Central European Journal of Economic Modelling and Econometrics

Bayesian Comparison of Bivariate Copula-GARCH and MGARCH Models

Justyna Mokrzycka*

Submitted: 16.07.2018, Accepted: 9.03.2019

Abstract

The aim of the study is to formally compare the explanatory power of Copula-GARCH and MGARCH models. The models are estimated for logarithmic daily rates of return of two exchange rates: EUR/PLN, USD/PLN and stock market indices: SP500, BUX. The analysis is performed within the Bayesian framework. The posterior model probabilities point to AR(1)-tSBEKK(1,1) for the exchange rates and VAR(1)-tCopula-GARCH(1,1) for the stock market indices, as the superior specifications. If the marginal sampling distributions are different in terms of tail thickness, the Copula-GARCH models have higher explanatory power than the MGARCH models.

Keywords: Bayesian model comparison, Copula-GARCH model, Multivariate GARCH model, Monte Carlo Importance Sampling

JEL Classification: G15

*Cracow University of Economics; e-mail: justyna.mokrzycka@uek.krakow.pl; ORCID: 0000-0003-0953-1169

1 Introduction

Modelling volatility and dependence structure of financial time series is one of the most interesting areas of research for both theorists and practitioners of financial markets. The well-known and often used models in the field are different types of Multivariate GARCH (MGARCH) models; for a comprehensive review see, e.g., Francq and Zakoïan (2010), Tsay (2010). The second important class of multivariate volatility models are Multivariate Stochastic Volatility Models (MSV, see e.g. Tsay (2010), Pajor (2010)). The different and useful properties of MGARCH and MSV models inspired Osiewalski and Pajor (2007) to develop a hybrid MSF-DCC structure. The model unifies Engle's DCC covariance structure (Engle, 2002) and the simplest MSV specification (i.e. the Multiplicative Stochastic Factor models). While the conditional correlation matrix has the same form as the one in DCC, due to the presence of latent process the model allows for fatter tails than in the DCC structure. Osiewalski and Pajor (2009) extended this specification to a hybrid MSF-SBEKK. As shown in the cited paper, latent AR(1) processes, typically featured by the MSV models, are crucial in modelling tail behavior. Next, Osiewalski and Osiewalski (2016) further generalized MSF-SBEKK to a hybrid GMSF-SBEKK specification, introducing as many latent processes as there are relatively homogeneous groups of assets being modelled.
Another approach to modeling volatility is based on copulas as functions which capture the dependence structure. The n-dimensional copula C is a n-dimensional distribution function on $[0,1]^n$ with standard uniform marginal distributions. Due to Sklar's theorem (1959) the multivariate distribution function can be represented as a copula and the cumulative probability functions of marginal distributions. This representation is unique when the random vector has a continuous distribution (Sklar, 1959). The class of bivariate copulas is large. It contains elliptical distributions, as well as copulas in which the distributions allow for asymmetries governed by different values of tail dependencies coefficients (Joe, 1993). In addition, for some copulas, tail dependencies coefficients are simple functions of their parameters (Nelsen, 1999). Therefore, by means of the copula it is possible to take into account the asymmetries in the tails of the distribution and, moreover, to model them dynamically. Patton (2006) and Jondeau and Rocklinger (2006) combined simple univariate GARCH structures with conditional copulas, thereby formulating multivariate Copula-GARCH models, which are currently used in the analysis of financial time series.
To the best of our knowledge, a formal, Bayesian comparison of all the four types of multivariate volatility models (i.e. MGARCH, MSV, hybrid GMSF-SBEKK and Copula-GARCH) has not been made so far, although some research in the area (yet, of a lesser extent) was indeed presented in the literature. The Bayesian comparison of bivariate MGARCH models was presented by Osiewalski and Pipień (2004). The MGARCH, MSV and hybrid MSF-SBEKK structures were formally compared by Osiewalski and Pajor (2009). Moreover, Osiewalski and Osiewalski (2016) have attempted to compare the hybrid MSF-SBEKK and GMSF-SBEKK structures. The

confrontation of the MSV and hybrid GMSF-SBEKK models using the Bayesian inference is difficult because of latent processes, which usually generate numerical problems with calculating the marginal data density (MDD) value. Hence, we limit our present work to a Bayesian comparison of the MGARCH and Copula-GARCH models, still, however, contributing to the research on assessing multivariate volatility models in terms of their empirical adequateness.

It is worth noting that the formal Bayesian model comparison (used here) amounts to computing MDD values, which are overall characteristics of Bayesian models adequacy. So our approach does not indicate models which are the best in all aspects; the models which are overall leaders of our ranking may not be very accurate in some particular aspects – such as tail behavior and risk measurement.

Over the last decade quite numerous works about modelling financial time series with Copula-GARCH and MGARCH was published (e.g. Patton, 2006, Jondeau and Rocklinger, 2006, Dias and Embrechts, 2010). In some of them the authors attempted to empirically compare various specifications of the models at hand, usually by means of information criteria. For instance, Dias and Embrechts (2010) use Copula-GARCH and BEKK structures to model dependencies between exchange rates and analyse this model within the sampling-theory (i.e. non-Bayesian) approach (resorting to analysis of realized correlations, R^2). Weiss (2013) compares the performance of Copula-GARCH and Dynamic Conditional Correlation GARCH (DCC-GARCH) models in the context of Value at Risk (VaR) and Expected Shortfall (ES) predictions, whereas Grziska (2013) applies Copula-GARCH and MGARCH structures to model emerging markets. In these two studies a non-Bayesian framework was also utilised. To the best of author's knowledge, no formal Bayesian comparison of Copula-GARCH and MGARCH specifications has been presented in the relevant literature so far. Our paper aims to fill this gap by this very first attempt to use Bayesian statistical tools of model comparison for the structures in question.

In this work we first compare nineteen bivariate models: thirteen Copula-AR(1)-GARCH(1,1) models and two kinds of VAR(1)-tSBEKK(1,1), VAR(1)-tCCC(1,1) and VAR(1)-tDCC(1,1) specifications. We apply these structures to model the dynamics of the logarithmic daily growth rates of two exchange rates: EUR/PLN and USD/PLN, over the period from August 1, 2005 through September 21, 2015. In the second part of our empirical study we shift our focus on modelling two stock market indices (SP500 and BUX) covering the same period as for the exchange rates. Taking into account the results obtained in the first part, we omit the Copula-GARCH models with the static copulas, following our conjecture that these specifications would have almost zero posterior probability. We compare four bivariate models: VAR(1)-tCopula-GARCH(1,1) with dynamic t-Student copula, VAR(1)-tSBEKK(1,1), VAR(1)-tCCC(1,1) and VAR(1)-tDCC(1,1). We formulate formal Bayesian statistical models for the structures at hand, and then design a relevant Monte Carlo algorithm with Importance Sampling (MCIS) to calculate marginal data density value, which is crucial to perform Bayesian model comparison by means of posterior model probabilities. The calculations are performed by the author's own procedures developed in MATLAB.

Previously, a common choice to compute MDD was the harmonic mean estimator (HME, Newton and Raftery, 1994), corrected by Lenk (2009) and further developed by Pajor and Osiewalski (2013-14). Nevertheless, this method, as shown by Pajor (2017), is still biased, resulting in overestimated MDD values. To overcome the problems inherent to HME, Pajor (2017) proposed the corrected arithmetic mean estimator (CAME) which proves to have far better numerical properties than the corrected HME (Lenk 2009), although requiring simulation from the prior density and calculation of the probability over some set in the parameter space, which is not needed when using MCIS. In many different applications, also the Chib and Jeliazkov (2001) estimator is used. However, as shown by Osiewalski and Pipień (2004) in a simple simulation experiment with very long MCMC chains, while the HME estimator stabilised close to the true value, Chib and Jeliazkov's estimator led to either over- or underestimation. Relying on MCIS in this study makes our comparison very reliable from the numerical perspective, which is due to a small number of parameters (18 at most) and lack of latent processes.

We limit the scope of our research only to this two-dimensional (bivariate) case not only because of numerical tractability, but also due to the fact that the class of bivariate copulas features a greater variety of functions allowing for asymmetry in tail dependencies. In more than two-dimensional cases of the Copula-GARCH model, the possibility of flexible modelling of the various asymmetric dependency structures between random variables is limited by the copula parameters: there are usually one or two of them. In this situation the pair-copula construction is proposed (Czado, 2010).

The paper has the following structure. In Sections 2 and 3, the Bayesian Copula-GARCH and MGARCH models and, respectively, the basics of Bayesian model comparison are briefly discussed. In Section 4 we present the data and the empirical results of model estimation. The article ends with conclusions.

