ORTHOGONAL TRANSFORMATION OF COORDINATES IN COPULA M-GARCH MODELS — BAYESIAN ANALYSIS FOR WIG20 SPOT AND FUTURES RETURNS

MATEUSZ PIPIEŃ

Department of Econometrics and Operations Research, Cracow University of Economics
e-mail: eepipien@cyf-kr.edu.pl

ABSTRACT

We check the empirical importance of some generalisations of the conditional distribution in M-GARCH case. A copula M-GARCH model with coordinate free conditional distribution is considered, as a continuation of research concerning specification of the conditional distribution in multivariate volatility models, see Pipień (2007, 2010). The main advantage of the proposed family of probability distributions is that the coordinate axes, along which heavy tails and symmetry can be modelled, are subject to statistical inference. Along a set of specified coordinates both, linear and nonlinear dependence can be expressed in a decomposed form.

In the empirical part of the paper we considered a problem of modelling the dynamics of the returns on the spot and future quotations of the WIG20 index from the Warsaw Stock Exchange. On the basis of the posterior odds ratio we checked the data support of considered generalisation, comparing it with BEKK model with the conditional distribution simply constructed as a product of the univariate skewed components. Our example clearly showed the empirical importance of the proposed class of the coordinate free conditional distributions.

STRESZCZENIE

M. Pipień. *Wielowymiarowe modele Copula M-GARCH o rozkładach niezmienniczych na transformacje ortogonalne — bayesowska analiza dla notowań spot i futures indeksu WIG20.* Folia Oeconomica Cracoviensia 2012, 53: 21–40.

W artykule przedstawiono uogólnienie rozkładu warunkowego w wielowymiarowym modelu typu GARCH, oraz poddano empirycznej weryfikacji skonstruowany model. Praca stanowi kontynuację badań prowadzonych przez Pipienia (2007, 2010) nad właściwą specyfikacją rozkładów warunkowych wektora stóp zmian instrumentów finansowych. Zasadniczym elementem określającym giętkość rozważanej klasy wielowymiarowych rozkładów jest możliwość zmiany układu współrzędnych, i – tym samym — kierunków w przestrzeni obserwacji, według których grube ogony i asymetria rozkładu mogą występować empirycznie. Zgodnie z przyjętą orientacją w przestrzeni obserwacji, możliwe jest modelowanie zależności pomiędzy elementami wektora losowego, zarówno o charakterze liniowym (stosowana transformacja liniowa) jak i nieliniowym (funkcja powiązań, copula).

W części empirycznej przedstawiamy wyniki modelowania dynamicznych zależności pomiędzy zwrotami z notowania spot i futures indeksu WIG20. Uzyskane rezultaty wskazują na zasadność proponowanego uogólnienia stosowanego w modelu BEKK. Model z proponowanym typem rozkładu warunkowego uzyskuje silne potwierdzenie empiryczne, mierzone ilorazem szans a posteriori i wartością brzegowej gęstości wektora obserwacji.

KEY WORDS — SŁOWA KLUCZOWE

Bayes factors, multivariate GARCH models, coordinate free distributions, Householder matrices

czynnik Bayesa, wielowymiarowe modele GARCH, macierze Householdera

1. INTRODUCTION

Most of contributions involved with multivariate GARCH (M-GARCH) models — for a survey see Bauwens, Laurent and Rombouts (2006) — rely on the assumption of the conditional Gaussian distribution. In spite of the fact that the M-GARCH models are applied in modelling and predicting temporal dependence in the second-order moments, some other properties of the conditional distribution, like for example fat tails and skewness, are also very important. This result was confirmed by Bayesian comparison of GARCH-type models with normal and Student-t conditional distributions presented by Osiewalski and Pipień (2004). In terms of the model data support, measured by posterior odds ratio and posterior probabilities, they clearly showed that conditional normality is completely unrealistic in modelling financial time series. Hence, long journey beyond normality is necessary — see Genton (2004) — for better understanding the dependence structure between related time series in general, and between financial returns particularly.

In the presence of empirical analyses decisively rejecting conditional normal distribution, a few studies concentrated on the application of the conditional distributions that allow both for heavy tails and asymmetry within M-GARCH models. Some developments on this subject present Bauwens and Laurent (2005). Modern propositions of modelling volatility and conditional dependence between financial returns try to resolve the problem by complicating stochastic structure of the model rather, than generalising explicitly conditional distribution. Recently Osiewalski and Pajor (2009, 2010) propose MSF-BEKK model, as an example of the process attributed with both, the flexibility of the Stochastic Volatility family of models, and parsimony of parameterisation of simple M-GARCH covariance structures. Some other, more complicated multifactor processes has been recently proposed by Osiewalski and Osiewalski (2011, 2012). Those hybrid processes can outperform pure M-GARCH specification, even in the case of conditional normality. As an alternative to approach investigated by Osiewalski and Pajor (2009, 2010) one may consider an explicit generalisation of the conditional distribution, also leading to more empirically important specifications.

