathcal{M}_i)}{J^{-1} \sum_{j=1}^{J} \alpha(\theta^*, \theta^{(j)}|\mathbf{y}, \mathcal{M}_i)}, \tag{29}$$

where $\alpha(\theta^{(s)}, \theta^*|\mathbf{y}, \mathcal{M}_i)$ and $q(\theta^{(s)}, \theta^*|\mathbf{y}, \mathcal{M}_i)$ are as defined earlier in this section, $\{\theta^{(s)}\}_{s=1}^{S}$ is a sample drawn from the posterior distribution and $\{\theta^{(j)}\}_{j=1}^{J}$ is drawn from $q(\theta^*, \theta^{(s)}|\mathbf{y}, \mathcal{M}_i)$.

To summarise, in order to test the second-order noncausality hypotheses with GARCH models, the investigator is advised to follow these steps:

Step 1: Specify the VAR-ECCC-GARCH model Choose the order of the VAR(p) process such that the linear relations between variables are modeled. Choose the order of the ECCC-GARCH(q, r) process with Bayes factors. Estimate the unrestricted model.

Step 2: Set the restrictions For the chosen model and for the hypotheses of interest, derive the restrictions on the parameters of the model using the determinant condition from Theorem 1.

Step 3: Estimate restricted models Using the Metropolis-Hastings algorithm, estimate all the models that represent the hypotheses of interest.

Step 4: Test the hypotheses of noncausality Estimate the marginal densities of data for all m models ($m-1$ restricted models and the unrestricted one). Compute the Bayes factors to compare the models. Chose the one that is best supported by the data model/hypothesis.

To test the existence of volatility spillovers between the variables, modify the four steps by setting the restrictions for appropriate hypothesis.

Discussion. The proposed approach to testing the second-order noncausality hypothesis for GARCH models has several appealing features. First of all, the proposed Bayesian testing procedure makes testing of the parameter conditions possible. It avoids the singularities that may appear in classical tests, in which the restrictions imposed on the parameters are nonlinear. Note that Bayesian tests have not been applied to testing of the second-order noncausality conditions so far.

Secondly, since the competing hypotheses are compared with Bayes factors, they are treated symmetrically. Thanks to the interpretation of the Bayes factors coming from the Posterior Odds Ratio, the outcome of the test is a positive argument *in favour of* the most likely *a posteriori* hypothesis. Moreover, contrary to classical testing, a choice is being made between all the competing hypotheses at once, not only between the unrestricted and one of the restricted models (see Hoogerheide, van Dijk & van Oest R.D., 2009, for a discussion of the argument).

Further, as the testing outcome is based on the posterior analysis, the inference has an exact finite sample justification. Thus, there is no need to refer to the asymptotic theory. In consequence, the assumptions required to test the restrictions may now be relaxed. In order to test the second-order noncausality hypothesis, the assumptions 1–5 must hold. They require the existence of the fourth-order unconditional moment that is ensured by the restrictions derived by He & Teräsvirta (2004). This is a significant improvement, in comparison with the result of Comte & Lieberman (2003). There, the asymptotic distribution of the QMLE is established under the existence of the eighth-order moments. Still, the asymptotic distribution of the Wald test statistic proposed by Comte & Lieberman (2000) and Hafner & Herwartz (2008b) for the original and nonlinear restrictions is unknown (as it is under this condition). Note that the Bayesian testing in the proposed form may be applied to BEKK-GARCH and vec-GARCH models without any complications, and while preserving all the advantages.

For the testing of volatility spillovers, the assumption may be further relaxed. Here, the strict assumption for the linear theory for noncausality of Florens & Mouchart (1985) need not hold. In fact, for testing the zero restrictions for the no volatility spillovers hypothesis, the only required assumption about the moments of the process is that the conditional variances must exist and be bounded. Not even the existence of the second unconditional moments of the process is required. Again, this result is an improvement, in comparison with the test of Nakatani & Teräsvirta (2009), which required the existence of fourth-order moments for the Lagrange Multiplier test statistic to be asymptotically χ^2-distributed.