2 Bivariate Bayesian Copula-GARCH and MGARCH models

Let us consider a bivariate observation on return rates $y_t = (y_{1,t}, y_{2,t})'$, $t = 1, \dots, T$, which follows the VAR(1) process:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \varphi_{1,0} \\ \varphi_{2,0} \end{bmatrix} + \begin{bmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix}, \ t = 1, \dots, T. \quad (1)$$

The parameters of (1) are collected in $\varphi_0 = [\varphi_{1,0}, \varphi_{2,0}]'$ and $\varphi = \begin{bmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{bmatrix}$. We assume that all eigenvalues of φ lie inside the unit circle. The vector $z_t = [z_{1,t}, z_{2,t}]'$

represents a bivariate white noise defined as some conditionally heteroskedastic process.

Let $\theta\in\Theta\subset R^m$ be a vector of parameters consisting of the elements of φ_0 and φ as well as the parameters of the volatility process. The Bayesian statistical model is uniquely determined by the joint probability (density) function of observations and parameters:

$$p(y,\theta) = p(y|\theta)\, p(\theta),$$

where $y = [y_1, \ldots, y_T]$ is the matrix of observations, $p(y|\theta)$ is the sampling density and $p(\theta)$ represents the prior density. The posterior density of θ is

$$p(\theta|y) = \frac{p(y,\theta)}{p(y)} = \frac{p(y|\theta)p(\theta)}{p(y)},$$

where $p(y)$ is the marginal data density (MDD), defined as

$$p(y) = \int_{\Theta} p(y|\theta)\, p(\theta)\, d\theta.$$

In this work, whenever it is possible, the parameters are assumed *a priori* mutually independent. The following subsections present selected and alternative bivariate specifications of multivariate time-varying volatility processes and dependence structures.

2.1 Copula-GARCH(1,1) models

Let ψ_{t-1} be the set of information up to the moment $t-1$. Stochastic processes $\{z_{i,t}\}$, $i = 1, 2$ follow the GARCH(1,1) structure:

$$z_{i,t} = \varepsilon_{i,t}\sqrt{h_{i,t}}, \tag{2}$$

$$h_{i,t} = \alpha_{i,0} + \alpha_{i,1} z_{i,t-1}^2 + \beta_{i,1} h_{i,t-1},$$

where $\alpha_{i,0} > 0$, $\alpha_{i,1} \geq 0$, $\beta_{i,1} \geq 0$, $\alpha_{i,1} + \beta_{i,1} < 1$. For independent and identically distributed random variables $\varepsilon_{i,t}$ we assume either the symmetric or skewed t-Student distribution with zero mean and unit precision. Note that we do not standardise the noise, so that $E\left(\varepsilon_{i,t}^2\right) = \nu_i/(\nu_i - 2)$ if $\nu_i > 2$.

The Copula-GARCH model, proposed by Patton (2006) and Jondeau and Rocklinger (2006), uses a conditional copula and Sklar's theorem (Sklar, 1959) to describe dependence between the components of $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t})'$. The conditional density function of ε_t has the following representation:

$$\begin{aligned} p_{\varepsilon_t}(\varepsilon_{1,t}, \varepsilon_{2,t}|\psi_{t-1}) \quad = \quad & c\left(t_{\nu_1}(\varepsilon_{1,t}|\psi_{t-1}), t_{\nu_2}(\varepsilon_{2,t}|\psi_{t-1})|\psi_{t-1}\right) \times \\ & f_{t_{\nu_1}}(\varepsilon_{1,t}|\psi_{t-1})\, f_{t_{\nu_2}}(\varepsilon_{2,t}|\psi_{t-1}) \end{aligned}, \tag{3}$$

where $c(\cdot|\psi_{t-1})$ is the density of (either static or dynamic) conditional copula, whereas $t_{\nu_1}(\cdot|\psi_{t-1})$ and $f_{t_{\nu_1}}(\cdot|\psi_{t-1})$ are the univariate (symmetric or skewed) t-Student cumulative distribution and density function, respectively, with v_1 degree of freedom. This formula combines Sklar's theorem and Patton's work (2006), in which the theorem was transferred also into the context of conditional distributions.

In this research we use eleven time-invariant copulas: Frank, Gumbel, Clayton, Clayton-Gumbel, Rotated Gumbel, Rotated Clayton, Joe-Clayton, symmetrized Joe-Clayton, Normal, t-Student and copula of independent random variables. For exact formulae of the density functions of a variety of time-invariant copulas we refer the reader to Nelsen (1999) or Doman and Doman (2014). The above selection of copula functions includes cases allowing for both symmetric and asymmetric tail dependencies.

The tail dependencies coefficients of the copulas follow rather simple formulae dependent on the copula parameters. Let us consider some random variables having continuous distributions, and some copula C. Then the tail dependencies coefficients $(\lambda^U,\ \lambda^L)$ and the Kendall τ coefficient are calculated as follows:

$$\lambda^U = \lim_{\alpha\to 1^-} \frac{2\alpha - 1 + C(1-\alpha, 1-\alpha)}{\alpha},\ \lambda^L = \lim_{\alpha\to 0^+} \frac{C(\alpha,\alpha)}{\alpha},$$

$$\tau(X_1, X_2) = 4\int_{[0,1]^2} C(u_1,u_2)\, dC(u_1,u_2) - 1.$$

Exact formulae of the above coefficients for a selection of copulas can be found in Doman and Doman (2014).

Apart from the static copulas (constituting a majority of the selection listed above), in this research we also consider two kinds of time-varying copula: the normal and t-Student copulas (in the latter case, a time-invariant number of degrees of freedom is assumed). Let ρ be the parameter of either the normal or t-Student copula. To introduce time-variation into ρ, we follow the approach proposed by Tse and Tsui (2002):

$$\rho_t = (1-\alpha-\beta)\rho + \alpha\xi_{t-1} + \beta\rho_{t-1} \tag{4}$$

where $\alpha \geq 0$, $\beta \geq 0$, $\alpha+\beta < 1$ and $\xi_{t-1} = \int_{i=0}^{1} \varepsilon_{1,t-i}\varepsilon_{2,t-i} / \sqrt{\int_{i=0}^{1} \varepsilon_{1,t-i}^2 \varepsilon_{2,t-i}^2}$.

Finally, a general formulation of the Copula-AR(1)-GARCH(1,1) models can be obtained by combining (1) (with restriction $\varphi_{12} = \varphi_{21} = 0$) and (2), (3).

Let $\mu_{1,t} = \varphi_{1,0} + \varphi_{11} y_{1,t-1}$ and $\mu_{2,t} = \varphi_{2,0} + \varphi_{22} y_{2,t-1}$ be the conditional expectations of $y_{1,t}$ and $y_{2,t}$, respectively. Then the joint conditional density of y_t can be represented as

$$p_{y_t}(y_{1,t}, y_{2,t}|\psi_{t-1}) = p_{\varepsilon_t}\left(\frac{y_{1,t}-\mu_{1,t}}{\sqrt{h_{1,t}}}, \frac{y_{2,t}-\mu_{2,t}}{\sqrt{h_{2,t}}}|\psi_{t-1}\right) / \sqrt{h_{1,t}h_{2,t}}, \tag{5}$$

In order to formulate the Bayesian Copula-AR(1)-GARCH(1,1) model, let $\theta = (\theta_G, \theta_c)' \in \Theta \subset R^m$, where

$$\theta_G = (\varphi_{1,0}, \varphi_{11}, \alpha_{1,0}, \alpha_{1,1}, \beta_{1,1}, \gamma_1, \nu_1, \varphi_{2,0}, \varphi_{22}, \alpha_{2,0}, \alpha_{2,1}, \beta_{2,1}, \gamma_2, \nu_2)$$

is the vector of parameters of the AR-GARCH structure (with γ_i denoting the asymmetry parameter in the skewed t-Student distribution), while θ_c – the vector of the copula parameters. Then the joint sampling probability density function admits the form

$$p\,(y|\theta_G, \theta_c) = \prod_{t=1}^{T} p_{y_t}\,(y_{1,t}, y_{2,t}|\psi_{t-1}). \tag{6}$$