In modelling volatility and dynamic dependence of returns of different financial assets, a linear dependence is economically interpretable and popular. Standard empirical exercises in financial econometrics, like controlling and pricing risks, optimal portfolio allocation, analysing volatility transmission mechanism or contagion and building hedging strategies, rely on solutions that are strictly connected with measures of stochastic dependence of the linear nature. However last decade have seen particularly strong attention in modelling dependence in a nonlinear setting. One of the important topic of financial econometrics that made substantial progress during last decade, relates to making inference about measures of stochastic dependence that are alternatives to the conditional correlation.

It seems that both, definition of a nonstandard distribution of observables, and a more detailed analysis of dependence are crucial in proper modelling of financial returns. One of the approaches that may resolve to some extent both issues involves copula functions. The approach was intensively developed by Patton (2001, 2009), Jondeau and Rockinger (2006) and, in the case of Polish financial market, by Doman (2008), Doman and Doman (2009), Jaworski and Pitera (2012) and others. Vast empirical literature clearly indicate that volatility models built within framework of copula functions contribute substantially to standard empirical issues in financial econometrics stated above; see Embrechts, McNeil and Straumann (2002), Bradley and Taqqu (2004), Rodriguez (2007), Chavez-Demoulin and Embrechts (2010), Balkema, Nolde, Embrechts (2012).

The main goal of this paper is to check the empirical importance of some generalisations of the conditional distribution in M-GARCH case. We generalise the M-GARCH model proposed and empirically analysed by Pipień (2006, 2007) who applied a novel class of probability distributions, which is coordinate free in the sense formulated by Fang, Kotz and Ng (1990). Pipień (2010) considered a multivariate distribution with independent components, with skewness imposed according to the inverse probability integral transformations, discussed in details by Ferreira and Steel (2006) and Pipień (2006). In the next step, orthogonal transformation was incorporated in order to assure that fat tails and also possible skewness can be imposed along a set of coordinate axes. Consequently, the construct postulated the existence of a set of coordinate axes, along which the univariate components are independent and the densities of the marginal distributions are known analytically. Now we additionally consider a generalisation, by imposing copula function that captures possible dependence of nonlinear nature between elements of the random vector. The main advantage of the proposed family of probability distributions is that the coordinate axes are subject to statistical inference and can be very different from the ones defined by canonical basis. Along a set of coordinates, supported by the data, both, linear and nonlinear dependence can be modelled.

In the empirical part of the paper we consider the bivariate series of the returns on the spot and futures quotations of the WIG20 index (WIG20 and

FWIG20 instruments) covering the period from 21.12.1999 till 27.02.2008; $t = 2053$ observations. In modelling the conditional dependence of the components of the bivariate time series we consider Copula-BEKK(1,1) model with coordinate free conditional distribution according to the postulates of the construct. For a comparison we also consider some restricted cases, leading to the much simpler conditional distribution. We apply formal approach to test explanatory power of a set of competing specifications, based on the posterior odds ratio, and discuss superiority and possible practical usefulness of the considered coordinate free conditional distribution. Additionally the posterior inference about coordinate axes is also presented.

2. A CLASS OF COORDINATE FREE CONDITIONAL DISTRIBUTIONS

The main goal of this chapter is to present a family of multivariate skewed distributions and apply it in the multivariate GARCH setting. The basic notion considered here is the unified representation of the univariate skewness that applies inverse probability integral transformation, proposed initially by Ferreira and Steel (2006). We follow the setting presented in the univariate case by Pipień (2006, 2007) and by Pipień (2010) in multivariate case. The skewed version of originally symmetric and unimodal density $f(.|\theta)$ (with cumulative distribution function $F(.|\theta)$) can be defined as follows:

$$s(x \,|\theta,\eta)=f(x|\theta) \cdot p(F(x|\theta) \,|\eta), \text{ for } x \in R, \tag{1}$$

where $p(.|\eta)$ denotes the density of the distribution defined on the unit interval. The asymmetric distribution $s(.|\theta,\eta)$ is obtained by application of the density $p(.|\eta)$ as a weighting function of the density $f(.|\theta)$. The case, when $p(.|\eta) = 1$, restores symmetry. Any family of densities $p(.|\eta)$, for $\eta \in H$, defined over unit interval, is called skewness mechanism. For a review of skewing mechanisms that incorporate hidden truncation mechanism, some approaches based on the inverse scale factors, order statistics concept, Beta or Bernstein distribution transformation or a constructive method see Pipień (2006). The empirical importance of the conditional skewness in modelling the relationship between risk and return was also studied in the univariate case by Pipień (2007). Some recent developments confirm results presented by Pipień (2007) that it is possible to restore the relationship, mentioned above, once a highly nonstandard stochastic process is considered in volatility modelling; see for example Markov switching-in-mean Stochastic Volatility model, proposed by Kwiatkowski (2010).

