



## Granger-causal analysis of GARCH models: A Bayesian approach

Tomasz Woźniak

To cite this article: Tomasz Woźniak (2018) Granger-causal analysis of GARCH models: A Bayesian approach, *Econometric Reviews*, 37:4, 325-346, DOI: [10.1080/07474938.2015.1092839](https://doi.org/10.1080/07474938.2015.1092839)

To link to this article: <https://doi.org/10.1080/07474938.2015.1092839>



Accepted author version posted online: 13 Oct 2015.  
Published online: 09 Apr 2016.



Submit your article to this journal [↗](#)



Article views: 148



View related articles [↗](#)



View Crossmark data [↗](#)

# Granger-causal analysis of GARCH models: A Bayesian approach

Tomasz Woźniak 

Department of Economics, The University of Melbourne, Carlton, Victoria, Australia

## ABSTRACT

A multivariate GARCH model is used to investigate Granger causality in the conditional variance of time series. 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, are derived. These novel conditions are convenient for the analysis of potentially large systems of economic variables. To evaluate hypotheses of noncausality, a Bayesian testing procedure is proposed. It avoids the singularity problem that may appear in the Wald test, and it relaxes the assumption of the existence of higher-order moments of the residuals required in classical tests.

## KEYWORDS

Bayesian hypotheses assessment; second-order noncausality; VAR-GARCH models

## 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. This study investigates 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 and Lieberman, 2000; Robins et al., 1986). If one variable does not second-order Granger-cause the other, then past information about the variability of the former is dispensable for the forecasting of the conditional variances of the latter. If Granger causality in the conditional mean is investigated at the same time, then *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 (VAR) 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, due to observed lack of autocorrelations in returns of financial assets the integration of the financial markets (assets) is commonly understood as dynamic cross-dependencies in conditional variances. 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 asset allocation, financial risk management, and option pricing.

Second-order causality is investigated within a popular Extended CCC-GARCH model by Jeantheau (1998), Ling and McAleer (2003), and He and Teräsvirta (2004), also analysed in Karolyi (1995), Conrad and Karanasos (2010), Nakatani and Teräsvirta (2009), and others.

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 market dynamics that exhibit volatility persistence due to private information or heterogeneous beliefs (see Hong, 2001, and references therein). Finally, the *meteor showers* hypothesis for intradaily exchange rates returns, which reflects cooperative or competitive monetary policies (see Engle et al., 1990), can be presented as a sufficient condition for second-order Granger 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 et al. (1994) use the term to describe international transmissions between stock returns and their volatilities. Further, Nakatani and Teräsvirta (2009) and Koutmos and 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 and Karanasos, 2010), as well as in others. However, parameters referred to in this way do not determine Granger causality or noncausality themselves. This study presents parameter conditions for the precisely defined Granger noncausality concept for conditional variances. In particular, the framework of the linear Granger noncausality of Florens and Mouchart (1985) is referred to, and which defined the noncausality relationship in terms of the orthogonality in the Hilbert space of square integrable variables<sup>1</sup>.

The contribution of this study is twofold. Firstly, conditions for second-order Granger noncausality for a family of GARCH models are derived. 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 and Lieberman, 2000; Hafner and Herwartz, 2008; Woźniak, 2015). 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, a Bayesian testing procedure of the conditions for Granger noncausality in the conditional variance processes is proposed. It is easily applicable and solves some of the drawbacks of the classical testing. In comparison with the Wald test of Boudjellaba et al. (1992), adapted to testing noncausality relations in the 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.

The reminder of this article is organized as follows. The notation and the GARCH model used in the analysis is set in Section 2. Section 3 presents the main theoretical findings of the article, deriving the conditions for Granger noncausality in the conditional variance process. Section 4 discusses classical testing for noncausality in the VAR-GARCH models, and then proposes 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.

<sup>1</sup> Among other possible approaches one could investigate conditional independence proposed by Haugh (1976) and applied to conditional variances by Cheung and Ng (1996) and Hong (2001), or Bayesian exogeneity as defined by Florens and Mouchart (1977, 1980) and applied by Pajor (2011) to the analysis of multivariate Stochastic Volatility models.

## 2. Notation and the GARCH model

**Notation.** 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  $\mathbb{Z}$ . Write

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

where  $y_{it}$  is a  $N_i \times 1$  vector such that  $y_{1t} = (y_{1t}, \dots, y_{N_1,t})'$ ,  $y_{2t} = (y_{N_1+1,t}, \dots, y_{N_1+N_2,t})'$ , and  $y_{3t} = (y_{N_1+N_2+1,t}, \dots, y_{N_1+N_2+N_3,t})'$  ( $N_1, N_2 \geq 1, N_3 \geq 0$  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  $y_\tau$ , 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, where  $P(y_{t+h} | I(t))$  denotes the affine projection of  $y_{t+h}$  on  $I(t)$ , i.e., the best linear prediction of  $y_{t+h}$ , based on the variables in  $I(t)$  and a constant term.

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  $(y'_{2\tau}, y'_{3\tau})'$ .  $I_{y,-1}^2$  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  $I_\epsilon^2(t)$  generated by the variables  $\epsilon_{i\tau}\epsilon_{j\tau}$ , where  $N_1+1 \leq i, j \leq N$  and for  $\tau \leq t$ .

**GARCH(r,s) Model and Its Properties.** Suppose that  $y_t$  follows a  $N$ -dimensional VARMA(p,q) process:

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

for all  $t = 1, \dots, T$ , where  $L$  is a lag operator such that  $L^i y_t = y_{t-i}$ ,  $\alpha(z) = I_N - \alpha_1 z - \dots - \alpha_p z^p$ ,  $\beta(z) = I_N + \beta_1 z + \dots + \beta_q z^q$  are matrix polynomials.  $I_N$  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 and Lieberman (2000) states that Granger noncausality in the mean can be analyzed also if  $E[\epsilon_t \epsilon_t' | I_\epsilon^2(t-1)] = H_t$ , i.e., if the conditional covariance matrix of  $\epsilon_t$  is time-varying, provided that the unconditional covariance matrix,  $E[H_t] = \mathbf{V}$ , is constant and nonsingular. Without loss of generality, it is 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. The process in Eq. (2) is assumed to be stationary and identified.

The residual term  $\epsilon_t$  follows a conditional variance process:

$$\epsilon_t = D_t r_t, \quad (3a)$$

$$r_t \sim i.i.d.(0, \mathbf{C}), \quad (3b)$$

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 follow independent and identically distributed (i.i.d.) with zero mean and a correlation matrix  $\mathbf{C}$ .

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

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

for all  $t = 1, \dots, T$ , where  $h_t$  is a  $N \times 1$  vector of conditional variances of  $\epsilon_t$ ,  $\omega$  is a  $N \times 1$  vector of constant terms,  $\epsilon_t^{(2)} = (\epsilon_{1t}^2, \dots, \epsilon_{Nt}^2)'$  is a vector of squared residuals, and  $A(L) = \sum_{i=1}^r A_i L^i$  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  $\epsilon_t$  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  $\nu > 2$  degrees of freedom, the conditional variances exist and are  $E[\epsilon_t^{(2)} | I_y^2(t-1)] = \frac{\nu}{\nu-2} h_t$ . 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), (3), and (4), which is the 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 Eqs. (3) and (4) with diagonal matrices  $A(L)$  and  $B(L)$ . Its extended version, with nondiagonal matrices  $A(L)$  and  $B(L)$ , was used in Karolyi (1995) and analyzed by Jeantheau (1998). He and Teräsvirta (2004) called this model extended CCC-GARCH (ECCC-GARCH).

Jeantheau (1998) proves that the ECCC-GARCH(r,s) model, as in (4), 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 and Teräsvirta (2004) give sufficient conditions for the existence of the fourth moments and derive complete structure of these moments.

The VARMA-ECCC-GARCH model has well established asymptotic properties. They can be set under the following assumptions.