Apart from the likelihood function, given by (6), one also needs to specify the prior distribution, $p\,(\theta)$. As regards the choice of the prior for θ_G, we follow the works by Osiewalski and Pipień (1998):

$$p(\theta_G) = p\,(\varphi_{1,0}, \varphi_{2,0})\, p\,(\varphi_{11}, \varphi_{22}) \prod_{i=1}^{2} p\,(\alpha_{i,0})\, p(\alpha_{i,1}, \beta_{i,1}) p(\gamma_i) p(\nu_i),$$

$$p\,(\varphi_{1,0}, \varphi_{2,0}) = f_N\,(\varphi_{1,0}, \varphi_{2,0}|0, I_2)\,,\ p\,(\varphi_{11}, \varphi_{22}) = \frac{1}{4} I_{(-1,1)^2}\,(\varphi_{11},\ \varphi_{22}\ )\,,$$

$$p\,(\alpha_{i,0}) = f_{Exp}\,(\alpha_{i,0}|1)\,,\ p(\alpha_{i,1}, \beta_{i,1}) = \frac{1}{2} I_B\,(\alpha_{i,1}, \beta_{i,1})\,,$$

$$B = [0,1]^2 \cap \left\{(x,y)' :\ x + y < 1\right\},$$

$$p\,(\gamma_i) = f_{LN}\,(\gamma_i|0,1)\,,\ \text{a lognormal density with parameters} \mu = 0, \sigma = 1;$$

$$p\,(\nu_i) = \frac{1}{\sigma_\nu} \exp\left(-\frac{x - \mu_\nu}{\sigma_\nu}\right) I_{(\mu_\nu, \infty)}\,(\nu_i)\,,\ \mu = 2,\ \sigma = 8,\ E\,(\nu_i) = 10,\ i = 1, 2.$$

In the cases of static copulas, our prior distribution for θ_c coincides with the one specified by Mokrzycka and Pajor (2016). Otherwise, that is in models with dynamic copulas, for θ_c we use the same prior as the one proposed for the tDCC model by Osiewalski and Pipień (2005).
In our paper, the Copula-GARCH(1,1) models discussed above are formally compared with MGARCH structures.

2.2 tSBEKK(1,1), tCCC(1,1) and tDCC(1,1) models

As regards modelling volatility by means of Multivariate GARCH models, one of possible specifications of the conditional distribution of z_t is zero-centered bivariate t-Student distribution:

$$z_t|\psi_{t-1} \sim St\,(0, H_t, \nu)\,,\ H_t = \begin{bmatrix} h_{11,t} & h_{12,t} \\ h_{12,t} & h_{22,t} \end{bmatrix},$$

Table 1: List of prior distributions for the copulas' parameters

Copula	Parameter	Prior distribution
Frank	$\theta \in \mathbb{R}\backslash 0$	$\theta \sim N(0,100)$
Clayton	$\theta > 0$	$\theta \sim Exp(1)$
Gumbel	$\theta \geq 1$	$\theta \sim Exp(1,1)I_{(1,+\infty)}$
Clayton-Gumbel (BB1)	$\theta > 0$, $\delta \geq 1$	$\theta \sim Exp(1)$, $\delta \sim Exp(1,1)I_{(1,+\infty)}$
Joe-Clayton(BB7)	$\kappa \geq 1$,$\gamma > 0$	$\kappa \sim Exp(1,1)I_{(1,+\infty)}$, $\gamma \sim Exp(1)$
Symmetrized Joe-Clayton	$\kappa \geq 1$,$\gamma > 0$	$\kappa \sim Exp(1,1)I_{(1,+\infty)}$,$\gamma \sim Exp(1)$
Rotated Clayton	$\theta > 0$	$\theta \sim Exp(1)$
Rotated Gumbel	$\theta \geq 1$	$\theta \sim Exp(1,1)I_{(1,+\infty)}$
Normal	$\rho \in (-1,\ 1)$	$\rho \sim U(-1,\ 1)$
t-Student	$\rho \in (-1,\ 1)$,$v > 2$	$\rho \sim U(-1,\ 1)$,$v \sim Exp(2,8)I_{(2,+\infty)}$
Normal (time-varying)	$\rho \in (-1,1)$, $\alpha > 0$, $\beta > 0$, $\alpha + \beta \leq 1$	$\rho \sim U(-1,\ 1)$, $\alpha \sim U(0,\ 1)$, $\beta \sim U(0,\ 1)$
t-Student (time-varying)	$\rho \in (-1,\ 1)$, $\alpha > 0$, $\beta > 0$, $\alpha + \beta \leq 1$, $\nu > 2$	$\rho \sim U(-1,\ 1)$, $v \sim Exp(2,8)I_{(2,+\infty)}$, $\alpha \sim U(0,\ 1)$,, $\beta \sim U(0,\ 1)$

Note: $N(0,100)$ denotes normal distribution with zero mean and standard deviation of 10, $Exp(1)$ – exponential distribution with parameter 1, $Exp(a,b)I_{(a,+\infty)}$– truncated exponential distribution with mean $a+b$ and variance b^2; $U(A)$ – uniform distribution over A.

with H_t denoting the inverse precision matrix, while $\nu > 2$ is the degrees of freedom. Osiewalski *et al.* (2006) proposed to use some simple formula for time-varying H_t as a special case of the BEKK structure (Baba, Engle, Kraft and Kroner, 1989). This specification is termed "the scalar BEKK", and its formula in the two-dimensional case is as follows:

$$H_t = A + b^2 z_{t-1} z'_{t-1} + c^2 H_{t-1}, \tag{7}$$

where A is a positive semi-definite and symmetric matrix, while $b > 0$ and $0 < c < 1$ are some independent scalar parameters. This model is further referred to as tSBEKK(1,1). Following by Osiewalski *et al.* (2006), the priors for the elements of A are specified as: $a_{11} \sim Exp\,(1)$, $a_{22} \sim Exp\,(1)$, $a_{12} \sim N\,(0,1)$, whereas the priors of b, c and ν: $b{\sim}N(0.5,1)$, $c \sim N(0.5,1)$, $\nu \sim Exp\,(2,8)\,I_{(2,+\infty)}$, are truncated according to the aforementioned inequality restrictions. $Exp(a,\ b)I_{(a,+\infty)}$ denotes the truncated exponential distribution with mean $(a+b)$ and variance b^2.
In our paper, we also consider the constant conditional correlation (CCC) model of Bollerslev (1990), in which each of the conditional variances is described by a separate GARCH(1,1) process, while the conditional correlation coefficient is constant in time. The elements of H_t have the following formulae:

$$h_{11,t} = a_{10} + a_{11} z_{1,t-1}^2 + b_{11} h_{11,t-1},$$

$$h_{22,t} = a_{20} + a_{22} z^2_{2,t-1} + b_{22} h_{22,t-1},$$
$$h_{12,t} = \rho_{12} \sqrt{h_{11,t} h_{22,t}},$$

where ρ_{12} is the conditional correlation coefficient.
For the model-specific parameters we assume the same prior as in Osiewalski *et al.* (2006), i.e.

$$a_{10} \sim Exp(1),\ a_{20} \sim Exp(1),\ (a_{11},\ a_{22},\ b_{11},\ b_{22}) \sim U\left([0,1]^4\right),\ \rho_{12} \sim U([-1,1]),$$

with $a_{ii} + b_{ii} < 1$ $(i = 1, 2)$.
Another possible specification of H_t is the one proposed by Engle (2002). Similarly to the CCC structure, the conditional variances in Engle's model follow separate GARCH(1,1) processes, yet the conditional correlation coefficient is allowed to be time-varying and specified as:

$$\rho_{12,t} = \frac{q_{12,t}}{\sqrt{q_{11,t} q_{22,t}}},$$

where $q_{ij,t}$ are elements of a symmetric positive-definite matrix Q_t:

$$Q_t = (1 - \alpha - \beta) S + \alpha \xi_{t-1} \xi'_{t-1} + \beta Q_{t-1};\ \alpha \geq 0,\ \beta \geq 0,\ \alpha + \beta < 1,$$

with $S = [s_{ij}]_{i,j=1,\ 2}$ $(s_{12} = s_{21} = \rho)$ denoting the unconditional correlation matrix of standardized errors defined as

$$\xi_{i,t} = z_{i,t} \sqrt{\frac{\nu - 2}{\nu h_{ii,t}}},\ i = 1, 2$$

and collected in vector $\xi_t = (\xi_{1,t}, \xi_{2,t})'$. The above structure is termed as the dynamic conditional correlation (DCC) model.
For parameters of this model we specify the same prior as in Osiewalski and Pipień (2005):

$$s_{11} = s_{22} \sim U([-1,1]),\ (\alpha, \beta) \sim U\left([0,1]^2\right) \text{ with } \alpha + \beta < 1.$$

The logic behind the above prior distribution is to limit the prior information about those parameters.
As regards our specification of the models' conditional means, each volatility model (BEKK, CCC, DCC) is considered in two variants, being equipped with either a full VAR(1) structure given by (1) or separate AR(1) processes resulting from (1) under $\varphi_{12} = \varphi_{21} = 0$. For elements of φ_0 and φ, multivariate standardized Normal priors are assumed: $N(0, I_2)$ and $N(0, I_4)$, respectively.
More detailed discussions about the BEKK, CCC and DCC structures can be found in Baba *et al.* (1989), Bollerslev (1990), Engle (2002), and Osiewalski and Pipień(2005) or Osiewalski *et al.* (2006).