Now let consider m-dimensional random vector $\varepsilon = (\varepsilon_1,...,\varepsilon_m)^{\textasciigrave}$ and let denote by $f_1(.|\theta_1),..., f_m(.|\theta_m)$ a set of unimodal (with mode at zero) univariate densities, parameterised by vectors $\theta_1,...,\theta_m$ respectively. In the first step, for $i = 1,...,m$, we

impose skewness mechanisms $p_i(.|\eta_i)$ on densities $f_i(.|\theta_i)$. Note that in general the construct does not require imposing the same type of skewness mechanism for each $i=1,\ldots,m$. For simplicity, in the empirical part of the paper, we consider the case, where the same skewness mechanism is considered for each of the coordinates. Possible different asymmetry effects will result from different values of parameters η_i. The resulting density $s_i(.|\theta_i,\eta_i)$ takes the form presented in equation (1):

$$s_i(x\,|\theta_i,\eta_i)=f_i(x|\theta_i)\cdot p_i(F_i(x|\theta_i)\,|\eta_i),\text{ for } x\in R \text{ and } i=1,\ldots,m,$$

where $F_i(.|\theta_i)$ denotes cumulative distribution function. Initially, for the random vector $\varepsilon=(\varepsilon_1,\ldots,\varepsilon_m)`$ we define the distribution with independent asymmetric components:

$$p(\varepsilon|\theta,\eta)=\prod_{i=1}^{m}s_i(\varepsilon_i\,|\,\theta_i,\eta_i),\qquad\qquad(2)$$

where $\theta=(\theta_1`,\ldots,\theta_m`)`$, $\eta=(\eta_1`,\ldots,\eta_m`)`$.

Pipień (2010) shows examples of distributions in bivariate cases indicating that possible outliers and asymmetry can be captured by distribution (2) only if those features of the data will occur along original coordinate axes, defined by canonical basis in R^m. Also, any family of distributions (2) is not closed with respect to the orthogonal transformations of the components. Hence, in order to improve flexibility of our class of distributions, a special mechanism that would make the coordinate axes varying is incorporated according to the idea proposed by Ferreira and Steel (2006). We provide it on the basis of the following linear (affine) transformation of the random vector ε:

$$y = A`\,\varepsilon + \mu \qquad\qquad(3$$

for a nonsingular matrix $A_{[m\times m]}$ and location vector $\mu_{[m\times 1]}\in R^m$. The density of the distribution of the random vector y is defined by the following formula:

$$p(y|\theta,\eta,A,\mu)=|\det(A)^{-1}|\prod_{i=1}^{m}s_i((y-\mu)'A_i^{-1}\,|\,\theta_i,\eta_i),\qquad\qquad(4)$$

where A_i^{-1} denotes the i-th column of A^{-1}. If the densities $f_i(.|\theta_i)$ are unimodal, with mode at zero, then the distribution the vector of y in (4) is unimodal, with mode defined by μ and skewing mechanisms $p_i(.|\eta_i)$. Transformation matrix A introduces the dependence between components of y, while η determines the skewness of the independent components of ε. Assuring the variability of the parameters, equation (4) generates a flexible class of multivariate distributions that is closed under orthogonal transformations. Hence, the construct (4) is coordi-

nate free, in the sense defined in Fang, Kotz and Ng (1990). In our approach we do not restrict the distribution to the case that A is a square root of the symmetric and positive definite covariance matrix. Consequently, practical application of specific families of multivariate distributions (4) requires interpreting the effect of the transformation matrix A. With no loss of generality let assume in (3) that $\mu = 0_{[mx1]}$:

$$y = A`\varepsilon.$$

According to the theorem presented in Golub and Van Loan (1993) any nonsingular matrix $A_{[mxm]}$ can be written as the product of mxm orthogonal matrix O_m and upper triangular matrix $U_{[mxm]}$ with positive diagonal elements:

$$A = O_m U,$$

and such a decomposition (called the QR decomposition) is unique. Now the results of the transformation matrix A can be considered in two steps:

$$y = A`\varepsilon = (O_m U)`\varepsilon = U`O_m`\varepsilon. \tag{5}$$

Initially, the random vector ε in (5) is subject to the rotation (if $\det O_m = 1$) or rotoinversion (if $\det O_m = -1$). Then the vector $\xi = O_m`\varepsilon$ is transformed according to the covariance-type linear transformation. The distribution of the vector ξ postulates that there exist a set of coordinate axes, along which the components of ξ are independent and the densities of the marginal distributions are known analytically. The main difference between distribution of ε and ξ is that those coordinate axes can vary from the axes defined by canonical basis in R^m. The distribution of y is then obtained by imposing scale transformation on the distribution of ξ, because matrix U can be interpreted as the Cholesky square root of the symmetric and positive definite matrix defining covariance structure.

A parametric sampling model that incorporates distributions described by equation (5) requires unique (one-to-one) parameterisation of the family of orthogonal matrices O_m in R^m. Also some restrictions have to be imposed, in order to assure identification. The one-to-one parameterisation was provided by Steward (1980) and Ferreira and Steel (2007), by an application of the Householder matrices decomposition. Let denote $\upsilon = (\upsilon_1,...,\upsilon_m)` \in R^m$, the m-dimensional column vector. The Householder matrix $H(\upsilon)$ (Householder reflection or Householder transformation) is defined as follows:

$$H(\upsilon) = I_m \frac{2}{\upsilon'\ \upsilon} \upsilon\upsilon'.$$

Golub and Van Loan (1983) show some useful properties of $H(\upsilon)$. Firstly, for each $\upsilon \in R^m$ $H(\upsilon)$ is orthogonal, and secondly $H(\upsilon) = H(-\upsilon) = H(a \cdot \upsilon)$, for any scalar $a \neq 0$. From the second property if we restrict the vector υ to the unit half sphere in R^m