#### Assumption 1.

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

Under Assumptions 1.1 and 2, 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 and McAleer (2003) proved the strong consistency of the Quasi maximum likelihood estimator (QMLE) for the VARMA-GARCH model under Assumptions 1 and 2 and these of stationarity and identifiability of the VARMA process. 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 the conditional covariance matrix,  $H_t$ ,  $h_t$  has to be positive for all  $t$ . Usual sufficient parameter conditions for  $h_t$  to be positive proposed by Bollerslev (1990) are  $\omega > 0$  and  $[A_i]_{jk}, [B_l]_{jk} \geq 0$  for  $i = 1, \dots, r$ ,  $l = 1, \dots, s$ , and  $j, k = 1, \dots, N$ . Conrad and Karanasos (2010) derived conditions such that some elements of  $A_i$ ,  $B_l$  ( $i = 1, \dots, r$ ;  $l = 1, \dots, s$ ) and even  $\omega$  are allowed to be negative. Still, it is not known whether asymptotic results hold under these conditions. However, their empirical usefulness has been proven, as Conrad and Karanasos (2010) have found that some parameters of the model responsible for volatility transmissions are negative. For a feasible Bayesian estimation the sufficient conditions of Bollerslev (1990) are used in the empirical example of the current paper.

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  $\epsilon$ s are, respectively,

$$L_N(\theta; \mathbf{y}) = (2\pi)^{-TN/2} \prod_{t=1}^T |H_t|^{-1/2} \exp(\epsilon_t' H_t^{-1} \epsilon_t)^{-1/2}, \quad \text{and} \quad (5a)$$

$$L_{St}(\theta; \mathbf{y}) = \prod_{t=1}^T \frac{\Gamma(\frac{\nu+N}{2})}{\Gamma(\frac{\nu}{2})} ((\nu-2)\pi)^{-\frac{N}{2}} |H_t|^{-\frac{1}{2}} \left(1 + \frac{1}{\nu-2} \epsilon_t' H_t^{-1} \epsilon_t\right)^{-\frac{\nu+N}{2}}, \quad (5b)$$

where  $\epsilon_t$  is defined in Eqs. (2) and (3).  $\Gamma(\cdot)$  is Euler's gamma and  $|\cdot|$  a matrix determinant. Algorithms maximizing the likelihood function, such as the algorithm by (see Berndt et al., 1974), use analytical derivatives. Fiorentini et al. (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 et al. (2003) propose a Metropolis–Hastings algorithm (see Chib and Greenberg, 1995, and references therein) for the estimation of the model.

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. For instance, in a recent study Ben Omrane and Hafner (2009) use the trivariate model for volatility spillovers between exchange rates. Conrad and Karanasos (2010) and Nakatani and 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 and 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  $B(L)$  nondiagonal) assumes the existence of sixth-order moments of the residual term,  $E|\epsilon_t^6| < \infty$ . Woźniak (2015) introduces the notion of Granger second-order causality and causality in variance for ECCC-GARCH models for the setting similar to that of Comte and Lieberman (2000), in which the vector of variables is partitioned in two parts. The current article extends 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 the notion of Granger noncausality for conditional variances is presented, the GARCH(r,s) model, (4), is rewritten into its VARMA and VAR representations. Define a process  $v_t = \epsilon_t^{(2)} - h_t$ . Then  $\epsilon_t^{(2)}$  follows a VARMA process given by

$$\phi(L)\epsilon_t^{(2)} = \omega + \psi(L)v_t, \quad (6)$$

for all  $t = 1, \dots, T$ , 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 (6) can be written in the form

$$\begin{bmatrix} \phi_{11}(L) & \phi_{12}(L) & \phi_{13}(L) \\ \phi_{21}(L) & \phi_{22}(L) & \phi_{23}(L) \\ \phi_{31}(L) & \phi_{32}(L) & \phi_{33}(L) \end{bmatrix} \begin{bmatrix} \epsilon_{1t}^{(2)} \\ \epsilon_{2t}^{(2)} \\ \epsilon_{3t}^{(2)} \end{bmatrix} = \begin{bmatrix} \omega_{1t} \\ \omega_{2t} \\ \omega_{3t} \end{bmatrix} + \begin{bmatrix} \psi_{11}(L) & \psi_{12}(L) & \psi_{13}(L) \\ \psi_{21}(L) & \psi_{22}(L) & \psi_{23}(L) \\ \psi_{31}(L) & \psi_{32}(L) & \psi_{33}(L) \end{bmatrix} \begin{bmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{bmatrix}. \quad (7)$$

Given Assumption 1.2, the VARMA process (6) is invertible and can be written in the VAR form

$$\Pi(L)\epsilon_t^{(2)} - \omega^* = v_t, \quad (8)$$

for all  $t = 1, \dots, T$ , 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, rewrite (8) in the form

$$\begin{bmatrix} \Pi_{11}(L) & \Pi_{12}(L) & \Pi_{13}(L) \\ \Pi_{21}(L) & \Pi_{22}(L) & \Pi_{23}(L) \\ \Pi_{31}(L) & \Pi_{32}(L) & \Pi_{33}(L) \end{bmatrix} \begin{bmatrix} \epsilon_{1t}^{(2)} \\ \epsilon_{2t}^{(2)} \\ \epsilon_{3t}^{(2)} \end{bmatrix} - \begin{bmatrix} \omega_{1t}^* \\ \omega_{2t}^* \\ \omega_{3t}^* \end{bmatrix} = \begin{bmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{bmatrix}. \quad (9)$$

Under Assumption 1, both processes (6) and (8) are stationary.

### 3. Restrictions for second-order noncausality

This section presents the main theoretical findings of the article, that is the derivation of the conditions for second-order Granger noncausality for the ECCC-GARCH model. Two concepts are defined: Granger noncausality in variance and second-order Granger noncausality. Further, the parametric conditions in Theorems 1 and 2 are derived and their novelty is discussed.

Robins et al. (1986) introduced the concept of Granger causality for conditional variances. Comte and Lieberman (2000) call this concept *second-order Granger causality* and distinguish it from *Granger causality in variance*. In the current article, we define these noncausalities slightly differently than Comte and Lieberman (2000). In the definition for second-order noncausality below, the information set  $I_\epsilon^2(t)$  is used, whereas Comte and Lieberman use  $I_y^2(t)$ . Our choice easily allows investigating the causality on different levels (first and/or second conditional moments) and is easily implementable for the ECCC-GARCH model. The definitions are in the following forms.

**Definition 1.**  $y_1$  does not second-order Granger-cause  $y_2$  given  $y_3$ , denoted by  $y_1 \xrightarrow{so} y_2 | y_3$ , if

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

**Definition 2.**  $y_1$  does not Granger-cause  $y_2$  in variance given  $y_3$ , denoted by  $y_1 \xrightarrow{v} y_2 | y_3$ , if

$$P([y_{2t+1} - P(y_{2t+1} | I(t))]^{(2)} | I_y^2(t)) = P([y_{2t+1} - P(y_{2t+1} | I_{-1}(t))]^{(2)} | I_{y,-1}^2(t)) \quad \forall t \in \mathbb{Z}, \quad (10)$$

where  $[\cdot]^{(2)}$  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 1, we take the affine projection of  $y_{2t+1}$  on  $I(t)$ , whereas on the right-hand side of Definition 2 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  $y_2$ ,  $y_1 \xrightarrow{G} y_2 | y_3$ . The relation between Granger noncausality in mean, noncausality in variance, and second-order noncausality have been established by Comte and Lieberman (2000) and are as follows:

$$y_1 \xrightarrow{v} y_2 | y_3 \Leftrightarrow (y_1 \xrightarrow{G} y_2 | y_3 \text{ and } y_1 \xrightarrow{so} y_2 | y_3). \quad (11)$$

One implication of this statement is that Definitions 1 and 2 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 and 2, the ECCC-GARCH model is stationary, identifiable, and invertible. One more assumption is needed in order to state noncausality relations in the conditional variances process:

**Assumption 3.** The process  $v_t$  is covariance stationary with covariance matrix  $V_v$ .

A theorem that introduces second-order Granger noncausality relations is as follows.