The improvements in moment conditions are, therefore, established for both kinds of hypothesis. This fact may be crucial for the testing of the hypotheses on the financial time series. In multiple applied studies, such data are shown to have the distribution of the residual term, with thicker tails than those of the normal distribution. Then, distributions modeling this property, such as the t distribution function, are employed. I follow this methodological finding, assuming exactly this distribution function.

As the main costs of the proposed approach, I name the necessity to estimate of all the unrestricted and restricted models. This simply requires some time-consuming computations. While bivariate GARCH models may (depending on the order of the process, and thus on the number of the parameters) be estimated reasonably quickly, trivariate models require significant amounts of time and computational power.

4. Granger causal analysis of exchange rates

The restrictions derived in Section 2 for second-order noncausality for GARCH models, along with the Bayesian testing procedure described in Section 3, are now used in an analysis of the bivariate system of two exchange rates.

Data. The system under consideration consists of daily exchange rates of the British pound (GBP/EUR) and the US dollar (USD/EUR), both denominated in Euro. I analyze logarithmic rates of return expressed in percentage points, $y_{it} = 100(\ln x_{it} - \ln x_{it-1})$ for $i = 1, 2$, where x_{it} are levels of the assets. The data spans the period from 16 September 2008 to 22 September 2011, which gives $T = 777$ observations, and 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 data set contains the two most liquid exchange rates in the Eurozone. The chosen period of analysis starts just after an event that had a very strong impact on the turmoil in the financial markets; the bankruptcy of Lehman Brothers Holding Inc. The proposed analysis of the second-order causality between the series may, therefore, be useful for financial institutions as well as public institutions located in the Eurozone

Table 1: Data: summary statistics

	GBP/EUR	USD/EUR
Mean	0.012	-0.006
Median	0.011	0.016
Standard Deviation	0.707	0.819
Minimum	-2.657	-4.735
Maximum	3.461	4.038
Excess kurtosis	2.430	2.683
Excess kurtosis (robust)	0.060	0.085
Skewness	0.344	-0.091
Skewness (robust)	0.010	-0.016
LJB test	206.525	234.063
LJB p-value	0.000	0.000
T	777.0	777.0

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).

whose performance depends on the forecast of exchange rates. Such institutions include the governments of the countries belonging to the Eurozone that keep their debts in currencies, mutual funds and banks, and all the participants of the exchange rates market.

Figure B.1 from Appendix B plots the time series. It clearly shows the first period of length – of nearly a year – which may be characterized by the high level of volatility of the exchange rates. The subsequent period is characterized by a slightly lower volatility for both of the series. The evident heteroskedasticity, as well as the volatility clustering, seem to provide a strong argument in favor of the employment of the GARCH models that are capable of modeling such features of the data.

Table 1 reports the summary statistics of the two considered series. Both of the series of the rates of returns have sample means and medians close to zero. The US dollar has a slightly larger sample standard deviation than the British pound. Both the series are leptokurtic, which is evidenced by the excess kurtosis coefficient of around 2.5. The pound is slightly positively and the dollar slightly negatively skewed. Neither series can be well described with a normal distribution function. These features of the series seem to confirm the choice of a Student-t likelihood function, which, however, neglects the small skewness of the series.

I have chosen four different models for estimation and performance of the inference about the second-order noncausality between the series. The first unrestricted model is the VAR(1)-ECCC-GARCH(1,1) model defined with equations (2), (3) and (4). Its likelihood function is defined in equation (6) and the prior distributions are specified as in equation (12). The three remaining models are nested within the unrestricted model, according the restrictions on the parameters reflecting the hypotheses of second-order noncausality. Then, the models and, therefore, the hypotheses are compared, using the marginal distributions of the sample data and the Bayes factors. All the models are estimated with the Metropolis-Hastings algorithm, as described in Section 2. However, in order to obtain good properties of the sample drawn from the posterior distribution of the parameters, I keep every 100th draw of the original Markov chain in the final sample which is further analyzed. Also, in order to assure the positivity of the conditional variances, I use the usual solution, assuming that all the parameters of the GARCH process are non-negative. This assumption is, however, stricter than the conditions derived by Conrad & Karanasos (2010).