(denoted by HS^{m-1}) we will keep the coverage of the whole family of Householder matrices. Parameterisation of the unit half sphere is easily obtained if we write down the vector $\upsilon_\omega = (\upsilon_1,...,\upsilon_m)` \in HS^{m-1}$ in polar coordinates:

$$\upsilon_1 = \sin(\omega_1),\ \upsilon_j = \sin(\omega_j)\prod_{s=1}^{j}\cos(\omega_s),\ \text{for } j < m,\quad \upsilon_m = \prod_{s=1}^{m-1}\cos(\omega_s), \tag{6}$$

where

$$\omega = (\omega_1,...,\omega_{m-1}) \in \Omega_m = \begin{cases} (\ \pi/2, \pi/2) & \text{if}\quad m = 2 \\ (0, \pi/2) \times (\ \pi/2, \pi/2)^{m-3} \times (\pi,\pi) & \text{if}\quad m > 2. \end{cases}$$

Now, for any $[m \times m]$ orthogonal matrix O_m with $\det O_m = -1^{m+1}$, there exist unique decomposition:

$$O_m = H(\omega^m)\cdot...\ H(\omega^2), \tag{7}$$

to $m{-}1$ Householder reflections $H(\omega^j)$ defined by vectors $\omega^j_{[m \times 1]}$ of the form:

$$\omega_j = (0_{m-j},\ \upsilon_{\omega j})`,\ j = 2,...,m,$$

for m-j dimensional vector of zeros, $0_{m-j} = (0,...,0)`$ if $j < m$ and for an empty vector for $j = m$. The vectors $\upsilon_{\omega j} \in HS^{j-1}$ are parameterised in terms of the polar coordinates applied in (6). The interesting case is $m = 2$, where the class of Householder reflections provide parametric family of orthogonal matrices of dimension $[2 \times 2]$ with identification restrictions imposed; see Steward (1980), Golub and Van Loan (1983).

3. ANOTHER STEP — INTRODUCING COPULA FUNCTIONS

Distribution of y, defined by the density (4), where $A = O_m U$, with orthogonal matrix O_m, parameterised according to decomposition (7), is obtained on the basis of the linear transformation of a random vector ε with the density (2). Consequently, only linear dependence between random variables, representing coordinates, can be modelled. Possible changes in coordinates that may be subject to statistical inference, enriched flexibility of the family, however the nature of dependence of elements of the vector y may still be linear. In order to model a more complicated dependence structure in vector y we follow the approach that involves copula functions.

Let consider a bivariate random variable $z = (z_1,z_2)`$, with cumulative density function (*cdf*) F and density function f, and with f_i and F_i the density and *cdf* of the marginal distribution of z_i respectively ($i = 1,2$). According to Sklar (1959), there exists a function $C:[0,1]^2 \rightarrow [0,1]$, with the following properties:

1. $C(u_1, u_2)$ is increasing in u_1 and u_2
2. $C(0, u_2) = C(u_1, 0) = 0$, $C(1, u_2) = u_2$, $C(u_1, 1) = u_1$
3. For each $(u_1, u_1`, u_2, u_2`) \in [0,1]^4$, such $u_1 < u_1`$ and $u_2 < u_2`$:

$$C(u_1`, u_2`) - C(u_1`, u_2) - C(u_1, u_2`) + C(u_1, u_2) \geq 0,$$

such:

$$F(z_1, z_2) = C(F_1(z_1), F_2(z_2)).$$

The density of the joint distribution of z (if exist) is defined as follows:

$$f(z_1, z_2) = f_1(z_1) \, f_2(z_2) \, c_d(F_1(z_1), F_2(z_2)),$$

where:

$$c_d(u_1, u_2) = \frac{\partial^2 C}{\partial u_1 \partial u_2}(u_1, u_2).$$

Function C is called copula, and restores dependence reflected in the joint distribution F, when marginal distributions F_1 and F_2 are considered. Function $c_d(\cdot, \cdot)$ is called the density of the copula C. In the case with $C(u_1, u_2) = u_1 u_2$, we have $F(z_1, z_2) = F_1(z_1) F_2(z_2)$, $c_d(u_1, u_2) = 1$ and $f(z_1, z_2) = f_1(z_1) \, f_2(z_2)$, hence $C(u_1, u_2) = u_1 u_2$ defines stochastic independence between z_1 and z_2. For detailed theory of copula functions and of the concept of measuring stochastic dependence within copula framework see Joe (1997) and Nelsen (2006).