**Theorem 1.** Let  $\epsilon_t^{(2)}$  follow a stationary vector autoregressive process, as in (8), partitioned, as in (9), that is identifiable (Assumptions 1–3). Then,  $y_1$  does not second-order Granger-cause  $y_2$  given  $y_3$  (denoted by  $y_1 \xrightarrow{so} y_2 | y_3$ ) if and only if

$$\Pi_{21}(z) \equiv 0 \quad \forall z \in \mathbb{C}. \quad (12)$$

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 (12), 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 (4). 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 (6), partitioned, as in (7), which is identifiable and invertible (Assumptions 1–3). Then  $y_1$  does not second-order Granger-cause  $y_2$  given  $y_3$  (denoted by  $y_1 \xrightarrow{so} y_2 | y_3$ ), if and only if

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

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

Condition (13) 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  $i = 1, \dots, \max(r, s) + (N_1 + N_3)s$ . Therefore, the joint number of restrictions for a second-order noncausality hypothesis is  $N_1 N_2 (\max(r, s) + s(N_1 + N_3))$ . Such restrictions are ready to be tested.

The innovation of condition (13) 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. Such restrictions are particularly useful 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 articles 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 (13) generalizes results from other studies. Comte and Lieberman (2000) derive similar restriction for the BEKK-GARCH model, with the difference that vector  $y_t$  is partitioned only into two subvectors  $y_{1t}^*$  and  $y_{2t}^*$ . Woźniak (2015) 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,  $y_{1t}^* \xrightarrow{so} y_{2t}^*$  (which can be written as  $y_{1t} \xrightarrow{so} (y_{2t}, y_{3t})$ ) is at most a sufficient condition for  $y_{1t} \xrightarrow{so} y_{2t} | y_{3t}$  or that  $y_{1t} \xrightarrow{so} y_{3t} | y_{2t}$  (see Example 4 of the working paper version of this work Woźniak, 2012b, for the illustration in a trivariate case). The sufficient condition is however a much stronger restriction, especially in a multivariate settings. Moreover, the results of Woźniak (2015) are nested in condition (13) by setting  $N_3 = 0$ .

The restrictions for second-order noncausality relate to the definitions based on the linear theory proposed by Florens and Mouchart (1985) and successfully applied to the analysis of multivariate time series models (see, e.g., Boudjellaba et al., 1992, 1994; Comte and Lieberman, 2000; Woźniak, 2015). The main cost arising from this choice is the assumed linearity of the model. Such features of the model as the dependence of the parameters on time, and especially time-varying conditional correlations, are not allowed in this approach. Contrarily, Granger-causal analysis relating to the concept of conditional independence or to definitions stated in terms of the expected value operator could potentially allow for nonlinearities in the analyzed models.



To conclude this section, two examples illustrating the derivation of the parameter restrictions for several processes that are often used in empirical works are given.

**Example 1.** 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:

$$\begin{aligned} & \begin{bmatrix} 1 - (A_{11} + B_{11})L & -(A_{12} + B_{12})L & -(A_{13} + B_{13})L \\ -(A_{21} + B_{21})L & 1 - (A_{22} + B_{22})L & -(A_{23} + B_{23})L \\ -(A_{31} + B_{31})L & -(A_{32} + B_{32})L & 1 - (A_{33} + B_{33})L \end{bmatrix} \begin{bmatrix} \epsilon_{1t}^2 \\ \epsilon_{2t}^2 \\ \epsilon_{3t}^2 \end{bmatrix} \\ &= \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} + \begin{bmatrix} 1 - B_{11}L & -B_{12}L & -B_{13}L \\ -B_{21}L & 1 - B_{22}L & -B_{23}L \\ -B_{31}L & -B_{32}L & 1 - B_{33}L \end{bmatrix} \begin{bmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{bmatrix}. \end{aligned} \quad (14)$$

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

$$\mathbf{R}_1(\theta) = A_{11}(B_{23}B_{31} - B_{21}B_{33}) + A_{31}(B_{13}B_{21} - B_{11}B_{23}) = 0, \quad (15a)$$

$$\mathbf{R}_2(\theta) = A_{11}B_{21} + A_{31}B_{23} = 0, \quad (15b)$$

$$\mathbf{R}_3(\theta) = A_{21} = 0. \quad (15c)$$

**Example 2.** 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

$$\mathbf{R}(\theta) = A_{i,21} = 0 \quad \text{for } i = 1, \dots, r. \quad (16)$$

#### 4. Bayesian testing of noncausality

In the following section, the problem of testing restrictions imposed on the original parameters of the ECCC-GARCH model is considered. Moreover, a Bayesian Lindley-type testing procedure is proposed as a solution for some of the drawbacks of classical tests.

**The Wald Test.** Consider the classical Wald test of Boudjellaba et al. (1992) for the parameter restrictions for Granger noncausality in the GARCH process. The Wald test has the desirable feature that it requires the estimation of only the most general model. Notice, that the estimation of a model restricted according to the noncausality restrictions (e.g., as these from Eq. (15)) is cumbersome and requires further studies. Thus, by estimating just one unrestricted model, one can do both: perform the testing procedure, and analyze the parameters responsible for possible volatility transmissions. The Wald statistic is then given by

$$\mathbf{W}(\hat{\theta}_m) = T\mathbf{R}(\hat{\theta}_m)' [\mathbf{T}(\hat{\theta}_m)' \mathbf{V}(\hat{\theta}_m) \mathbf{T}(\hat{\theta}_m)]^{-1} \mathbf{R}(\hat{\theta}_m), \quad (17)$$

where  $\theta_m$  is a subvector of  $\theta$ , 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  $\theta_m$ :

$$\mathbf{T}(\hat{\theta}_m) = \left. \frac{\partial \mathbf{R}(\theta_m)}{\partial \theta_m} \right|_{\theta_m = \hat{\theta}_m}. \quad (18)$$

Under the null hypothesis of second-order noncausality  $\mathbf{W}(\hat{\theta}_m)$  has an asymptotic  $\chi^2(l_m)$  distribution. However, in Eq. (17)  $\mathbf{T}(\theta_m)$  must be of full rank. Otherwise, the asymptotic covariance matrix is singular and the asymptotic distribution is no longer  $\chi^2(l_m)$ . In the context of Granger causality for VARMA models, Boudjellaba et al. (1992), testing the nonlinear restrictions, show that there are cases when  $\mathbf{T}(\theta_m)$  is not of full rank under the null hypothesis. The same problem applies to the Wald test in

the context of the current article. This is shown for the bivariate case in Woźniak (2015). The example below considers the trivariate case.

**Example 1 (continued).** The Jacobian matrix for the restrictions in Eq. (15) is

$$\mathbf{T}(\theta)' = \begin{bmatrix} B_{23}B_{31} - B_{21}B_{33} & B_{21} & 0 \\ 0 & 0 & 1 \\ B_{13}B_{21} - B_{11}B_{23} & B_{23} & 0 \\ -A_{31}B_{23} & 0 & 0 \\ A_{31}B_{21} & 0 & 0 \\ A_{31}B_{21} - A_{11}B_{33} & A_{11} & 0 \\ A_{11}B_{31} - A_{31}B_{11} & A_{31} & 0 \\ A_{11}B_{23} & 0 & 0 \\ -A_{11}B_{21} & 0 & 0 \end{bmatrix}.$$

If a sufficient condition for (15) to hold

$$A_{11} = A_{21} = A_{31} = B_{21} = B_{23} = 0 \quad (19)$$

holds as well, then the rank of matrix  $\mathbf{T}(\theta)$  is one, and therefore, under the null hypothesis the Jacobian matrix is not of full rank. Condition (19) is reducible and the null hypothesis is also implied by either  $A_{11} = A_{21} = A_{31} = 0$  or  $A_{21} = B_{21} = B_{23} = 0$ . Therefore, the asymptotic properties of the Wald test depend on auxiliary parameters, i.e.,  $B_{21}$  and  $B_{23}$  for the former, and  $A_{11}$  and  $A_{31}$  for the latter case.