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

	Vector Autoregression				
	α_0	α_1			
GBP/EUR	0.0091	0.0786	-0.0213		
	(0.0211)	(0.0391)	(0.0330)		
USD/EUR	0.0067	0.0606	-0.0029		
	(0.0261)	(0.0487)	(0.0413)		
	GARCH(1,1)				
	ω	A		B	
GBP/EUR	0.0059	0.0411	0.0154	0.8290	0.0659
	(0.0045)	(0.0170)	(0.0087)	(0.0643)	(0.0447)
USD/EUR	0.0539	0.0954	0.0598	0.8451	0.2015
	(0.0411)	(0.0529)	(0.0353)	(0.2472)	(0.1753)
	Degrees of freedom and correlation				
	ν	ρ_{12}			
	14.9316	0.4101			
	(1.5495)	(0.0317)			

The table summarises the estimation of the VAR(1)-ECCC-GARCH(1,1) model described by the equations (2), (3), (4) and the likelihood function (6). The prior distributions are specified as in equation (12). The parameters are grouped in the matrices. The posterior means and the posterior standard deviations (in brackets) are reported.

Table 2 reports the posterior means and the posterior standard deviations of the parameters of the unrestricted model. When the parameters of the GARCH process are analyzed, it is striking that the British pound impacts on the volatility of the US dollar at a great rate. This is evidenced by the values of the posterior means of the parameters A_{21} and B_{21}. In particular, the value of the posterior mean of the latter parameter, which is equal to 0.845, with the posterior standard deviation equal to 0.247, shows that the previous conditional variance of the British pound impacts on the current conditional variance of the US dollar at a rate much higher than the previous conditional variance of the dollar. The case is the same for the impact of the previous squared innovations of the pound and the dollar on the current conditional variance of the US dollar. This phenomenon is further analyzed and interpreted in subsequent parts of this section.

The properties of the numerical simulation of the posterior distribution of the parameters are reported in Table C.4 in Appendix C. In the table, we read that the sample of draws of the parameters of the vector autoregression and the GARCH processes have good properties and are efficiently drawn from the stationary posterior distribution of the parameters.

Table C.5 of Appendix C presents the results of the parameter estimation for all four models. The high values of the posterior means of the parameters A_{21} and B_{21} are reproduced in model $\mathcal{M}_2$, representing the hypothesis of second-order noncausality from dollar to pound. These two parameters are, in models $\mathcal{M}_3$ and $\mathcal{M}_4$, set to zero in order to represent the hypothesis of second-order noncausality from pound to dollar. Note that in these two cases, when the previous conditional variance of the pound does not impact on the current conditional variance of the dollar by assumption, the value of the parameter B_{22} changes. It increases from a level of around 0.2 for the models $\mathcal{M}_1$ and $\mathcal{M}_2$, which allow for volatility spillovers from pound to dollar, to a level of around 0.93 for the models $\mathcal{M}_3$ and $\mathcal{M}_4$, in which this spillover is not allowed. This finding illustrates an interesting pattern in the parameters responsible for modeling the persistence in conditional variances in models in which volatility spillovers may be allowed or disallowed.

The properties of the numerical simulation of the posterior distributions of all the models are reported in Table C.6 in Appendix C. All the models are reasonably well estimated.

Finally, Table 3 reports the results of the testing of the hypotheses. In the first row, the models and the hypotheses they represent are stated. Refer to Table C.5 for the restrictions imposed on the parameters of the models. In order to test the second-order noncausality hypotheses, the marginal densities of data for each model were computed, using the three estimators described in Section 3. The first panel of Table C.6 reports their natural logarithms. The second panel reports the logarithms of basis 10 of the Bayes factors, comparing all models to the best one, according to the estimator of the marginal density of data. The best model has the biggest value for the marginal density of data, and it also has the value of the logarithm of the Bayes factor equal to zero, as in both numerator and the denominator there is the same value of the marginal density of data.