Now, in the bivariate case ($m = 2$), we generalise our distribution of y, defined by the density (4), by incorporating copula function in the distribution of the random vector ε. We consider a random vector y of the form:

$$y = U`O_m`z, \tag{8}$$

with upper triangular matrix U and the orthogonal matrix O_m defined by (7) and the bivariate random variable z with the following density:

$$p(z|\theta, \eta, \theta_{cop}) = s_1(z_1|\theta_1, \eta_1) \, s_2(z_2|\theta_2, \eta_2) \, c_d(S_1(z_1), S_2(z_2)|\theta_{cop}), \tag{9}$$

for the density c_d of a particular copula function parameterised by the vector θ_{cop}, and skewed univariate densities s_i, considered initially in (2). In (9) by S_1 and S_2 we denote *cdf* functions of those skewed univariate distributions. Introducing copula function in the distribution of y, according to (9), provides another source of possible stochastic dependence in the random vector y, not involved with linear transformation with matrix A, considered initially. The case with $C(u_1, u_2) = u_1 u_2$ (or equivalently $c_d(u_1, u_2) = 1$) restores independence in the vector z, and hence

the distribution is defined just like for ε in (2). In this case only a linear dependence between coordinates of y can be modelled.

4. THE SET OF COMPETING SPECIFICATIONS

By y_j we denote the two-dimensional vector of logarithmic returns at time j, i.e. $y_j = (y_{j1}, y_{j2})$`, where $y_{ji} = 100\ln(x_{ji}/x_{j-1,i})$ and x_{ji} denotes the value of i-th financial instrument at time j. In order to model conditional dependence between components of y_j we assume the following structure:

$$y_j = H_j^{0.5}(\beta, \psi_{j-1})` H(\upsilon_\omega)` z_j, \; j = 1, ..., t, \tag{10}$$

where $\psi_{j-1} = (...., y_{j-2}, y_{j-1})$` denotes the information set at time $j$. Random variables $z_j = (z_{j1}, z_{j2})$` follow the distribution defined in (9), where components $s_i(.|\theta_i, \eta_i)$ are the skewed versions of the standardised Student-t densities with $\nu_i > 0$ degrees of freedom parameter (hence $\theta_i = \nu_i$), and skewness parameters η_i. Matrix $H(\upsilon_\omega)$ in (10) is a Householder reflection defined by:

$$H(\upsilon_\omega) = I_m 2\frac{\upsilon_\omega \upsilon_\omega '}{\upsilon_\omega ' \upsilon_\omega},$$

where $\upsilon_\omega = (\sin\omega_1, \cos\omega_1)$, and $\omega_1 \in (-\pi/2; \pi/2)$. Symmetric and positive definite matrix $H_j(\beta, \psi_{j-1})$ follows BEKK(1,1) specification:

$$H_j(\beta, \psi_{j-1}) = A + B \cdot y_{j-1} y_{j-1}` \cdot B` + C \cdot H_{j-1}(\beta, \psi_{j-2}) \cdot C`,$$

and β groups all required parameters, namely

$$\beta = (a_{11}, a_{12}, a_{22}, b_{11}, b_{12}, b_{21}, b_{22}, c_{11}, c_{12}, c_{21}, c_{22}).$$

Rewriting (10) in the following form:

$$y_j = W_j` z_j, \; j = 1, ..., t,$$

where $W_j = H(\upsilon_\omega) H_j^{0.5}(\beta, \psi_{j-1})$, just like in (8), we can formulate the conditional distribution of y_j (with respect to ψ_{j-1}) as a result of linear transformation of distribution of z_j, with transformation matrix W_j:

$$p(y_j | \psi_{j-1}, \nu_1, \nu_2, \eta_1, \eta_2, \omega_1, \beta, M_1) =$$
$$= |\det W_j|^{-1} s_1(y_j` W^{-1}_{j(1)} | \nu_1, \eta_1)\, s_2(y_j` W^{-1}_{j(2)} | \nu_2, \eta_2)\, c_d(S_1(y_j` W^{-1}_{j(1)}), S_2(y_j` W^{-1}_{j(2)}) | \theta_{cop}),$$

where $W^{-1}_{j(i)}$ denotes i-th column of W_j^{-1}, and $s_i(.|\nu_i, \eta_i)$ are skewed Student-t densities:

$$s_i(z\,|\,\nu_i,\eta_i) = f_{st}(z\,|\,0,1,\nu_i)\cdot p(F_{st}(z\,|\,0,1,\nu_i)\,|\,\eta_i),\ z\in R,$$

with the density and cdf of the standardised Student-t distribution with zero mode, unit precision and degrees of freedom parameter $\nu_i>0$ denoted by $f_{st}(.\,|\,0,1,\nu_i)$ and $F_{st}(.\,|\,0,1,\nu_i)$ respectively.

We considered five different single parameter copula functions, namely Gaussian, Clayton, Frank, Plackett and Gumbel, together with the case of no copula function. This gives us six competing sampling models collected in the set denoted by H_1. The analytic form of copulas and its densities can be found in Joe (1997) and Nelsen (2006). The set of copula function applied in the empirical part of the paper is restricted to only to the cases where only a single parameter in θ_{cop} describes dependence in the random vector. Some other copula functions attributed with richer parameterisation can be found in Joe (1997).