The asymptotic distribution of the test statistics in cases such as those described in Example 1 is unknown. Note that this feature is not a direct consequence of the form of the restrictions  $\mathbf{R}(\theta)$  (by construction, they are linearly independent), but rather it comes from approximating the asymptotic covariance matrix using the delta method. Several works coping with this problem have appeared (Dufour, 1989; Boudjellaba et al., 1992; Lütkepohl and Burda, 1997; Dufour et al., 2006), in the context of testing Granger noncausality for conditional mean processes.

In fact, the Wald test was applied to test the restrictions for the second-order noncausality in the BEKK-GARCH models by Comte and Lieberman (2000) and Hafner and Herwartz (2008). The Wald statistics, proposed by Comte and Lieberman and Hafner and Herwartz, is  $\chi^2$ -distributed, given the asymptotic normality of the QMLE of the parameters (see Comte and Lieberman, 2003). The asymptotic distribution of the test statistic, however, could only be obtained due to the simplifying approach undertaken in which sufficient linear conditions are tested. Among the classical solutions proposed for the problem of testing the Granger noncausality in conditional means, only the modified Wald test of Lütkepohl and Burda (1997) seems applicable for testing second-order noncausality in the GARCH models. Nevertheless, further research of this topic is required.

The asymptotic normality of the QMLE of the parameters for the VARMA-ECCC-GARCH models was derived by Ling and McAleer (2003), under the assumption of the existence of sixth-order unconditional moments of  $y_t$ . A similar result was obtained by Hafner and Preminger (2009) for the BEKK-GARCH models, also under the assumption of the existence of sixth-order unconditional 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.

**Bayesian Testing.** In this study, an alternative approach to testing is undertaken following the Lindley-type test used in Osiewalski and Pipień (2002) and Marzec and Osiewalski (2008). Firstly, we propose the method of testing the original restrictions on the parameters for the second-order noncausality, and secondly, the Bayesian procedure presented in the subsequent part is shown to overcome 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.

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 considered set of hypotheses is represented by

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

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

The formulation of the hypotheses is general and encompasses second-order noncausality and other hypotheses resulting in either linear or nonlinear restrictions.

Given the posterior distribution of the parameters in the Bayesian approach, the posterior distribution of the function  $\mathbf{R}(\theta)$  is available,  $p(\mathbf{R}(\theta) | \mathbf{y})$ . 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  $p(\mathbf{R}(\theta) | \mathbf{y})$ . Define a scalar function  $\kappa : \mathbb{R}^l \rightarrow \mathbb{R}^+$  by

$$\kappa(\mathbf{R}) = [\mathbf{R} - E[\mathbf{R}(\theta) | \mathbf{y}]]' V[\mathbf{R}(\theta) | \mathbf{y}]^{-1} [\mathbf{R} - E[\mathbf{R}(\theta) | \mathbf{y}]], \quad (20)$$

where  $\mathbf{R}$  is the argument of the function. In order to distinguish the argument of the function  $\mathbf{R} = \mathbf{R}(\theta)$ , the simplified notation is used, neglecting the dependence on the vector of parameters. In the 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.

The function  $\kappa$  is a positive semidefinite quadratic form of a real-valued vector. It gives a measure of the deviation of the value of the vector of restrictions from its posterior mean,  $\mathbf{R} - E[\mathbf{R}(\theta) | \mathbf{y}]$ , 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. Alternatively, one could avoid using the delta method by drawing a sample directly from the asymptotic distribution of the parameters and transforming it to a sample drawn from the asymptotic distribution of the restrictions that could be used to compute the covariance matrix in (17). Another potential solution to the problems with Wald testing would be the employment of the resampling methods (e.g., bootstrap) to repeat the Bayesian procedure described in this section. Nevertheless, the application of these solutions would require further studies. Notice, however, that the function  $\kappa$  is not a test statistic, but a scalar function that summarizes multiple restrictions on the parameters of the model.

Moreover, if  $\mathbf{R}$  follows a normal density function, then  $\kappa(\mathbf{R})$  has a  $\chi^2(l)$  distribution with  $l$  degrees of freedom (see, e.g., Proposition B.3 (2) of Lütkepohl, 2005, p. 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  $\chi^2$ -distributed  $\kappa$  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 VAR-GARCH models with Student's  $t$  likelihood function are analyzed, the posterior distribution of the parameters of neither the VAR nor GARCH parts are in the form of some known distributions (see Bauwens and 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. Luckily, using the Monte Carlo Markov Chain methods, the posterior distributions of the parameters of the model,  $\theta$ , can be simulated, and those of the restrictions imposed on them,  $\mathbf{R}$ , as well as of the function  $\kappa(\mathbf{R})$ , easily computed. The posterior distribution of the function  $\kappa$  is used in order to evaluate the hypothesis of noncausality.

Let  $\kappa(\mathbf{0})$  be the value of function  $\kappa$ , 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  $\kappa(\mathbf{R}) > \kappa(\mathbf{0})$ ,

$$p_0 = \Pr(\kappa(\mathbf{R}) > \kappa(\mathbf{0}) | \mathbf{y}) = \int_{\kappa(\mathbf{0})}^{\infty} p(\kappa(\mathbf{R}) | \mathbf{y}) d\kappa(\mathbf{R}). \quad (21)$$

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:

$$\hat{p}_0 = \frac{\#\{\kappa(\mathbf{R}^{(i)}) > \kappa(\mathbf{0})\}}{S_2}. \quad (22)$$

The probability,  $\hat{p}_0$ , should be compared to a tail probability,  $\pi_0$ , and  $1 - \pi_0$  is a measure of the reliability of a result. The values for probability  $\pi_0$  that one might consider are, e.g., 0.05 or 0.1. In Fig. 3, an illustration of the testing procedure for two of the hypotheses investigated in Section 5, when one gives rejection of the null hypothesis and the other does not, is presented.

Osiewalski and Pipień (2002) and Marzec and Osiewalski (2008) use the quadratic form as in (20) in order to assess restrictions on parameters. Osiewalski and Pipień (2002) proposed the *Lindley-type Bayesian counterpart of the usual F test* to compare a bivariate BEKK-GARCH specification with the vech-GARCH nesting BEKK-GARCH. They tested zero restrictions imposed on nonlinear functions of parameters. In order to assess the null hypothesis, they use the fact that the posterior distribution of the parameters is approximately asymptotically normal, and therefore, the quadratic form is approximately  $\chi^2$ -distributed. Therefore, the null hypothesis is rejected if the probability  $\Pr(\chi^2 \leq \kappa(\mathbf{0}))$  is greater than  $1 - \alpha$ , a confidence level of the classical test, where  $\chi^2$  is a chi-square random variable with appropriate degrees of freedom parameter. Marzec and Osiewalski (2008) test zero restrictions imposed on the parameters of a stochastic frontier model in order to compare different specifications of the cost function and the distributions for inefficiency effects nested in the estimated general formulation of the model. The null hypothesis is assessed by checking the probability content of the shortest Highest Posterior Density interval that includes value  $\kappa(\mathbf{0})$ . This procedure practically corresponds to the one used in the current study.

**Discussion.** The proposed Bayesian procedure allows testing of the noncausality restrictions resulting directly from the determinant condition (12). 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 and Lieberman (2000) or Hafner and Herwartz (2008). Second, the procedure requires the estimation of only one unrestricted model for the purpose of testing, which given the time required to estimate multivariate VAR-GARCH models is a significant gain as compared to the procedure proposed by Woźniak (2015).

Moreover, Lindley type inference based on Highest Posterior Density regions provides a summary of the Bayesian evidence for or against hypotheses. It strongly relates to the likelihood principle stating that the whole information contained in sample data about the parameters of the model is represented by the likelihood function. Since there is no need to refer to asymptotic results, the Bayesian test relaxes the assumptions of the existence of higher-order moments. Only the existence of fourth-order unconditional moments is assumed (as in Assumption 3), in comparison to the required existence of sixth-order moments in a classical derivation of the asymptotic distribution of the QMLE (see Ling and McAleer, 2003). Finally, this testing procedure could be employed for the restrictions of Comte and Lieberman (2000) for testing the noncausality in variance in the BEKK-GARCH models (that require the existence of the sixth-order moments for the classical testing; see Hafner and Preminger, 2009).