Despite the fact that the results are not fully robust to the chosen estimator of the marginal density of data some, patterns are clearly visible. First of all, the data favors the models that allow for second-order causality from GBP/EUR to USD/EUR. According to two of the estimators, models $\mathcal{M}_1$ and $\mathcal{M}_2$ are the best. Still, for the estimator of Geweke, the model $\mathcal{M}_2$ is the most probable *a posteriori*, whereas the model $\mathcal{M}_1$ is ranked third. Secondly, models that represent the hypothesis that GBP/EUR does not second-order Granger-cause USD/EUR are rejected by the data. The models $\mathcal{M}_3$ and $\mathcal{M}_4$ have the lowest values of the Bayes factors.

Table 3: Summary of the second-order noncausality testing: marginal densities of data and Bayes factors

Estimator	$\mathcal{M}_1$ unrestricted	$\mathcal{M}_2$ $y_2 \overset{so}{\nrightarrow} y_1$	$\mathcal{M}_3$ $y_1 \overset{so}{\nrightarrow} y_2$	$\mathcal{M}_4$ $y_2 \overset{so}{\nrightarrow} y_1$ and $y_1 \overset{so}{\nrightarrow} y_2$
	Marginal density of data, $\ln p(\mathbf{y}\vert\mathcal{M}_i)$:			
NR1994	-1595.1312	-1594.9256	-1598.4017	-1595.5036
Geweke1997	-1623.0599	-1617.9192	-1624.0413	-1621.7130
CJ2001	-1613.7263	-1614.8588	-1622.4544	-1615.5204
	Bayes factors, $\log_{10} \mathcal{B}_{ij}$:			
NR1994	-0.0893	0	-1.5097	-0.2510
Geweke1997	-2.2325	0	-2.6588	-1.6476
CJ2001	0	-0.4918	-3.7906	-0.7792

In the first panel, the natural logarithms of the marginal density of data, defined in equation (23), for all the models representing the hypotheses on second-order noncausality, are reported. The three considered estimators are: the Harmonic Mean estimator by Newton & Raftery (1994), defined in (26), the Modified Harmonic Mean estimator by Geweke (1997), defined in (27), and the estimator by Chib & Jeliazkov (2001), defined in (28) and (29). In the second panel, the models (and hypotheses) are compared with the logarithm of basis 10 of the Bayes factor, defined in equation (24). The jth model, from the denominator of the Bayes factor, is always a model of the biggest marginal density of data for each of the estimators. Description of the variables: y_1 - GBP/EUR, y_2 - USD/EUR.

To summarize, the model with most support by the data is model $\mathcal{M}_2$. According to two estimators, this model fits the data best, and according to the estimator of Chib & Jeliazkov, it is the second-best model. It represents the hypothesis that the US dollar does not second-order Granger-cause the British pound. Therefore, it restricts the parameters A_{12} and B_{12} to zero, whereas the parameters A_{21} and B_{21} which are responsible for the second-order causality from pound to dollar, remain unrestricted. Both these parameters have values of the posterior means distant from zero; thus, the second-order causal link is strong in this direction. Note that, according to the Corollary 1, the second-order noncausality from dollar to pound and the second-order causality from pound to dollar are established at all the future horizons. That is the only proper conclusion, given the bivariate system, irrespective of the order of the GARCH process.

However surprising this finding may seem, the main reason behind it seems to be the trading hours. The volatility of GBP/EUR at time t impacts on the volatility of USD/EUR at time $t + 1$, as the information appearing during trading hours in Europe (and thus the exchange between the Eurozone and the United Kingdom) impacts on the volatility of USD/EUR the day after. This may be caused by the fact that the trading hours of the financial markets in the United States that start at the end of the trading hours in Europe. Moreover, this main finding of the testing procedure resembles the recursive systems (triangular matrices A and B) of the *meteor shower* hypothesis of Engle, Ito & Lin

(1990). Although the setting of this work differs from the empirical model of Engle et al. (1990), the pattern of the second-order causal and noncausal relations associated with the non-synchronous trading hours of the financial markets is evident.

5. Conclusions

In this work I have derived the conditions for analyzing the Granger noncausality for the second conditional moments modeled with GARCH processes. The presented restrictions for one period ahead second-order noncausality, due to the specific setting of the system, in which all the considered variables belong to one of the vectors, appear to be the retractions for the second-order noncausality at all future horizons. These conditions may result in several nonlinear restrictions on the parameters of the model, which results in the fact that the available classical tests have limited uses.