3 Bayesian model comparison

Let us consider m competing Bayesian models (for the same data collected in y) with parameters $\theta_{(i)}$:

$$M_i: \; p_i\left(y, \; \theta_{(i)}\right) = p_i\left(y|\theta_{(i)}\right) p_i\left(\theta_{(i)}\right), \; i = 1, \ldots, m.$$

We assume that the models are complementary and non-nested.
Comparison of competing Bayesian models is based on their posterior probabilities, calculated as follows:

$$p\left(M_i|y\right) = \frac{p(M_i)p(y|M_i)}{\sum_{j=1}^{m} p(M_j)p(y|M_j)}, \; i = 1, \ldots, m, \tag{8}$$

where $p(M_i)$ is the prior probability of model M_i. In this research we assume equal prior probabilities of each model, $p\left(M_i\right) = \frac{1}{m}$, $i = 1, \ldots, m$. The model with the highest posterior probability is considered the best model for explaining the data. More about Bayesian model comparison can be found in Osiewalski and Steel (1993).
Equation (8) hinges upon the marginal data density in each model:

$$p\left(y|M_i\right) = \int_{\theta_{(i)}} p_i\left(y|\theta_{(i)}\right) p_i\left(\theta_{(i)}\right) \, d\theta_{(i)}.$$

Due to a rather complicated form of the joint data and parameters' distribution, $p_i\left(y, \theta_{(i)}\right)$, calculation of MDD usually requires numerical techniques of integration. To that end, in our paper we resort to the Monte Carlo method with Importance Sampling (MCIS). The details of this approach can be found in Geweke (1989).
Generally, let us consider calculation of the expected value of some function of the parameters, $g\left(\theta\right)$. The MCIS method is based upon the following identity:

$$I = E\left(g\left(\theta\right)|y\right) = \int_{\Theta} g(\theta)p(\theta|y)d\theta = \frac{\int_{\Theta} g(\theta)f(\theta)d\theta}{\int_{\Theta} f(\theta)d\theta} = \frac{\int_{\Theta} g(\theta)\frac{f(\theta)}{s(\theta)}s(\theta)d\theta}{\int_{\Theta} \frac{f(\theta)}{s(\theta)}s(\theta)d\theta}$$

where $f\left(\theta\right) = p\left(y|\theta\right) p\left(\theta\right)$ and $s\left(\theta\right)$ is the density of some auxiliary distribution from which one can generate draws in a straightforward manner. Usually, $s\left(\theta\right)$ is named the importance function. The MCIS estimator of I has the following form:

$$\widehat{I_n} = \frac{1}{n} \sum_{i=1}^{n} \frac{g\left(\theta^{(i)}\right) f\left(\theta^{(i)}\right)}{s\left(\theta^{(i)}\right)} / \frac{1}{n} \sum_{i=1}^{n} \frac{f\left(\theta^{(i)}\right)}{s\left(\theta^{(i)}\right)} = \frac{\sum_{i=1}^{n} g\left(\theta^{(i)}\right) w\left(\theta^{(i)}\right)}{\sum_{i=1}^{n} w\left(\theta^{(i)}\right)},$$

where $w\left(\theta\right) = f(\theta)/s(\theta)$ is the weight function, and $\theta^{(i)}$ denotes the i-th draw from $s(\theta)$ $(i = 1, 2, \ldots, n)$. Under some additional assumptions this Monte Carlo estimator is consistent and asymptotically normal (Geweke, 1989).
The use of MCIS significantly simplifies the calculation of MDD, since its estimator

assumes a simple form of the arithmetic mean of the weights: $\frac{1}{n}\sum_{i=1}^{n} w(\theta^{(i)})$.
In this work, for the importance function we take the one of the multivariate t-Student distribution with 3 degrees of freedom, which remains in accordance with Evans and Swartz (1995) or Osiewalski and Pipień (1998). The other parameters of this distribution (i.e. the mean vector and covariance matrix) were estimated iteratively, based on at least 100,000 initial passes of the algorithm, with monitoring the numerical standard error (NSE), the relative numerical efficiency (RNE) and the variation coefficient of the weight function (γ_n):

$$\widehat{NSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(g\left(\theta^{(i)}\right) - \widehat{I_n}\right)^2 w^2\left(\theta^{(i)}\right)}{\left(\sum_{i=1}^{n} w\left(\theta^{(i)}\right)\right)^2}},$$

$$\widehat{RNE} = \frac{1}{n\widehat{NSE}^2}\left[\frac{\sum_{i=1}^{n}\left(g\left(\theta^{(i)}\right)\right)^2 w\left(\theta^{(i)}\right)}{\sum_{i=1}^{n} w\left(\theta^{(i)}\right)} - \widehat{I_n}^2\right]$$

$$\gamma_n = \frac{\sqrt{\sum_{i=1}^{n} w^2\left(\theta^{(i)}\right)/n - \left(\sum_{i=1}^{n} w\left(\theta^{(i)}\right)/n\right)^2}}{\frac{1}{n}\sum_{i=1}^{n} w\left(\theta^{(i)}\right)} = \sqrt{\frac{n\sum_{i=1}^{n} w^2\left(\theta^{(i)}\right)}{\left(\sum_{i=1}^{n} w\left(\theta^{(i)}\right)\right)^2} - 1}.$$

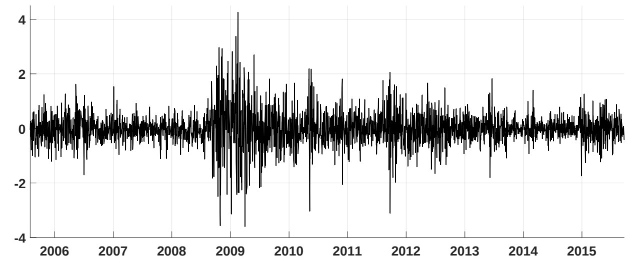
MCIS proves effective when $\widehat{NSE}$ converges to zero with an increase of n, while $\widehat{RNE}$ is close to one (then the draws are very close to the sample from the posterior distribution). The sequence of coefficients γ_n should stabilise with increasing n. Stable γ_n's smaller than one indicate very good quality of the importance function (Bauwens *et al.*, 1999).

4 Empirical study

In the first part of our empirical study we use the series of logarithmic daily returns on two exchange rates: EUR/PLN and USD/PLN, over the period August 1, 2005 till September 21, 2015. Nineteen different Bayesian models are under consideration: thirteen specifications of the Copula-GARCH models and six MGARCH structures. In the set of the Copula-GARCH structures we include the AR(1)-GARCH(1,1)-Copula models with symmetric t-Student conditional marginal distributions combined with the following copulas: Frank, Gumbel, Clayton, Joe-Clayton, Clayton-Gumbel, Rotated Gumbel, Rotated Clayton, Symmetrized Joe-Clayton, Normal (constant and time-varying), t-Student (constant and time-varying) and copula of independent random variables. Note that for the marginal distributions we choose the t-Student (symmetric and skewed) rather than normal

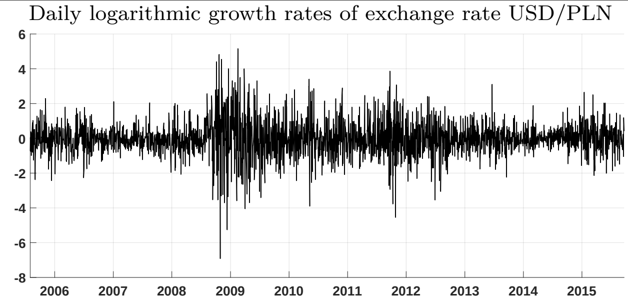
Figure 1: Daily logarithmic growth rates from August 1, 2005 to September 21, 2015

Daily logarithmic growth rates of exchange rate EUR/PLN



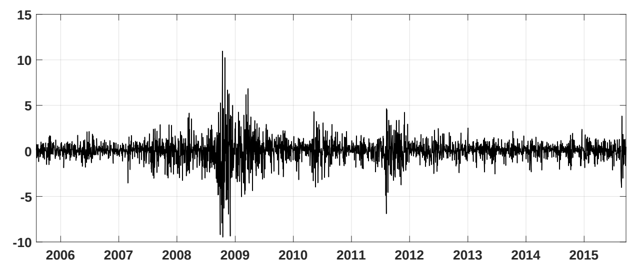
Mean: 0.0012, SD: 0.616, Skewness: 0.081; Kurtosis: 8.18; Min: −3.6, Max: 4.2