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.

## 5. Granger causal analysis of exchange rates

**Data.** In order to illustrate the use of the methods presented in previous sections three time series of daily exchange rates are chosen. A system of three variables is the smallest to use the innovation of the derived restrictions. The series, all denominated in Euro, are the Swiss franc (CHF/EUR), the British pound (GBP/EUR) and the United States dollar (USD/EUR). Logarithmic rates of return expressed in percentage points are used,  $y_{it} = 100(\ln x_{it} - \ln x_{it-1})$  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  $T = 777$  prices. 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.

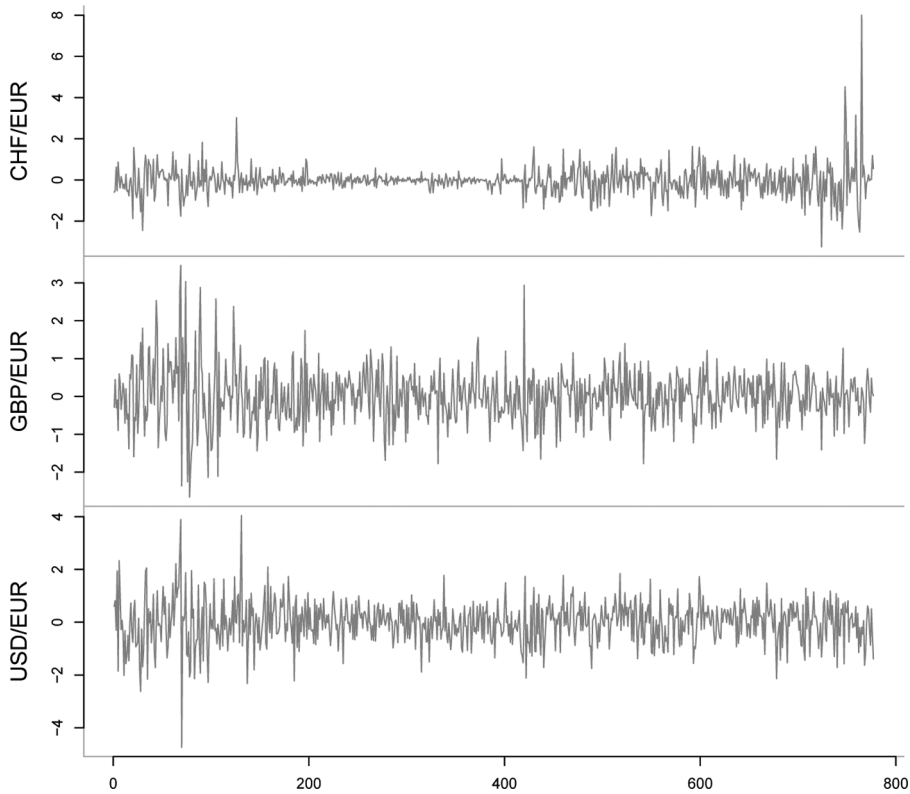
The motivation behind this choice of variables and the period of analysis is its usefulness for the institutions for which the forecast of the exchange rates is a crucial element of financial planning. For instance, suppose that the government of a country participating in the Eurozone is indebted in 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 (2012a).

Figure 1 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.

**Estimation of the Model.** The Bayesian estimation of the VAR-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.

The parameters of the VAR-GARCH model follow the subsequent prior specification. For the parameters of the vector autoregressive process of order one, VAR(1), and of the GARCH(1,1) model, the assumed prior distribution is proportional to a constant and constrained to a parameter space bounded according to Assumptions 1–3. Each of the parameters of the correlation matrix,  $C$ , collected in a  $N(N-1)/2 \times 1$  vector  $\rho = \text{vecl}(C)$ , follows a uniform distribution on the interval  $[-1, 1]$ , where a *vecl* operator stacks lower-diagonal elements of a matrix in a vector. Finally, the degrees of freedom parameter follows the prior distribution proposed by Deschamps (2006). Such a prior specification, with diffuse distributions for all the parameters but the degrees of freedom parameter  $\nu$ , guarantees the existence of the posterior distribution, where its existence is understood as the integrability of the product of the likelihood function and the prior distribution (see Bauwens and Lubrano, 1998). It does not discriminate any of the values of the parameters from within the parameter space. The prior distribution for the parameter  $\nu$  is a proper density function, and it gives as much as a 32% chance that its value is greater than 30. For such values of this parameter, the likelihood function given by Eq. (5b) is a close approximation of the normal likelihood function.



**Figure 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>).

**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	

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 and White (2004). LJB test and LJB  $p$ -values describe the test of normality by Lomnicki (1961) and Jarque and Bera (1980).

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

$$p(\theta) = p(\alpha'_0, \text{vec}(\alpha_1)) p(\omega', \text{vec}(A)', \text{vec}(B)') p(\nu) \prod_{i=1}^{N(N-1)/2} p(\rho_i), \quad (23)$$

where each prior distribution is specified by

$$\begin{aligned} p(\alpha'_0, \text{vec}(\alpha_1)') &\propto \mathcal{I}(\theta \in \Theta) \\ p(\omega', \text{vec}(A)', \text{vec}(B)') &\propto \mathcal{I}(\theta \in \Theta) \\ v &\sim .04 \exp[-.04(v-2)] \mathcal{I}(v \geq 2) \\ \rho_i &\propto \mathcal{I}(\mathbf{C} \text{ positive definite}), \end{aligned}$$

for  $i = 1, \dots, N(N-1)/2$ , where  $\mathcal{I}(\cdot)$  is an indicator 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 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 posterior distributions. Therefore, numerical methods need to be employed in order to simulate the posterior distribution. The random walk Metropolis–Hastings algorithm adapted for the GARCH models by Vrontos et al. (2003) and used in Osiewalski and Pipień (2002, 2004) is used also in this study. At each step of the algorithm, a candidate draw,  $\theta^*$ , 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,  $\Omega$ , 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 (see Vrontos et al., 2003). A new candidate draw,  $\theta^*$ , is accepted with the probability:

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

Every 100th state of the Markov Chain is kept 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 length of the Markov chain Monte Carlo (MCMC), the chain of desirable properties according to several criteria (see Geweke, 1989, 1992; Plummer et al., 2006) is obtained.

**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. 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. However, the 90% 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 the US dollar on the current value of the Swiss franc. This finding has its consequences in testing the Granger causality in mean hypothesis.

All of 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. The posterior probability mass of several of the parameters, however, is distant from zero. All the diagonal parameters of matrices  $A$  and  $B$ , except 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 the 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 (2012a) for the same time span. 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



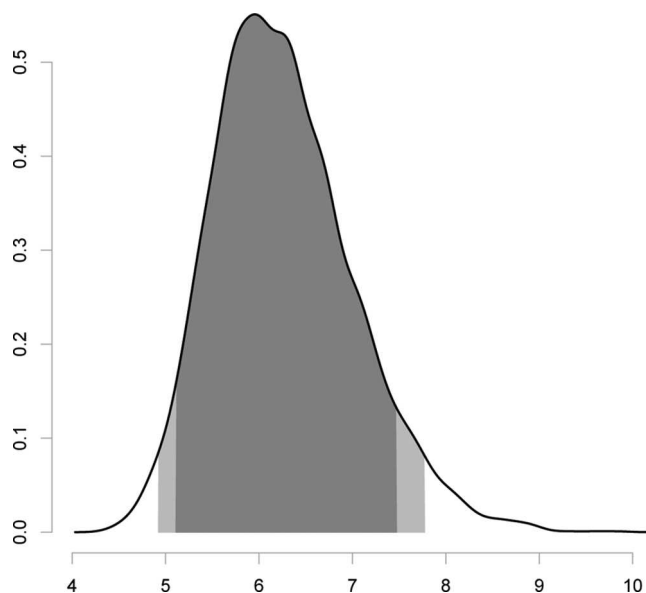
**Table 2.** Summary of the estimation of the VAR(1)-ECCC-GARCH(1,1) model.