Therefore, in order to test these restrictions, I have applied the basic Bayesian procedure that consists of the estimation of the models representing the hypotheses of second-order causality and noncausality, and then of the comparison of the models and hypotheses with Bayes factors. This well-known procedure overcomes the difficulties that the classical tests applied so far to this problem have met. The Bayesian inference about the second-order causality between variables is based on the finite-sample analysis. Moreover, although the analysis does not refer to the asymptotic results, the strict assumptions about the existence of the higher-order moments of the series that are required in the asymptotic analysis may be relaxed in the Bayesian inference. In effect, the existence of fourth unconditional moments is assumed for the second-order noncausality analysis, and of second conditional moments for the volatility spillovers analysis.

The proposed approach has several limitations, however. These come from the fact that all the variables in the system are divided into only two vectors, between which the causality inference is performed. With such a setting, not all the hypotheses of interest may be formulated in the system that contains more than two variables (see Example 2). Another limitation is the fact that the presented restrictions serve as the restrictions for the second-order noncausality at all future horizons at once. This feature is caused by the particular setting considered in this work.

This critique is a motivation for further research on the topic of Granger causality in second conditional moments. First, one might consider the setting in which the causality between two variables is analyzed, when there are also other variables in the system that might be used for the purpose of modeling and forecasting. This may be particularly necessary for the analysis of the robustness of the causal or noncausal relations found, as the values of the parameters in the GARCH models are exposed to the omitted variables problem. Second, the second-order noncausality could be analyzed separately at each of the future horizons. Such a decomposition could provide further insights into causal relations between economic relations. However, the setting considered in this work does not allow for such an analysis.

Appendix A. Proof

Proof of Theorem 1. The first part of the proof sets the second-order noncausality restrictions for the GARCH process in the VAR form (9). Let $\epsilon_t^{(2)}$ follow a stationary VAR process as in (9), partitioned as in (10), that is identifiable. Then, y_1 does not second-order Granger-cause y_2 if and only if:

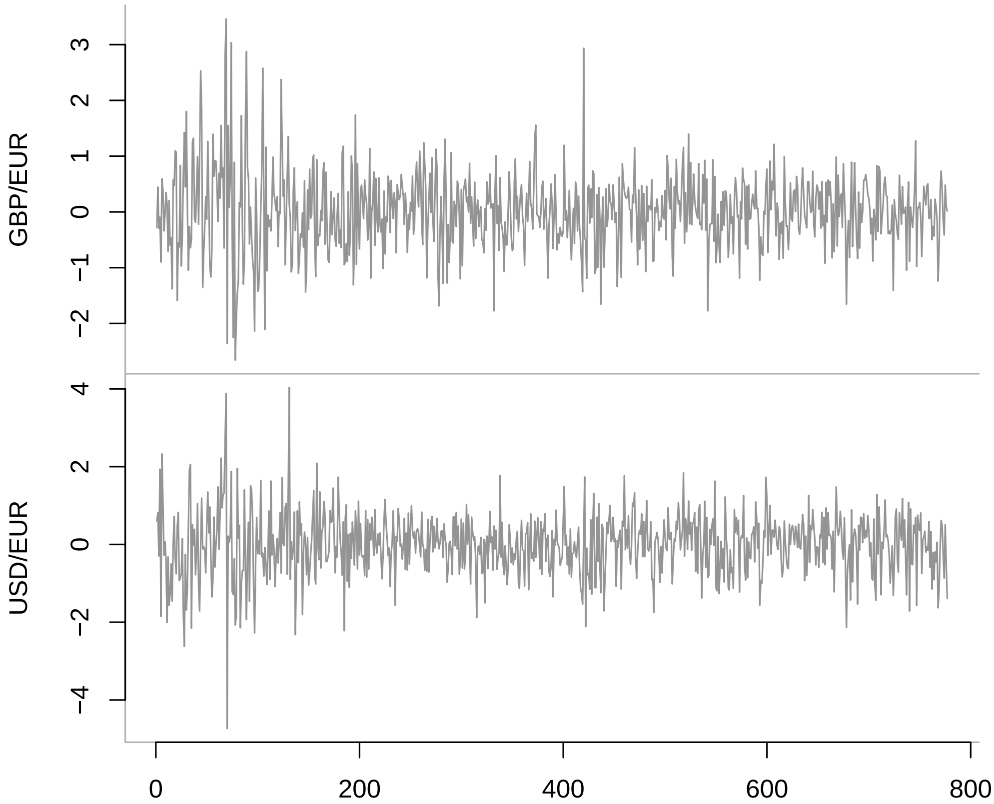
$$\Pi_{21}(z) \equiv 0 \qquad \forall z \in \mathbb{C}. \tag{A.1}$$