Daily logarithmic growth rates of exchange rate USD/PLN



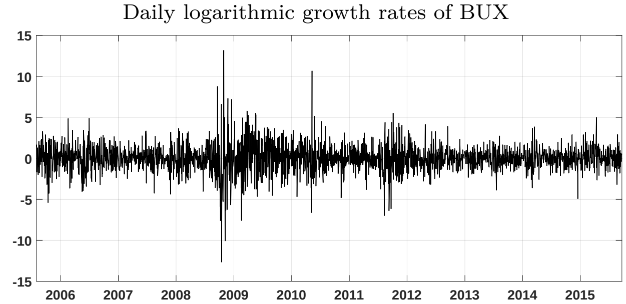
Mean: 0.0042, SD: 0.986, Skewness: 0.031; Kurtosis: 6.57; Min: −6.9, Max: 5.2

Daily logarithmic growth rates of SP500



Mean: 0.0178, SD: 1.27, Skewness: −0.305; Kurtosis: 14.15; Min: −9.5, Max: 10.9

Daily logarithmic growth rates of BUX



Mean: 0.0009, SD: 1.62, Skewness: −0.013; Kurtosis: 9.60; Min: −12.6, Max: 13.2

distribution, for the latter usually proves insufficient in terms of modelling the tails of the empirical data distribution. In the second part of our empirical study we analyse the series of logarithmic daily returns of two stock market indices: SP500 and BUX, over the same period as for exchange rates. This data set was modelled using only the VAR(1)-tCopula-GARCH(1,1) specifications with the dynamic t-Student copula. We omit the separate AR(1) structures because in the VAR(1) case the parameter φ_{12} is estimated as relatively far from zero. Moreover, as we mention in the introduction, we now omit the Copula-GARCH models with static copula.
The analysed MGARCH structures include: AR(1)-tSBEKK(1,1), AR(1)-tCCC(1,1), AR(1)-tDCC(1,1), VAR(1)-tSBEKK(1,1), VAR(1)-tCCC(1,1) and VAR(1)-tDCC(1,1). Selected estimation results for the models and their formal comparison (in terms of posterior probabilities) are presented in the following subsections.

Table 2: Posterior means and standard deviations of the tCopula-AR(1)-GARCH(1,1) models for exchange rates, along with numerical standard error (NSE) and relative numerical efficiency (RNE) of the means

	θ	Conditional marginal symmetric t-Student distributions				Conditional marginal skewed t-Student distributions			
		$E(\theta\|y)$	$D(\theta\|y)$	NSE	RNE	$E(\theta\|y)$	$D(\theta\|y)$	NSE	RNE
θ_G	$\varphi_{1,0}$	-0.017	0.008	2.2E-05	0.141	-0.041	0.013	3.3E-05	0.164
	φ_{11}	-0.010	0.016	3.9E-05	0.167	-0.013	0.016	4.2E-05	0.144
	$\alpha_{1,0}$	0.004	0.001	2.6E-06	0.134	0.004	0.001	2.9E-06	0.116
	$\alpha_{1,1}$	0.044	0.007	1.7E-05	0.158	0.044	0.007	2.0E-05	0.114
	$\beta_{1,1}$	0.920	0.011	3.1E-05	0.136	0.919	0.012	3.4E-05	0.114
	ν_1	7.799	1.036	4.3E-03	0.059	7.839	1.048	3.8E-03	0.077
	$\varphi_{2,0}$	-0.024	0.014	3.7E-05	0.147	-0.074	0.023	5.8E-05	0.159
	φ_{22}	-0.002	0.016	3.7E-05	0.181	-0.005	0.016	4.3E-05	0.135
	$\alpha_{2,0}$	0.007	0.002	4.5E-06	0.139	0.007	0.002	4.6E-06	0.132
	$\alpha_{2,1}$	0.035	0.005	1.2E-05	0.153	0.034	0.005	1.2E-05	0.143
	$\beta_{2,1}$	0.942	0.007	2.0E-05	0.139	0.942	0.007	2.1E-05	0.128
	ν_2	8.147	1.153	4.0E-03	0.081	8.257	1.205	6.7E-03	0.032
	γ_1	–	–	–	–	1.048	0.020	5.0E-05	0.162
	γ_2	–	–	–	–	1.057	0.021	5.2E-05	0.158
θ_c	$\theta_{c,1}$	0.746	0.010	2.5E-05	0.173	0.746	0.010	2.7E-05	0.147
	$\theta_{c,2}$	5.017	0.680	1.9E-03	0.133	5.056	0.685	2.1E-03	0.104

4.1 Data presentation and estimation results

In Figure 1 we present the modelled growth (or return) rate data on two exchange rates: EUR/PLN and USD/PLN, and two stock market indices: SP500 and BUX, along with basic descriptive statistics.

Table 3: Posterior means and standard deviations of the parameters of the VAR(1)-tCopula-GARCH(1,1) model with dynamic copula for the stock market indices, along with numerical standard error (NSE) and relative numerical efficiency (RNE) of the means

	θ	Conditional marginal symmetric t-Student distributions			
		$E(\theta\|y)$	$D(\theta\|y)$	NSE	RNE
θ_V	$\varphi_{1,0}$	0,083	0,014	9,6E-05	0,0217
	$\varphi_{2,0}$	0,031	0,022	1,7E-04	0,0181
	φ_{11}	-0,048	0,020	1,6E-04	0,0155
	φ_{12}	-0,006	0,012	7,6E-05	0,0237
	φ_{21}	0,246	0,028	2,5E-04	0,0125
	φ_{22}	-0,033	0,021	1,3E-04	0,0239
θ_G	$\alpha_{1,0}$	0,012	0,003	2,4E-05	0,0170
	$\alpha_{1,1}$	0,070	0,010	1,0E-04	0,0097
	$\beta_{1,1}$	0,879	0,015	1,3E-04	0,0141
	ν_1	5,438	0,630	4,5E-03	0,0193
	$\alpha_{2,0}$	0,038	0,011	8,1E-05	0,0178
	$\alpha_{2,1}$	0,071	0,011	7,8E-05	0,0195
	$\beta_{2,1}$	0,885	0,017	1,2E-04	0,0193
	ν_2	8,155	1,212	9,6E-03	0,0160
θ_c	$\theta_{c,1}$	16,062	5,048	5,1E-02	0,0097
	$\theta_{c,2}$	0,454	0,093	2,0E-03	0,0021
	$\theta_{c,3}$	0,018	0,007	6,7E-05	0,0102
	$\theta_{c,4}$	0,964	0,020	3,0E-04	0,0044

The time series in Figure 1 exhibit volatility clustering, which is typical to financial time series. The values of kurtosis: 8.18, 6.57, 14.15 and 9.60 indicate the leptokurtosis of empirical distributions (high concentration of the distribution around the modal value, accompanied by fat tails). For the conditional sampling distribution the t-Student density is assumed because as we can conclude with the basic descriptive statistics and empirical research (see, e.g., Osiewalski and Pipień 2004), the normal distribution would not be adequate.

The results presented below was obtained *via* the MCIS method discussed in the previous section. The estimation routine was monitored in terms of numeric errors, including the numerical standard error (NSE), the relative numerical efficiency (RNE) and the variation coefficient of the weights. The overall quality of the importance function is quite good. The variation coefficient of the weight function is stabilizing, but it is not less than one.

Table 2 displays the estimation results of the tCopula-AR(1)-GARCH(1,1) models for the logarithmic daily returns of exchange rates.