VAR(1)							
	$\alpha_0$		$\alpha_1$				
CHF/EUR	−0.022 (0.011)	−0.068 (0.037)	0.003 (0.019)	0.041 (0.018)			
GBP/EUR	0.014 (0.021)	−0.016 (0.030)	0.077 (0.040)	−0.019 (0.033)			
USD/EUR	0.027 (0.025)	−0.027 (0.041)	0.050 (0.045)	0.006 (0.041)			
GARCH(1,1)							
	$\omega$		$A$		$B$		
CHF/EUR	0.001 (0.001)	0.117 (0.029)	0.002 (0.002)	0.002 (0.002)	0.873 (0.030)	0.001 (0.001)	0.001 (0.001)
GBP/EUR	0.011 (0.009)	0.002 (0.002)	0.062 (0.024)	0.017 (0.011)	0.002 (0.003)	0.808 (0.158)	0.063 (0.098)
USD/EUR	0.086 (0.059)	0.018 (0.018)	0.117 (0.062)	0.051 (0.034)	0.031 (0.028)	0.787 (0.215)	0.164 (0.147)
Degrees of freedom and correlations							
	$\nu$	$\rho_{12}$	$\rho_{13}$	$\rho_{23}$			
	6.267 (0.743)	0.145 (0.038)	0.356 (0.035)	0.400 (0.034)			

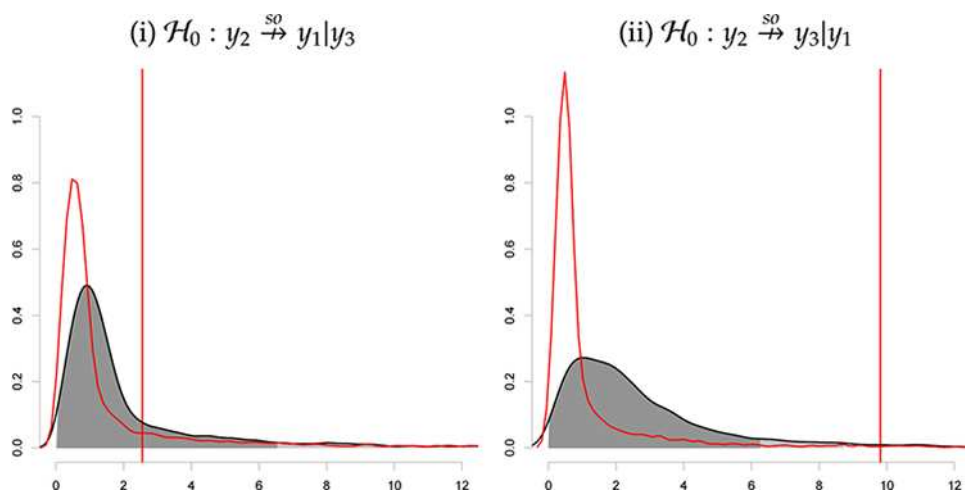
The table summarises the estimation of the VAR(1)-ECCC-GARCH(1,1) model described by the Eqs. (2), (3), (4), and the likelihood function (5b). The prior distributions are specified in Eq. (23). The posterior means and the posterior standard deviations (in brackets) are reported.

to the prior distribution specification. Woźniak (2012a) estimates two models with a truncated-normally distributed priors with two different variance parameters: 100 and 0.1.

Finally, Fig. 2 proves that the parameter of the degrees of freedom,  $\nu$ , of the  $t$ -distributed residuals cannot be considered strictly greater than 6. The posterior probability of such an event is just over 50%. Therefore, a hypothesis that moments of order six and higher of the conditional distribution do not exist has a non-negligible probability mass. Although  $\nu$  is a parameter of a conditional distribution, the



**Figure 2.** Marginal posterior distribution of the degrees of freedom parameter  $\nu$ . The marginal posterior distribution of parameter  $\nu$  is plotted with the black line. The shaded areas denote the 95% and 90% highest posterior density regions of the distribution.



**Figure 3.** Sample results of testing: second-order Granger causality hypotheses. The simulated posterior distributions of function  $\kappa$  for two hypotheses of second order noncausality are plotted with the black lines. The shaded areas denote the 95% and 90% highest posterior density regions of the distributions. The simulated prior distributions of  $\kappa$  implied by the prior distribution for  $\theta$  are plotted with red lines. The horizontal red bars denote the value of  $\kappa(\mathbf{0})$  representing the null hypotheses given above the graphs. For more detailed results, refer to Table 4. Description of the variables:  $y_1$  = CHF/EUR,  $y_2$  = GBP/EUR,  $y_3$  = USD/EUR.

existence of unconditional moments of order 6 and higher of the error term is questionable. In effect, classical testing of the VAR-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 and McAleer, 2003).

**Granger-Causality Testing Results.** Table 3 presents the results of testing the Granger noncausality in mean hypotheses represented by the standard zero restrictions on appropriate parameters of the

**Table 3.** Results of testing: Granger causality hypothesis.

$\mathcal{H}_0 :$	$\kappa(\mathbf{0})$	$\hat{p}_0$
$y_1 \xrightarrow{G} y_2   y_3$	0.294	0.586
$y_1 \xrightarrow{G} y_3   y_2$	0.427	0.522
$y_2 \xrightarrow{G} y_1   y_3$	0.022	0.883
$y_2 \xrightarrow{G} y_3   y_1$	1.226	0.270
$y_3 \xrightarrow{G} y_1   y_2$	5.013	0.023
$y_3 \xrightarrow{G} y_2   y_1$	0.336	0.561
$(y_1, y_2) \xrightarrow{G} y_3$	1.580	0.455
$(y_1, y_3) \xrightarrow{G} y_2$	0.884	0.642
$(y_2, y_3) \xrightarrow{G} y_1$	5.520	0.063
$y_1 \xrightarrow{G} (y_2, y_3)$	0.530	0.765
$y_2 \xrightarrow{G} (y_1, y_3)$	1.252	0.543
$y_3 \xrightarrow{G} (y_1, y_2)$	5.776	0.059
$y_1 \xrightarrow{G} y_2   y_3 \text{ \& } y_2 \xrightarrow{G} y_1   y_3$	0.315	0.858
$y_1 \xrightarrow{G} y_3   y_2 \text{ \& } y_3 \xrightarrow{G} y_1   y_2$	5.249	0.072
$y_2 \xrightarrow{G} y_3   y_1 \text{ \& } y_3 \xrightarrow{G} y_2   y_1$	1.402	0.490

The table presents the considered null hypotheses,  $\mathcal{H}_0$ , of Granger noncausality. The values of function  $\kappa$  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 (21). Description of the variables:  $y_1$  = CHF/EUR,  $y_2$  = GBP/EUR,  $y_3$  = USD/EUR.

**Table 4.** Results of testing: Second-order Granger causality hypothesis.

$\mathcal{H}_0$ :	$\kappa(\mathbf{0})$	$\hat{p}_0$
$y_1 \overset{SO}{\nrightarrow} y_2 \mid y_3$	1.521	0.279
$y_1 \overset{SO}{\nrightarrow} y_3 \mid y_2$	3.120	0.221
$y_2 \overset{SO}{\nrightarrow} y_1 \mid y_3$	2.550 <sup>(i)</sup>	0.247 <sup>(i)</sup>
$y_2 \overset{SO}{\nrightarrow} y_3 \mid y_1$	9.811 <sup>(ii)</sup>	0.039 <sup>(ii)</sup>
$y_3 \overset{SO}{\nrightarrow} y_1 \mid y_2$	2.379	0.231
$y_3 \overset{SO}{\nrightarrow} y_2 \mid y_1$	2.354	0.193
$(y_1, y_2) \overset{SO}{\nrightarrow} y_3$	11.539	0.113
$(y_1, y_3) \overset{SO}{\nrightarrow} y_2$	2.733	0.372
$(y_2, y_3) \overset{SO}{\nrightarrow} y_1$	3.926	0.386
$y_1 \overset{SO}{\nrightarrow} (y_2, y_3)$	3.491	0.246
$y_2 \overset{SO}{\nrightarrow} (y_1, y_3)$	10.714	0.061
$y_3 \overset{SO}{\nrightarrow} (y_1, y_2)$	4.084	0.227
$y_1 \overset{SO}{\nrightarrow} y_2 \mid y_3 \ \& \ y_2 \overset{SO}{\nrightarrow} y_1 \mid y_3$	3.481	0.386
$y_1 \overset{SO}{\nrightarrow} y_3 \mid y_2 \ \& \ y_3 \overset{SO}{\nrightarrow} y_1 \mid y_2$	4.741	0.324
$y_2 \overset{SO}{\nrightarrow} y_3 \mid y_1 \ \& \ y_3 \overset{SO}{\nrightarrow} y_2 \mid y_1$	11.633	0.099

The table presents the considered null hypotheses of second-order Granger causality, as in Definition 1. For a description of the notation, see the note to Table 3.