Condition (A.1) may be proven by the application of Proposition 1 of Boudjellaba et al. (1992). Several changes are, however, required to adjust the proof of that Proposition for the vector autoregressive process to the setting considered in Theorem 1 for the GARCH models. Here, one projects the squared elements of the residual term, $\epsilon^{(2)}(y_{2t+1}|I(t)) = [y_{2t+1} - P(y_{2t+1}|I(t))]^{(2)}$, on the Hilbert spaces $I^2(t)$ or $I^2_{-1}(t)$, both defined in Section 2.

The proven condition still leads to infinite number of restrictions on parameters. This property excludes the possibility of testing these restrictions. In order to obtain the simplified condition (14), apply to (A.1) the matrix transformations, first of Theorem 1 and then of Theorem 2 of Boudjellaba et al. (1994).

Appendix B. Data

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



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

Appendix C. Results of estimation

Table C.4: Summary of the posterior distribution simulation

		Mean	SD	Autocorrelations at lag 1	Autocorrelations at lag 10	RNE	Geweke's z
	Vector Autoregression						
α_0	$\alpha_{0.1}$	0.0091	0.0211	0.3225	0.0315	0.3803	0.6709
	$\alpha_{0.2}$	0.0067	0.0261	0.2897	-0.0207	0.6216	1.0137
α_1	$\alpha_{1.11}$	0.0786	0.0391	0.2631	0.0046	0.6467	0.1190
	$\alpha_{1.21}$	0.0606	0.0487	0.3581	0.0170	0.3685	1.4173
	$\alpha_{1.12}$	-0.0213	0.0330	0.3104	0.0188	0.4330	-1.6478
	$\alpha_{1.22}$	-0.0029	0.0413	0.4594	0.0278	0.2093	-0.5947
	GARCH(1,1)						
ω	ω_1	0.0059	0.0045	0.1185	0.0001	0.5631	-0.6871
	ω_2	0.0539	0.0411	0.6835	0.1797	0.0718	-1.0395
A	A_{11}	0.0411	0.0170	0.2030	0.0016	0.3986	-0.8394
	A_{21}	0.0954	0.0529	0.4131	0.0473	0.2446	-1.4520
	A_{12}	0.0154	0.0087	0.1437	0.0027	0.6219	-0.3682
	A_{22}	0.0598	0.0353	0.4529	-0.0023	0.2561	-1.2452
B	B_{11}	0.8290	0.0643	0.4259	0.0394	0.1794	-0.4076
	B_{21}	0.8451	0.2472	0.7482	0.1453	0.1032	-1.4820
	B_{12}	0.0659	0.0447	0.4183	0.0423	0.1628	0.8119
	B_{22}	0.2015	0.1753	0.8219	0.2423	0.0755	1.4204
	Degrees of freedom and correlation						
	ν	14.9316	1.5495	0.9832	0.9256	0.0151	6.4738
	ρ_{12}	0.4101	0.0317	0.4244	-0.0028	0.5334	-0.9674

Note: The table reports the posterior mean and the posterior standard deviation of the parameters of the model. Also, autocorrelations at lag 1 and 10 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 simulation, comparing the mean of the first 50% of the draws to the mean of the last 35% of the draws. The z scores follow the standard normal distribution. This method is presented in Geweke (1992). The numbers presented in this table were obtained using the package coda by Plummer et al. (2006).