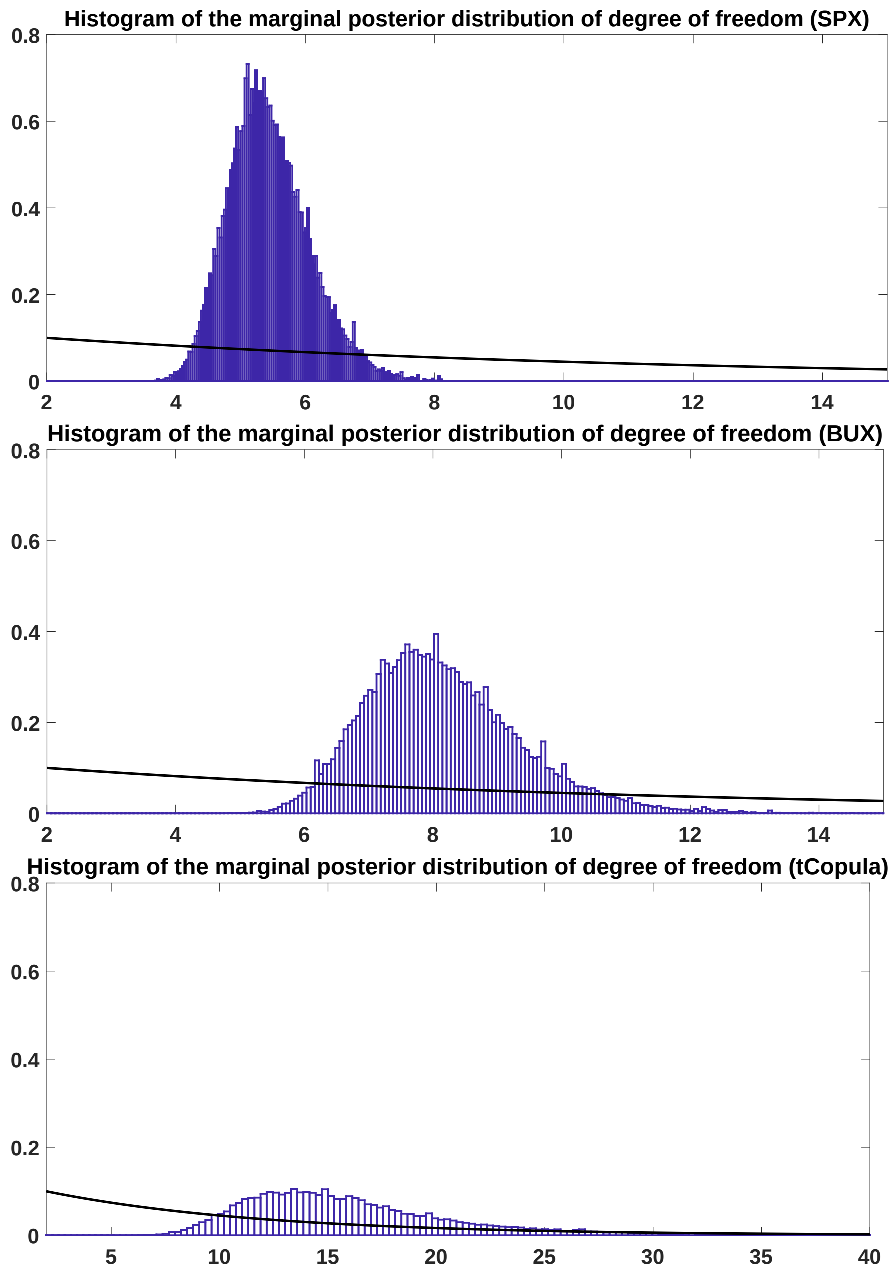
Table 4: Posterior means (and standard deviations) of the Kendall τ and tail dependencies coefficients in models for the exchange rates.

Copula	Conditional marginal symmetric t-Student distributions			Conditional marginal skewed t-Student distributions		
	Kendall τ	λ^U	λ^L	Kendall τ	λ^U	λ^L
Frank	0.5577 (0.0086)	0 (0)	0 (0)	0.5573 (0.0086)	0 (0)	0 (0)
Clayton	0.4573 (0.0093)	0 (0)	0.6626 (0.0103)	0.4593 (0.0107)	0 (0)	0.6648 (0.0177)
Rotated Clayton	0.4404 (0.0094)	0.6437 (0.0108)	0 (0)	0.4735 (0.0105)	0.68 (0.011)	0 (0)
Gumbel	0.5189 (0.0098)	0.6042 (0.0094)	0 (0)	0.5357 (0.0097)	0.6203 (0.0093)	0 (0)
Rotated Gumbel	0.5295 (0.0095)	0 (0)	6144 (0.0091)	0.5327 (0.0098)	0 (0)	0.6176 (0.0094)
Clayton-Gumbel, (BB1)	0.5855 (0.0137)	0.5117 (0.0152)	0.4247 (0.0327)	0.5827 (0.0137)	0.5333 (0.0158)	0.3595 (0.0428)
Joe-Clayton, (BB7)	0.5165 (0.0097)	0.582 (0.0149)	0.5554 (0.019)	0.5163 (0.0098)	0.6042 (0.0155)	0.5211 (0.023)
Symmetrized	–	0.5711 (0.0163)	0.569 (0.0171)	–	0.5946 (0.017)	0.5393 (0.0206)
Normal	0.5334 (0.0085)	0 (0)	0 (0)	0.5322 (0.0086)	0 (0)	0 (0)
t-Student	0.5365 (0.0099)	0.4761 (0.0199)	0.4761 (0.0199)	0.5359 (0.0099)	0.4744 (0.0202)	0.4744 (0.0202)

Note: λ^U-the upper tail dependence, λ^L-the lower tail dependence.

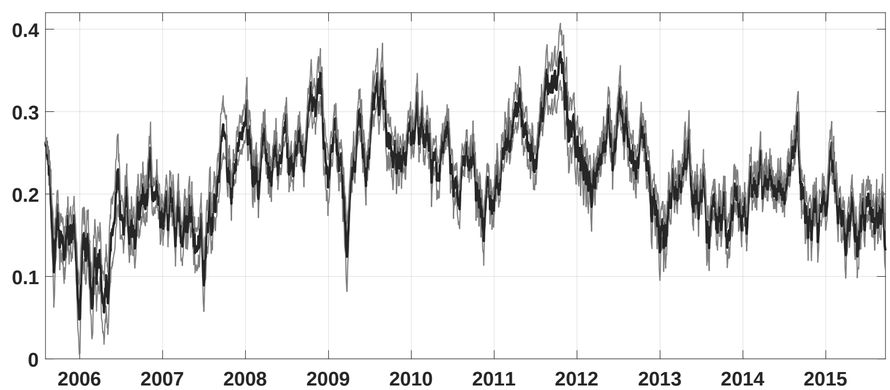
It is worth noting that the degrees of freedom for the univariate sampling distributions (ν_1, ν_2) and for the conditional copula $(\theta_{c,2})$ in this model are different, hence the copula's parameter $\theta_{c,1}$ cannot be interpreted as a linear correlation coefficient. With respect to the results for parameters γ_1 and γ_2, it may be conjectured that $\gamma_1 = 1$ and $\gamma_2 = 1$ fall into the highest posterior density intervals, thereby indicating no empirical need for allowing for skewness of the conditional distributions.
Table 3 presents the estimation results of the VAR(1)-tCopula-GARCH(1,1) model with dynamic copula for the logarithmic daily returns of stock market indices. It is worth noticing that the estimates of the degrees of freedom for the t-Student univariate sampling distributions are different. We can also notice discernible differences between their posterior distributions (see Figure 2). Therefore, we conclude that the tails of the sampling distributions differ substantially.
In Table 4 we present the posterior means and standard deviations of the Kendall τ and tail dependencies coefficients of twenty Copula-GARCH models for the exchange rates. All the results indicate positive dependence between the modelled time series. As regards the tail dependencies, the results appear somewhat less coherent, although in most cases symmetric tail behaviour is implied (with some exceptions of the models with the Clayton-Gumbel and Joe-Clayton copulas).

Figure 2: Histograms of the marginal posterior distributions of the degrees of freedom in the VAR(1)-tCopula-GARCH(1,1) model for the stock market indices



Figures 3 and 4 display the posterior means (with bands of two posterior standard deviations) of, respectively, the Kendall τ and tail dependencies coefficients in the VAR(1)-tCopula-GARCH(1,1) models with a dynamic copula. The values of the Kendall τ indicate that the static copula is not adequate (see Figure 3), thereby leading us to a conclusion that time-variability of the coefficients should not be dispensed with.
Also note that the upper and lower tail dependencies measures coincide in the case of the t-Student copula (see Figure 4). In this empirical study, the copula has quite high degree of freedom parameter, being indicative of a small value of the tails dependencies.

Figure 3: Posterior means (with bands of two posterior standard deviations) of the Kendall τ coefficient in the dynamic Bayesian VAR(1)-tCopula-GARCH(1,1) model for the stock markets indices



In Figure 5 we display the posterior means (with bands of two posterior standard deviations) of the constant and dynamic conditional correlation coefficients in, respectively, the AR(1)-tCCC(1,1) and AR(1)-tSBEKK(1,1) models. We choose to present the results for these two models, since the latter of them features the highest posterior model probability (as it is discussed in the following Subsection). For analogous reasons as in the case of Figures 3 and 4, also here one may infer that a time-variable pattern of conditional correlations should be accounted for.