<sup>(i)</sup> and <sup>(ii)</sup> are the visual representation of the results can be found in Fig. 3 (i) and (ii), respectively. Description of the variables:  $y_1$  = CHF/EUR,  $y_2$  = GBP/EUR,  $y_3$  = USD/EUR.

autoregressive matrices  $\alpha$ . The values of  $\kappa(\mathbf{0})$  and of the estimate of the probability  $p_0$  are reported. Only a few of the hypotheses of noncausality in mean are rejected at the levels for  $\pi_0$  equal to 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 franc, a result established at the tail probability being less than 0.05. These two exchange rates impact on each other as well. Further, the US dollar has a significant effect on the Swiss franc and the British pound taken jointly. Finally, the franc is significantly affected by both the pound and the dollar, taken jointly. All the three last results are established at the level of  $\pi_0 = 0.1$ .

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

$\mathcal{H}_0$ :	$\kappa(\mathbf{0})$	$\hat{p}_0$
$y_1 \overset{V}{\nrightarrow} y_2 \mid y_3$	1.755	0.502
$y_1 \overset{V}{\nrightarrow} y_3 \mid y_2$	3.791	0.297
$y_2 \overset{V}{\nrightarrow} y_1 \mid y_3$	2.591	0.430
$y_2 \overset{V}{\nrightarrow} y_3 \mid y_1$	11.177	0.040
$y_3 \overset{V}{\nrightarrow} y_1 \mid y_2$	8.007	0.095
$y_3 \overset{V}{\nrightarrow} y_2 \mid y_1$	2.795	0.324
$(y_1, y_2) \overset{V}{\nrightarrow} y_3$	13.412	0.128
$(y_1, y_3) \overset{V}{\nrightarrow} y_2$	3.728	0.545
$(y_2, y_3) \overset{V}{\nrightarrow} y_1$	10.350	0.205
$y_1 \overset{V}{\nrightarrow} (y_2, y_3)$	4.229	0.440
$y_2 \overset{V}{\nrightarrow} (y_1, y_3)$	12.098	0.083
$y_3 \overset{V}{\nrightarrow} (y_1, y_2)$	10.599	0.118
$y_1 \overset{V}{\nrightarrow} y_2 \mid y_3 \ \& \ y_2 \overset{V}{\nrightarrow} y_1 \mid y_3$	3.826	0.629
$y_1 \overset{V}{\nrightarrow} y_3 \mid y_2 \ \& \ y_3 \overset{V}{\nrightarrow} y_1 \mid y_2$	10.900	0.184
$y_2 \overset{V}{\nrightarrow} y_3 \mid y_1 \ \& \ y_3 \overset{V}{\nrightarrow} y_2 \mid y_1$	13.310	0.119

The table presents the considered null hypotheses of Granger causality in variance, 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.

**Table 6.** Summary of the posterior distribution simulation.

				Autocorrelations at			
		Mean	SD	lag 1	lag 50	RNE	Geweke's z
Vector Autoregression							
$\alpha_0$	$\alpha_{0,1}$	−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
$\alpha_1$	$\alpha_{1,11}$	−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
	$\alpha_{1,22}$	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)							
$\omega$	$\omega_1$	0.001	0.001	0.677	0.032	0.108	−0.773
	$\omega_2$	0.011	0.009	0.600	0.023	0.082	−1.154
	$\omega_3$	0.086	0.059	0.392	0.058	0.145	1.818
A	$A_{11}$	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
	$A_{22}$	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	$B_{11}$	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
	$B_{22}$	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							
C	$\nu$	6.267	0.743	0.415	0.017	0.447	0.123
	$\rho_{12}$	0.145	0.038	0.335	0.036	0.481	1.444
	$\rho_{13}$	0.356	0.035	0.359	−0.021	0.481	2.820
	$\rho_{23}$	0.400	0.034	0.328	−0.006	0.501	0.048

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).

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 the second-order causality than for the causality in mean. The rejected hypotheses of second-order noncausality relate to the British pound and the US dollar. Information about the history of the volatility of GBP/EUR has a significant effect on the current conditional variance of variable USD/EUR at  $\pi_0 = 0.05$ . The visual representation of this result is given on the RHS of Fig. 3. The same conclusion is found in Woźniak (2012a), which is particularly interesting, as Woźniak uses Bayes factors and Posterior probabilities in order to assess the hypotheses. This conclusion is, therefore, robust to the choice of the testing procedure. GBP/EUR has also a significant effect on CHF/EUR and USD/EUR, taken jointly for  $\pi_0 = 0.1$ .

The following interpretation of the testing results of second-order noncausality hypothesis is proposed. The Swiss franc 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 *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 levels and their volatility, affects the decisions of traders in North America. Such a pattern can be captured by the dataset and the model considered in this study.

One more hypothesis is rejected at the value of the tail probability less than 0.1: the pound is found to second-order cause the dollar, *and* the dollar second-order causes the pound. The result of the joint test of these two hypotheses is mainly driven by parameter  $B_{32}$ .

Finally, the results of testing hypotheses of noncausality in variance are reported in Table 5. Although the restrictions for this case include both, the restrictions for noncausality in mean and second-order noncausality, the results of the statistical inference are not just a simple intersection of the results of testing these hypotheses separately, as one might deduce from Eq. (11). The parameters of the VAR process are not independent of the parameters of the GARCH process: the posterior covariance matrix is not block-diagonal. In effect, only one of the hypotheses is rejected with the value of the tail probability less than 0.05: the hypothesis of noncausality in variance from the British pound to the US dollar. The other three hypotheses are rejected for  $\pi_0$  equal to 0.1. The following relations are found: the dollar causes the franc in variance; and the pound causes in variance the franc and the dollar taken jointly. For completeness, the summary of the properties of the final sample of draws from the posterior distribution is presented in Table 6.

## 6. Conclusions

This study 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 ECCC-GARCH model. The novelty of these conditions is that, contrary to the developments of Comte and Lieberman (2000) and Woźniak (2015), 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, the Bayesian procedure based on the posterior distribution of a function summarizing all the restrictions is proposed. 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. Note that the procedure proposed by Woźniak (2015), based on Bayes factors, 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 (2015), the existence of fourth-order moments is required in the proposed analysis, whereas the asymptotic derivations of Ling and 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 and Lieberman (2000) and Woźniak (2015), 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 nonempty, 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  $h = 1, 2, 3, \dots$ . 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 and 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: Proofs

### Proof of Theorem 1

The forecast error of  $y_{2,t+1}$  from the stationary conditional mean process is  $\epsilon_{2,t+1} = y_{2,t+1} - P(y_{2,t+1} | I(t))$ , and its square is denoted by  $\epsilon_{2,t+1}^{(2)}$ . Under Assumption 1, the VARMA representation of the ECCC-GARCH model for  $\epsilon_t^{(2)}$ , as in Eq. (6), exists and is invertible. Moreover, under Assumption 2, it is identifiable. Therefore, its uniquely defined VAR representation for  $\epsilon_t^{(2)}$ , as in Eq. (8), exists and is stationary. Having imposed the partitioning of vectors accordingly as in Eq. (1), the VAR representation of  $\epsilon_t^{(2)}$  becomes (9). A model in such a form is ready for second-order Granger causality analysis as stated in Definition 1. Theorem 1 is then proven straightaway by applying Proposition 1 of Boudjellaba et al. (1992). The theorem states under what conditions the orthogonal projection of  $\epsilon_{2,t+1}^{(2)}$  on  $I_\epsilon^2(t)$  and on  $I_{\epsilon,-1}^2(t)$  are equivalent. In other words, it states under what conditions the best linear predictor of  $\epsilon_{2,t+1}^{(2)}$  in terms of the objects included in  $I_\epsilon^2(t)$ , i.e.,  $\epsilon_{1,\tau}^{(2)}, \epsilon_{2,\tau}^{(2)}, \epsilon_{3,\tau}^{(2)}$  for  $\tau \leq t$  and constant terms, is equal to its best linear predictor in terms of the elements  $I_{\epsilon,-1}^2(t)$ , i.e.,  $\epsilon_{2,\tau}^{(2)}, \epsilon_{3,\tau}^{(2)}$  for  $\tau \leq t$  and constant terms.