Table C.5: Summary of the estimation of the VAR(1)-ECCC-GARCH(1,1) models

$\mathcal{M}_1$: *Unrestricted model*										
	α_0	α_1		ω	A		B		ν	ρ_{12}
y_1	0.0091	0.0786	-0.0213	0.0059	0.0411	0.0154	0.8290	0.0659	14.9316	0.4101
	(0.0211)	(0.0391)	(0.0330)	(0.0045)	(0.0170)	(0.0087)	(0.0643)	(0.0447)	(1.5495)	(0.0317)
y_2	0.0067	0.0606	-0.0029	0.0539	0.0954	0.0598	0.8451	0.2015		
	(0.0261)	(0.0487)	(0.0413)	(0.0411)	(0.0529)	(0.0353)	(0.2472)	(0.1753)		
$\mathcal{M}_2 : y_2 \overset{so}{\nrightarrow} y_1$										
	α_0	α_1		ω	A		B		ν	ρ_{12}
y_1	0.0087	0.0741	-0.0194	0.0070	0.0559	0	0.9284	0	11.5226	0.4152
	(0.0208)	(0.0421)	(0.0342)	(0.0048)	(0.0173)		(0.0230)		(3.0436)	(0.0329)
y_2	0.0062	0.0541	0.0066	0.1071	0.1069	0.0639	0.7058	0.2056		
	(0.0267)	(0.0500)	(0.0407)	(0.0611)	(0.0586)	(0.0365)	(0.2377)	(0.1743)		
$\mathcal{M}_3 : y_1 \overset{so}{\nrightarrow} y_2$										
	α_0	α_1		ω	A		B		ν	ρ_{12}
y_1	0.0149	0.0913	-0.0229	0.0317	0.0952	0.0222	0.5444	0.1903	12.1869	0.4000
	(0.0218)	(0.0407)	(0.0333)	(0.0283)	(0.0378)	(0.0190)	(0.2233)	(0.1253)	(3.7379)	(0.0327)
y_2	0.0146	0.0683	-0.0180	0.0102	0	0.0463	0	0.9381		
	(0.0258)	(0.0461)	(0.0399)	(0.0068)		(0.0163)		(0.0227)		
$\mathcal{M}_4 : y_1 \overset{so}{\nrightarrow} y_2$ and $y_2 \overset{so}{\nrightarrow} y_1$										
	α_0	α_1		ω	A		B		ν	ρ_{12}
y_1	0.0120	0.0806	-0.0237	0.0114	0.0672	0	0.9088	0	11.6345	0.4052
	(0.0207)	(0.0424)	(0.0338)	(0.0079)	(0.0222)		(0.0336)		(2.8701)	(0.0342)
y_2	0.0170	0.0570	-0.0137	0.0147	0	0.0513	0	0.9258		
	(0.0261)	(0.0447)	(0.0377)	(0.0086)		(0.0172)		(0.0262)		

Note: The table summarizes the estimation of the VAR(1)-ECCC-GARCH(1,1) models described by the equations (2), (3), (4) and the likelihood function (6). The prior distributions are specified as in equation (12). The parameters are grouped in the matrices. The posterior means and the posterior standard deviations (in brackets) are reported.
The restrictions for second-order noncausality (denoted by $\overset{so}{\nrightarrow}$) are imposed according to the determinant condition (**??**) for the hypotheses stated at the top of the panel for each model. Description of the variables: y_1 - GBP/EUR, y_2 - USD/EUR.

Table C.6: Summary of the properties of the simulations of the posterior densities of the models

Model	RNE			Autocorrelation at lag 1			Autocorrelation at lag 10			Geweke's z			S
	median	min	max	median	min	max	median	min	max	median	min	max	
$\mathcal{M}_1$	0.312	0.015	0.647	0.416	0.119	0.983	0.023	-0.021	0.926	0.062	-1.906	12.342	4226
$\mathcal{M}_2$	0.639	0.228	0.922	0.188	0.124	0.472	-0.004	-0.050	0.045	-0.279	-2.260	1.412	2800
$\mathcal{M}_3$	0.683	0.225	1.008	0.208	0.089	0.571	0.004	-0.017	0.026	0.708	-1.551	2.527	3000
$\mathcal{M}_4$	0.081	0.055	0.107	0.767	0.731	0.859	0.152	0.078	0.265	-0.026	-2.307	1.904	3000