4.2 Results of formal comparison of models

In this subsection we compare the analysed models in terms of their data fit performance measured by MDD values. To that end, three tables (5-7) are presented below, with each displaying the logarithms of MDD and the resulting posterior model probabilities (under the assumption of their equal prior probabilities). The three tables differ with respect to the set of models under comparison. Tables 5 and 6

Figure 4: Posterior means (with bands of two posterior standard deviations) of the tail dependencies coefficient in in the dynamic Bayesian VAR(1)-tCopula-GARCH(1,1) model for the stock markets indices

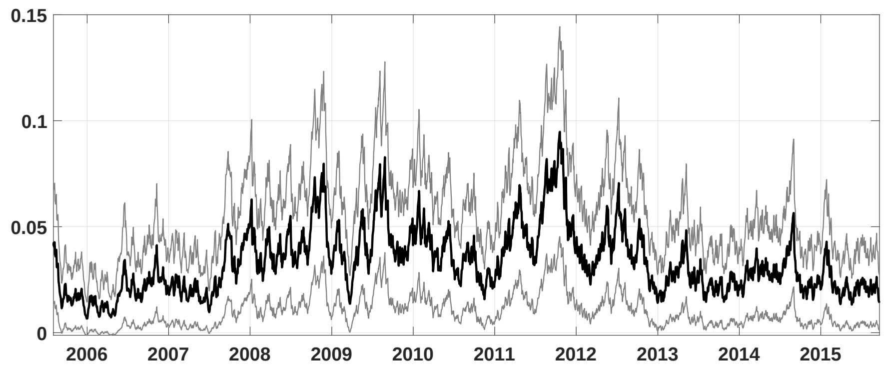
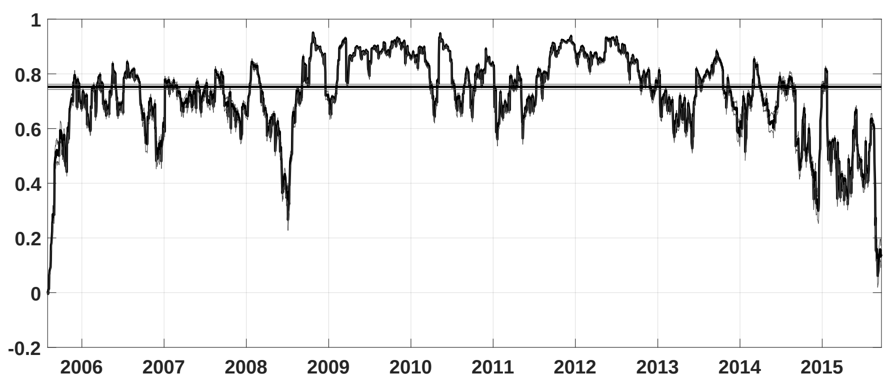


Figure 5: Posterior means (with bands of two posterior standard deviations) of the constant and dynamic conditional correlation coefficients in, respectively, the AR(1)-tCCC(1,1) and AR(1)-tSBEKK(1,1) models for the exchange rates



are for the case of exchange rates, and Table 7 concerns the case of stock market indices. In Table 5 we joined both sets of the models, i.e. the static-Copula-GARCH structures with either the symmetric or skewed t-Student conditionals, and we present the posterior model probabilities for the entire set. Next, Table 6 includes: eleven static-Copula-GARCH models with the symmetric t-Student conditionals (since conditional skewness appears empirically irrelevant for the data at hand, as it will be discussed below), two dynamic-Copula-GARCH structures with the symmetric t-Student conditionals, and six MGARCH specifications.

The results indicate that the tCopula-AR(1)-GARCH(1,1) specification wins within the class of eleven static-Copula-GARCH structures, with each of the remaining specifications gaining virtually zero posterior model probabilities. Note that the model with the independent copula is strongly rejected by the data, as evidenced by the lowest value of MDD.

Results displayed in Table 5 imply that allowing for conditional skewness in the static-Copula-GARCH models does not improve the model explanatory power. The tCopula-AR(1)-GARCH(1,1) specification with the symmetric t-Student conditionals gains almost entire posterior model probability, leaving all the remaining models far behind.

Table 5: Posterior probabilities of the Copula-AR(1)-GARCH(1,1) models for the exchange rates

No.	Copula	Conditional marginal distributions	$\ln(p(y\|M_i))$	$p(M_i\|y)$
1	Frank	t	$-4172,911$	0
2		t-sk	$-4170,854$	0
3	Clayton	t	$-4333,228$	0
4		t-sk	$-4340,842$	0
5	Gumbel	t	$-4136,109$	0
6		t-sk	$-4116,994$	0
7	Clayton-Gumbel	t	$-4079,713$	**0,0002**
8	(BB1)	t-sk	$-4081,203$	**0,0001**
9	Joe-Clayton	t	$-4097,268$	0
10	(BB7)	t-sk	$-4099,434$	0
11	Symmetrized	t	$-4095,495$	0
12	Joe-Clayton	t-sk	$-4098,123$	0
13	Rotated	t	$-4287,122$	0
14	Clayton	t-sk	$-4270,215$	0
15	Rotated	t	$-4148,386$	0
16	Gumbel	t-sk	$-4155,486$	0
17	Normal	t	$-4123,058$	0
18		t-sk	$-4126,36$	0
19	t-Student	t	$-4071,356$	**0,9686**
20		t-sk	$-4074,795$	**0,0311**
21	Independent	t	$-5128,058$	0
22		t-sk	$-5125,314$	0

Note: 't' – symmetric t-Student conditional marginal distributions, 't-sk'– skewed t-Student conditional marginal distribution, $\ln(p(y|M_i))$ – the logarithm of the marginal data density, $p(M_i|y)$ – the posterior probability of model M_i.

Table 6: Posterior probabilities of the Copula-AR(1)-GARCH(1,1) models with symmetric t-Student conditional distributions (no. 1–13), and the MGARCH models for the exchange rates

No.	Model	$\ln(p(y\|M_i))$	$p(M_i\|y)$
	Copula-AR(1)-GARCH(1,1)		
1	Frank	$-4172,91$	0
2	Clayton	$-4333,23$	0
3	Gumbel	$-4136,11$	0
4	Clayton-Gumbel, (BB1)	$-4079,71$	0
5	Joe-Clayton, (BB7)	$-4097,27$	0
6	Symmetrized Joe-Clayton	$-4095,50$	0
7	Rotated Clayton	$-4287,12$	0
8	Rotated Gumbel	$-4148,39$	0
9	Normal	$-4123,06$	0
10	Normal (time-varying)	$-4023,45$	0
11	t-Student	$-4071,36$	0
12	t-Student (time-varying)	$-3988,39$	0
13	Independent	-5128,07	0
14	**AR(1)-tSBEKK(1,1)**	$\mathbf{-3968,18}$	**0,9979**
15	AR(1)-tCCC(1,1)	$-4071,12$	0
16	AR(1)-tDCC(1,1)	$-3974,60$	0,0016
17	VAR(1)-tSBEKK(1,1)	$-3975,79$	0,0005
18	VAR(1)-tCCC(1,1)	$-4078,81$	0
19	VAR(1)-tDCC(1,1)	$-3982,05$	0

Note: $\ln(p(y|M_i))$ – the logarithm of the marginal data density, $p(M_i|y)$ – the posterior probability of model M_i.

Table 7: Posterior probabilities of the VAR(1)-tCopula-GARCH(1,1) model with symmetric t-Student conditional distributions, and the MGARCH models for the stock market indices

No.	Model	$\ln(p(y\|M_i))$	$p(M_i\|y)$
1	**VAR(1)-tCopula-GRACH(1,1) with time-varying copula**	$\mathbf{-7920,938}$	**0,997**
2	VAR(1)-tSBEKK(1,1)	$-7949,168$	0
3	VAR(1)-tCCC(1,1)	$-7931,957$	0
4	VAR(1)-tDCC(1,1)	$-7926,803$	0,003

Note: $\ln(p(y|M_i))$ – the logarithm of the marginal data density, $p(M_i|y)$ – the posterior probability of model M_i.

The ultimate comparison of the models considered in this paper is presented in Table 6 and 7, which also include several MGARCH structures. For the exchange rates, the highest and almost unit posterior model probability is scored by the AR(1)-tSBEKK(1,1) specification, with each of the remaining models having practically zero posterior probability (perhaps except for the the AR-tDCC and VAR-tSBEKK cases). Therefore, it appears that the data does not require the entire VAR structure to capture adequately the dynamics of the sampling conditional means. As regards volatility and conditional correlations, the constant conditional correlation specification is strongly rejected by the data. Although time-variant conditional correlations need to be allowed for, the specification of their dynamics should be kept relatively simple, for SBEKK is *a posteriori* preferred to DCC.
For the stock market indices, the highest and almost unit posterior model probability is scored by the VAR(1)-tCopula-GARCH(1,1) specification. This data is better described by a model allowing for different tail thickness of the univariate conditional sampling distributions. Note that a standard MGARCH structure does not allow for such a property, which in current situation makes the tCopula-GARCH model by far superior in terms of its explanatory power.