### Proof of Theorem 2

In order to prove the simplified conditions for second-order Granger noncausality, (13), apply to Eq. (12) from Theorem 1 the matrix transformations of Theorem 3 and then of Theorem 4 of Boudjellaba et al. (1994).

## Acknowledgments

The author acknowledges the editor and the two anonymous referees for their comments. He thanks Professors Helmut Lütkepohl, Jacek Osiewalski, William Griffiths, Lennart Hoogerheide, Timo Teräsvirta, Michael Smith, Vance Martin, Massimiliano Marcellino, Giampiero Gallo, Mateusz Pipień, and Russell Cooper for useful comments and remarks on the study. The author is grateful to the participants of the seminars at the Cracow University of Economics, Monash University, the University of Queensland, the University of Melbourne, Università di Pisa, VU University Amsterdam, the Centre for Financial Econometrics at the Singapore Management University, and Freie Universität Berlin. The paper was presented at the SMYE 2011 Conference at the University of Groningen, the ISBA 2012 World Meeting, the Bayes on the Beach 2012 Conference, as well as at the Fifth Italian Congress of Econometrics and Empirical Economics 2013, Seventh International Conference on Computational and Financial Econometrics 2013, the Econometric Society Australasian Meeting 2013, and the First Conference on Recent Developments in Financial Econometrics and Applications 2014. Also, the author thanks Michał Markun and Norbert Metiu and his colleagues from the EUI Time Series Econometrics Working Group, Pierre Guérin, Matthieu Droumaguet, and Stelios Bekiros, for multiple discussions and suggestions.

## ORCID

Tomasz Woźniak  <http://orcid.org/0000-0003-2212-2378>

## References

- Bauwens, L., Lubrano, M. (1998). Bayesian inference on GARCH models using the Gibbs sampler. *Econometrics Journal* 1(1):C23–C46.
- Ben Omrane, W., Hafner, C. M. (2009). Information spillover, volatility and the currency markets. *International Econometric Review* 1:47–59.
- 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(4):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(3):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(420):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(1):33–48.
- Chib, S., Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician* 49(4):327.
- Comte, F., Lieberman, O. (2000). Second-order noncausality in multivariate GARCH processes. *Journal of Time Series Analysis* 21(5):535–557.
- Comte, F., Lieberman, O. (2003). Asymptotic theory for multivariate GARCH processes. *Journal of Multivariate Analysis* 84(1):61–84.
- Conrad, C., Karanasos, M. (2010). Negative volatility spillovers in the unrestricted Eccc-Garch model. *Econometric Theory* 26(3):1–25.
- Deschamps, P. J. (2006). A flexible prior distribution for Markov switching autoregressions with student-t errors. *Journal of Econometrics* 133(1):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(534):158–171.
- Dufour, J.-M. (1989). Nonlinear hypotheses, inequality restrictions, and Non-nested hypotheses: Exact simultaneous tests in linear regressions. *Econometrica* 54(2):335–355.
- Dufour, J.-M., Pelletier, D., Renault, E. (2006). Short run and long run causality in time series: Inference. *Journal of Econometrics* 132(2):337–362.
- Engle, R. F., Ito, T., Lin, W.-I. (1990). Meteor showers or heat waves? Heteroskedastic intra-daily volatility in the foreign exchange market. *Econometrica* 58(3):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 and Economic Statistics* 21(4):532–546.
- Florens, J. P., Mouchart, M. (1977). Reduction of bayesian experiments. Technical report, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.
- Florens, J. P., Mouchart, M. (1980). Initial and sequential reduction of bayesian experiments. Technical report, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.
- Florens, J. P., Mouchart, M. (1985). A linear theory for noncausality. *Econometrica* 53(1):157–176.
- Geweke, J. (1989). Bayesian inference in econometric models using Monte Carlo integration. *Econometrica* 57(6):1317–1339.
- Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: Bernardo, J. M., Berger, J. O., Dawid, A., Smith, A. F. M., eds. *Bayesian Statistics 4*, Vol. 148. Oxford: Clarendon Press.
- Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37(3):424–438.
- Hafner, C. M., Herwartz, H. (2008). Testing for causality in variance using multivariate GARCH models. *Annales d'Economie et de Statistique* 89:215–241.
- Hafner, C. M., Preminger, A. (2009). On asymptotic theory for multivariate GARCH models. *Journal of Multivariate Analysis* 100(9):2044–2054.
- Haugh, L. D. (1976). Checking the independence of two covariance-stationary time series: A univariate residual cross-correlation approach. *Journal of the American Statistical Association* 71(354):378–385.
- 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.



- Jarque, C. M., Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters* 6:255–259.
- Jeantheau, T. (1998). Strong consistency of estimators for multivariate ARCH models. *Econometric Theory* 14(01):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(1):11–25.
- Kim, T., White, H. (2004). On more Robust estimation of skewness and kurtosis. *Finance Research Letters* 1(1):56–73.
- Koutmos, G., Booth, G. (1995). Asymmetric volatility transmission in international stock markets. *Journal of International Money and Finance* 14(6):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(02):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(2–3):367–378.
- Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer: Berlin Heidelberg.
- Lütkepohl, H., Burda, M. M. (1997). Modified Wald tests under nonregular conditions. *Journal of Econometrics* 78(1):315–332.
- Marzec, J., Osiewalski, J. (2008). Bayesian inference on technology and cost efficiency. *Bank i Kredyt* 39(9):1–32.
- 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(1):147–163.
- Osiewalski, J., Pipień, M. (2002). Multivariate t-GARCH models - Bayesian analysis for exchange rates. In: *Modelling Economies in Transition - Proceedings of the Sixth AMFET Conference*. Absolwent, Łódź, pp. 151–167.
- 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.
- Pajor, A. (2011). A Bayesian analysis of exogeneity in models with latent variables. *Central European Journal of Economic Modelling and Econometrics* 3(2):49–73.
- Plummer, M., Best, N., Cowles, K., Vines, K. (2006). Coda: Convergence diagnosis and output analysis for MCMC. *R News* 6(1):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*. Cambridge, Massachusetts & London: The MIT Press, pp. 1–17.
- Sims, C. A. (1972). Money, income, and causality. *The American Economic Review* 62(4):540–552.
- Sims, C. A. (1980). Macroeconomics and reality. *Econometrica* 48(1):1–48.
- Taylor, M. P. (1995). The economics of exchange rates. *Journal of Economic Literature* 33(1):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.
- Woźniak, T. (2012a). Testing Causality Between Two Vectors in Multivariate GARCH Models. Working Paper Series, European University Institute, Florence, Italy.
- Woźniak, T. (2012b). Granger-Causal Analysis of VARMA-GARCH Models. Working Paper Series, European University Institute, Florence, Italy.
- Woźniak, T. (2015). Testing causality between two vectors in multivariate GARCH models. *International Journal of Forecasting* 31(3):876–894.