The sampling model is represented by the following product of the conditional densities:

$$p(y,y_f|\nu_1,\nu_2,\eta_1,\eta_2,\omega_1,\beta, M_1) = \prod_{j=1}^{t+k} p(y_j|\psi_{j-1},\nu_1,\nu_2,\eta_1,\eta_2,\omega_1,\beta,M_1), \qquad (11)$$

where $y = (y_1,...,y_t)$ denotes the matrix of observed daily returns, while $y_f = (y_{t+1},...,y_{t+k})$ groups forecasted observables. In order to complete Bayesian models, the prior distributions of all parameters must be stated. For the vector β we adopted prior used in Osiewalski and Pipień (2004), for skewness parameters η_i and degrees of freedom parameters ν_i we applied prior distribution studied by Pipień (2007). Since the orthogonal component $H(\upsilon_\omega)$ in (10) is parameterised by a single parameter $\omega_1\in(-\pi/2;\ \pi/2)$, we assumed for simplicity uniform prior over the whole interval. Less trivial probability distributions, with some interesting topological properties, adopted for a subset of the orthogonal matrices, were proposed by Steward (1980).

All prior densities, except the one imposed on the parameter $\omega_1\in(-\pi/2;\ \pi/2)$, were investigated previously in our papers. As it was clearly shown by Osiewalski and Pipień (2004) and Pipień (2007) the prior information included in the Bayesian models is very weak, as the prior distributions of parameters are very diffuse. For parameters in copula functions we imposed normal distributions truncated to the appropriate domain, with the prior mode at the point assuring independence. Consequently, we do not specify any type of dependence between coordinates and imposed appropriately diffused distributions. Consequently, the conclusions drawn from the empirical analysis does not seem to be biased by the prior knowledge, which is vague and not precisely stated in our case.

The main goal of the empirical part of the paper is to discuss the importance of orthogonal component $H(\upsilon_\omega)$ and its form with respect to the type of the copula function included in the sampling model. As an alternative to models

in class H_1 we also considered Copula-BEKK(1,1) specifications written in the following way:

$$y_j = H_j^{0.5}(\beta, \psi_{j-1})^{\grave{}} z_j, \, j = 1, ..., t, \tag{12}$$

with no orthogonal mechanism, changing coordinates, included. The assumptions concerning z_j and $H_j(\beta, \psi_{j-1})$ are remained unchanged. In particular the distribution of z_j may involve five different form of copula function and also may not involve copula. This gives us additional set of six competing specifications, denoted by H_0. The model 0 can be interpreted as a special case of (10), obtained by imposing zero restriction on Householder vector $\upsilon_\omega = (0,0)^{\grave{}}$, leading to the case, where $H(\upsilon_\omega) = I_2$.

5. EMPIRICAL ANALYSIS

In the empirical part of the paper we analyse bivariate time series of the logarithmic returns of the spot and futures quotations of the WIG20 index, covering the period from 21.12.1999 till 27.02.2008; $t = 2053$ observations. The dataset, depicted on Figure 1, together with some descriptive statistics, exemplifies rather complicated nature of the dependence between both univariate time series. The possible dependence is clearly determined by the coincidence of outliers, making the empirical distribution considerably more dispersed along first and second quarter of the Cartesian product, as compared with relative stronger concentration of daily returns of spot and futures quotations with different sing at the same day. The modelled time series covers rather long history of spot and futures trading on the Warsaw Stock Exchange. But, we cut the dataset at the end of the February 2008 in order to compare our results of model comparison with those presented in a much simpler model setting by Pipień (2010).

Another reason to focus on the considered time series is that possible empirical importance of copula function in sampling model received so far attention only during the financial crisis. There is vast literature suggesting that during last global financial crisis, the dependence between financial time series become very complicated and nonstandard. Hence, many authors clearly indicated that copula functions are a promising tool in modelling time series during crises and market crashes; see Bradley and Taqqu (2004), Rodriguez (2007), Patton (2009). However, there is a little evidence in favour of the existence of nonlinear dependence prior to the latest financial crisis.Consequently, we did not updated our dataset and focus on the pre-crisis period. The empirical importance of copula construct in the sampling model presented in this paper will be much greater, if the data support will be obtained on the basis of the time series that ends before global financial crisis.

Table 1 presents the results of model comparison. We considered 12 competing specifications, imposing 5 different copula functions (Normal, Clayton, Frank, Plackett and Gumbel) and no copula function. In all cases respectively, we considered existence of orthogonal transformation against conditional distribution with marginal densities for both series defined as simply skewed Student-t distribution. We denote by H_1 the subset of models with orthogonal transformation included, while by H_0 a class of Copula-BEKK models with no free coordinates in the conditional distribution. In Table 1 we put decimal logarithms of the mar-
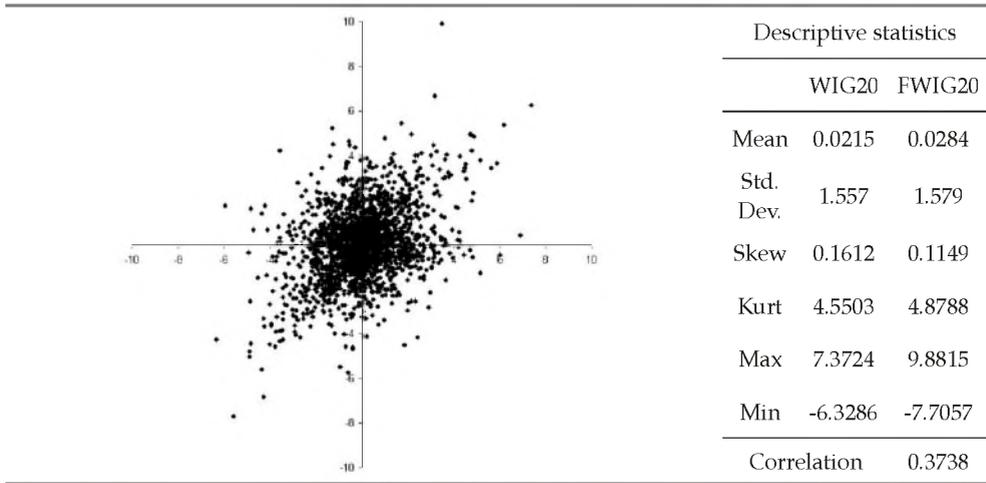