Note: The table summarizes the properties of the numerical simulation of the posterior densities of all the considered models. For each of the statistics, the median of all the parameters of the model, as well as minimum and maximum, are reported. The table reports the relative numerical efficiency coefficient (RNE), by Geweke (1989), autocorrelations of the MCMC draws at lags 1 and 10, as well as Geweke's z scores for the hypothesis of equal means of the first 10% and the last 50% of draws that follow the standard normal distribution (see Geweke, 1992). The numbers presented in this table were obtained using the package coda by Plummer et al. (2006).

References

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

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.

Cheung, Y. W., & Ng, L. K. (1996). A causality-in-variance test and its application to financial market prices. *Journal of Econometrics*, *72*, 33–48.

Chib, S., & Greenberg, E. (1995). Understanding the Metropolis-Hastings Algorithm. *The American Statistician*, *49*, 327.

Chib, S., & Jeliazkov, I. (2001). Marginal Likelihood from the Metropolis-Hastings Output. *Journal of the American Statistical Association*, *96*, 270– 281.

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. (2010). Negative Volatility Spillovers in the Unrestricted ECCC-GARCH Model. *Econometric Theory*, *26*, 838–862.

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

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

Dufour, J.-M., & Renault, E. (1998). Short Run and Long Run Causality in Time Series : Theory. *Econometrica*, *66*, 1099–1125.

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.

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 & Economic Statistics*, *21*, 532–546.

Florens, J. P., & Mouchart, M. (1985). A Linear Theory for Noncausality. *Econometrica*, *53*, 157–176.

Frühwirth-Schnatter, S. (2004). Estimating marginal likelihoods for mixture and Markov switching models using bridge sampling techniques. *Econometrics Journal*, *7*, 143–167.

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.

Geweke, J. (1995). Bayesian Comparison of Econometric Models. *Federal Reserve Bank of Minneapolis*, (pp. 1–36).

Geweke, J. (1997). Posterior Simulators in Econometrics. In *Advances in economics and econometrics: theory and applications* chapter Posterior. Cambridge University Press volume 3 of *Econometric Society Monographs*.

Geweke, J. (1999). Using simulation methods for bayesian econometric models: inference, development,and communication. *Econometric Reviews*, *18*, 1–73.

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, *37*, 424–438.

Hafner, C. M., & Herwartz, H. (2006). A Lagrange multiplier test for causality in variance. *Economics Letters*, *93*, 137–141.

Hafner, C. M., & Herwartz, H. (2008a). Analytical quasi maximum likelihood inference in multivariate volatility models. *Metrika*, *67*, 219–239.

Hafner, C. M., & Herwartz, H. (2008b). 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.

Hoogerheide, L. F., van Dijk, H. K., & van Oest R.D. (2009). Simulation Based Bayesian Econometric Inference: Principles and Some Recent Computational Advances. In *Handbook of Computational Econometrics* Handbook of Computational Econometrics chapter 7. (pp. 215–280). Wiley.

Jarociński, M., & Maćkowak, 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.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773–795.

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

Ling, S., & McAleer, M. (2003). Asymtotic 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. (1993). *Introduction to Multiple Time Series Analysis*. Springer-Verlag.

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.

Miazhynskaia, T., & Dorffner, G. (2006). A comparison of Bayesian model selection based on MCMC with an application to GARCH-type models. *Statistical Papers*, *47*, 525–549.

Nakatani, T., & Teräsvirta, T. (2009). Testing for volatility interactions in the Constant Conditional Correlation GARCH model. *Econometrics Journal*, *12*, 147–163.

Newton, M. A., & Raftery, A. E. (1994). Approximate Bayesian Inference with the Weighted Likelihood Bootstrap. *Journal of the Royal Statistical Society. Series B (Methodological)*, *56*, 3–48.

Osiewalski, J., & Pipień, M. (2004). Bayesian comparison of bivariate ARCH-type models for the main exchange rates in Poland. *Journal of Econometrics*, *123*, 371–391.

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.

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