5 Conclusions

In this paper bivariate Bayesian Copula-AR(1)-GARCH(1,1) and VAR(1)-tSBEKK(1,1), VAR(1)-tDCC, VAR(1)-tCCC(1,1) models are employed to model logarithmic daily returns of two exchange rates: EUR/PLN and USD/PLN. Also, the bivariate Bayesian VAR(1)-tCopula-GARCH(1,1) model with dynamic t-Student copula and bivariate Bayesian VAR(1)-tSBEKK(1,1), VAR(1)-tDCC, VAR(1)-tCCC(1,1) specifications are used to model logarithmic daily returns of two stock market indices: SP500 and BUX. The main aim of the research was to formally compare the empirical performance of these models. To that end, we calculated the values of marginal data density using the Monte Carlo method with Importance Sampling, which also enabled us to approximate posterior means and standard deviations of relevant quantities in each model, with particular interest in the Kendall τ and tail dependencies coefficients. The results point to AR(1)-tSBEKK(1,1) as *a posteriori* the most likely model for the exchange data. The more complicated dependence structures, such as the Copula-AR(1)-GARCH(1, 1) or AR(1)/VAR(1)-tDCC models gain virtually zero posterior probability. Moreover, the results obtained in both sets of models, i.e. Copula-GARCH and MGARCH, strongly corroborate dynamic rather than constant dependency structures. Notice that the results for MGARCH are quite compatible with the ones presented by Osiewalski and Pipień (2004), where the t-BEKK(1,1) model ranks first. The authors estimated and compered ten Bayesian MGARCH models, with BEKK(1,1), VECH(1,1) and CCC(1,1) among others.

On the other hand, the second empirical study delivers qualitatively different results. For the stock markets data, the best explanatory power is achieved by the VAR(1)-tCopula-GARCH(1,1) model with dynamic t-Student copula. This indicates that the Copula-GARCH may prove superior to standard MGARCH specifications if the data requires different tail thickness for each univariate series.

It appears worthwhile to examine the models considered in this paper not only with respect to the in-sample fit, but also in terms of their forecasting power, for example in the context of Value at Risk and Expected Shortfall predictions, as in Pajor and Osiewalski (2012). We leave this exercise for further research. Also, any formal Bayesian comparison between dynamic copula specifications and the models from the GMSF-SBEKK class, proposed in Osiewalski and Osiewalski (2016), would be very interesting.

Acknowledgments

The author would like to thank Jacek Osiewalski, Anna Pajor and Łukasz Kwiatkowski for valuable comments on the manuscript. This work was supported by the funds of Ministry of Science and Higher Education granted to the Faculty of Finance and Law at Cracow University of Economics for research conducted by young researchers and PhD students. The work development was co-finansed by the funds granted to the Faculty of Finance and Law at Cracow University of Economics, within the framework of the subsidy for the maintenance of research potential.

References

[1] Baba Y., Engle R.F., Kraft D.F., Kroner K.F., (1989), Multivariate Simultaneous Generalized ARCH, Departament of Economics, *Working Paper*, University of California at San Diego.

[2] Bauwens L., Lubrano M., Richard J.F., (1999), *Bayesian inference in dynamic econometric models*, Oxfors University Press, New York.

[3] Bollerslev T., (1990), Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Approach, *Review of Economics and Statistics*, 72, 498–505.

[4] Chib S., Jeliazkov I., (2001), "Marginal likelihood from the Metropolis–Hastings output." *Journal of the American Statistical Association*, 96: 270–281.

[5] Czado C., (2010), Pair-Copula Construction of Multivariate Copulas, [in:] *Copula Theory and Its Application*, [eds.:] Jaworski P., Durante F., Hardle W., Rychlik T., Lecture Notes in Statistics-Proceedings, Springer, Berlin, 93–109.

[6] Dias A., Embrechts P., (2010), Modeling exchange rate dependence dynamics at different time horizons, *Journal of International Money and Finance*, 29, 8, 1687–1705.

[7] Doman M., Doman R., (2014), *Dynamika zależności na globalnym rynku finansowym (Dynamic of Dependence Structure on Global Financial Market)*, Difin SA, Warszawa.

[8] Engle R.F., (2002), Dynamic Conditional Correlation –A simple class of Multivariate GARCH Models, *Journal of Business and Economic Statistics*, 20, 339–350.

[9] Evans M., Swartz T., (1995), Methods for approximating integrals in statistics with special emphasis on Bayesian integration problems, *Statistical Science*, 10(3), 254–272.

[10] Francq C., Zakoïan J-M., (2010), *GARCH Models: Structure, Statistical Inference and Financial Applications*, John Wiley & Sons, United Kingdom.

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

[12] Grziska M., (2013), Multivariate GARCH and Dynamic Copula Models for Financial Time Series With an Application to Emerging Markets, *Dissertation an der Fakultät Mathematik*, Informatik und Statistik der Ludwig–Maximilians–Universität München, München, `https://edoc.ub.uni-muenchen.de/17921/1/Grziska_Martin.pdf` (access: 2017-12-08).

[13] Jondeau E., Rockinger M., (2006), The Copula-GARCH model of conditional dependencies: An international stock market application, *Journal of International Money and Finance*, 25, 827–853.

[14] Nelsen R.B., (1999), *An Introduction to Copulas*, Springer-Verlag, New York.

[15] Newton M., Raftery A. (1994), Approximate Bayesian inference by Weighted Likelihood Bootstrap (with discussion), *Journal of the Royal Statistical Society*, series B 56, 3–48.

[16] Mokrzycka J., Pajor A., (2016), Formalne porównanie modeli Copula-AR(1)-GRACH(1,1) dla subindeksów indeksu WIG, *Przegląd Statystyczny*, R. LXIII-2, 123–148.

[17] Osiewalski J., Osiewalski K., (2016), Hybrid MSV-MGARCH Models – General Remarks and the GMSF-SBEKK Specification, *Central European Journal of Economic Modelling and Econometrics*, 8, 241–271.

[18] Osiewalski J., Pajor A., (2007), Flexibility and parsimony in multivariate financial modelling: a hybrid bivariate DCC-SV model, [in:] *Financial Markets. Principles of Modeling, Forecasting and Decision-Making* (FindEcon Monograph Series No.3), [ed.:] W. Milo, P. Wdowinski, Łódz University Press, Łódz, 11–26.

[19] Osiewalski J., Pajor A., (2009), Bayesian Analysis for Hybrid MSF-SBEKK Models of Multivariate Volatility, *Central European Journal of Economic Modelling and Econometrics* 1, 179–202.

[20] Osiewalski J., Pajor A., Pipień M., (2006), Bayesian analysis of main bivariate GARCH and SV models for PLN/USD and PLN/DEM (1996-2001), *Dynamic Econometric Models*, 7, 25–35.

[21] Osiewalski J., Pipień M., (1999), Bayesian forecasting of exchange rates using GARCH models with skewed t conditional distributions, [in:] *Marcromodels'98, Proceedings of the 25-th International Conference*, vol. 2, 195–218; Absolwent, Łódź.

[22] 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.

[23] Osiewalski J., Pipień M., (2005), Bayesian analysis of dynamic conditional correlation using bivariate GARCH models, *Acta Universitatis Lodziensis Folia Oeconomica*, 192, 213–227.

[24] Osiewalski J., Steel M.F.J., (1993), A Bayesian perspective on model selection [publication in Spanish: Una perspectiva bayesiana en selección de modelos, *Caudernos Economicos*, 55/3, 327–351]; English version available at `http://www.cyfronet.krakow.pl/~eeosiewa/pubo.htm`.

[25] Pajor A., (2010), *Wielowymiarowe procesy wariancji stochastycznej w ekonometrii finansowej. Ujęcia Bayesowskie*, Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie, Kraków.

[26] Pajor A., Osiewalski J., (2012), Bayesian Value-at-Risk and Expected Shortfall for a Large Portfolio (Multi- and Univariate Approaches), *Acta Physica Polonica A*, 121, 2-B, 101–109.

[27] Pajor A., Osiewalski J., (2013-14), A note on Lenk's correction of the harmonic mean estimator, *Central European Journal of Economic Modelling and Econometrics* 5 (2013), 271–275; correction: vol. 6 (2014), 69.

[28] Patton A.J., (2006b), Modelling asymmetric exchange rate dependence, *International Economic Review*, 47(2), 527–556.

[29] Sklar A., (1959), Fonctions de répartition à n dimensions et leurs marges', *Publications de l'Institut de Statistique de L'Université de Paris*, 8, 229–231.

[30] Tsay R.S., (2010), *Analysis of financial time series*, John Wiley & Sons, New Jersey.

[31] Tse Y.K., Tsui A.K.C., (2002), A Multivariate GARCH Model with Time-Varying Correlations, *Journal of Business & Economic Statistics*, 20, 351–362.

[32] Weiss G.N.F., (2013), Copula-GARCH versus dynamic conditional correlation: an empirical study on VaR and ES forecasting accuracy, *Review of Quantitative Finance and Accounting*, 41, 2, 179–202.