	Descriptive statistics	
	WIG20	FWIG20
Mean	0.0215	0.0284
Std. Dev.	1.557	1.579
Skew	0.1612	0.1149
Kurt	4.5503	4.8788
Max	7.3724	9.8815
Min	-6.3286	-7.7057
Correlation	0.3738	

Figure 1. The plot of the daily returns on WIG20 (vertical coordinate) and on FWIG20 (horizontal coordinate) from 21.12.1999 till 27.02.2008; t=2053 observations.

Table 1

Decimal logarithms of the marginal data density values in all competing specifications, and of the Bayes factor in favour of the existence of orthogonal component in model

Copula function applied in sampling model	Orthogonal component included (H_1)	No orthogonal component (H_0)	Bayes factor in favour of model from H_1 against model from H_0
No Copula	-2974.9263	-2977.5126	2.5863
Normal	-2971.2150	-2976.2267	5.0117
Clayton	-2972.2007	-2977.7896	5.5889
Frank	-2970.3253	-2973.0979	1.7726
Plackett	-2966.0346	-2968.1112	2.0766
Gumbel	-2973.3153	-2975.0409	4.1973

ginal data density values in case of all models, and also decimal logarithms of Bayes factors in favour of the existence of orthogonal component. The results clearly indicate the empirical importance of copula functions in sampling model. The model without the construct receives a little data support in both subsets H_1 and H_0 invariantly within H_1 and H_0 subset. The greatest data support both, in case of H_1 and H_0, receives model with Plackett copula incorporated in sampling function.

Another interesting issue concerning model comparison is that orthogonal component always improves the explanatory power of models. In case of no copula sampling models, and also for all copula functions, decimal logarithm of the Bayes factor against pure Copula-BEKK specification is greater than one, indicating in most cases the decisive support of this component in the sampling model. This result seems to be invariant with respect to all remained parts of the sampling model, and was suggested previously by Pipień (2010). Table 2 presents the results of posterior inference about tail parameters in all models. We focus on posterior mean and standard deviations of the degrees of freedom parameters of the conditional distributions of univariate series. Within subsets of models H_1 and H_0, the inference about the tails of the conditional distribution is relatively the same. In case of models, where orthogonal component excluded in the sampling model, posterior means of parameters ν_1 and ν_2 indicate that the conditional distribution is not of Gaussian type, however the posterior uncertainty, as measured by the posterior standard deviation, does not preclude strongly

Table 2

Posterior inference about tails of the conditional distribution in all competing specifications

| Copula function applied in sampling model	Orthogonal component included (subclass of models H_1)	No orthogonal component (subclass of models H_0)
No Copula	ν_1 5.64 (1.03) ν_2 18.93 (3.45)	ν_1 7.49 (1.98) ν_2 10.85 (1.98)
Normal	ν_1 6.94 (1.26) ν_2 18.37 (3.40)	ν_1 7.49 (1.35) ν_2 10.84 (1.97)
Clayton	ν_1 5.77 (1.05) ν_2 19.13 (3.49)	ν_1 7.25 (1.32) ν_2 11.00 (2.01)
Frank	ν_1 6.93 (1.27) ν_2 19.65 (3.59)	ν_1 8.57 (1.59) ν_2 11.42 (2.08)
Plackett	ν_1 6.61 (1.21) ν_2 18.95 (3.43)	ν_1 8.82 (1.61) ν_2 12.00 (2.20)
Gumbel	ν_1 5.51 (1.01) ν_2 19.34 (3.55)	ν_1 7.46 (1.33) ν_2 10.36 (1.83)

the same type of tail behaviour for both coordinates. If we include orthogonal transformation in sampling model, the posterior inference changes substantially, but in the same way in case of all copula functions and also in no-copula case. If we consider the mechanism that enables search for a set of coordinates, along which possible heavy tails and asymmetry can be modelled, the results of estimation of the properties of the conditional distribution in tails are changing. In all cases in subset H_1, invariantly with respect to the type of copula function, tails of the conditional distribution of univariate coordinates are different. The data clearly support heavy tails for the first coordinate, while the second one exhibit the Gaussian type tails.

In order to illustrate changes in conditional distribution, when orthogonal mechanism is included in the sampling model, we plotted the isodensities of z_j in case of models from subset H_0 (Table 3) and isodensities of a random variable $H(v_\omega)`z_j$ in case of models from subset H_1 (Table 4). All parameters required to draw the plots we chosen as posterior means. On the plots in Table 3 and 4, we draw vectors representing coordinates appropriate in sampling models. In case of models from subset H_0 we draw vectors proportional to the vectors from canonical basis in R^2, namely $e_1=(10,0)$ and $e_2=(0,10)$. In case of model from H_1 (Table 4) a set of coordinates are subject to posterior inference and hence we present posterior means, together with the bands of the 95% HPD (Highest Posterior Density) intervals for $H(v_\omega)`e_1$ and $H(v_\omega)`e_2$ respectively.

Analysing isodensities plotted in Table 3 and 4 it is clear that the data support different directions, than canonical, along which heavy tails and possible asymmetry can be modelled. Copula functions change the shape of isodensities strongly. However the most important feature of the sampling model seems to be the existence of the orthogonal mechanism changing coordinates. Only in case of models from subset H_1, a more complicated dependence between observed time series can be discovered, as the shapes of isodensities in Table 4 exhibit considerable excess from regular "elliptical" shape. For models from subset H_0, without orthogonal mechanism, differences between shapes of isodensities of the distribution of z_j are rather minor among models. New, estimated, directions in the sampling models from subset H_1 (Table 4) are different from initial, canonical, ones. Taking into account dispersion of the posterior distribution, the bands of the HPD intervals for $H(v_\omega)`e_1$ and $H(v_\omega)`e_2$ are located far away from the case, where $H(v_\omega)=I_2$. This clearly makes models without orthogonal component improbable in the view of the data. Additionally, changing directions in models from subset H_1 is nontrivial and does not only involve rotation. Comparing vectors e_1 and e_2 with its corresponding images, we see that canonical basis is subject to inversion and then to appropriate clock-wise rotation. This is due to the properties of the Householder reflections applied in the construct. It enables to search for optimal orientation in a more composed way.

Table 3

The plots of the isodensities of z_j in models from class H_0, i.e. in sampling models with
no orthogonal component included. Isodensities are plotted on the basis of values of parameters equal to posterior means

No Copula	Normal	Clayton
Frank	Plackett	Gumbel

Table 4   36
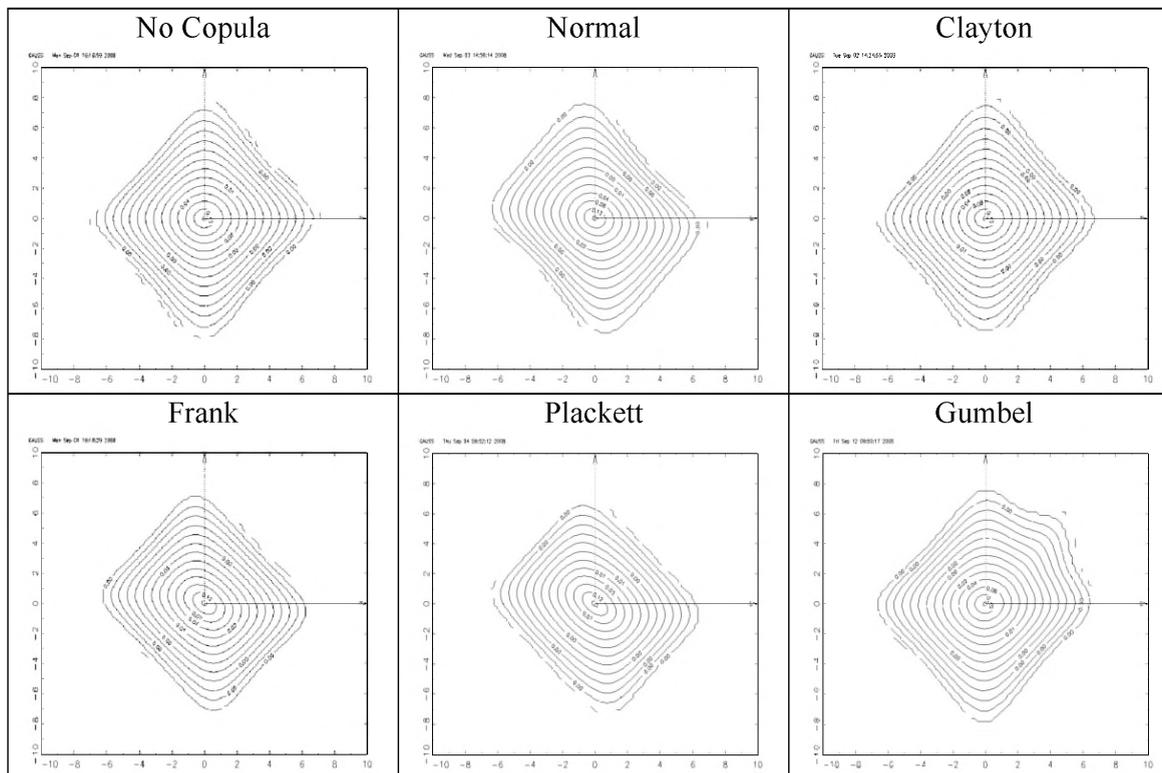
The plots of the isodensities of $H(v_\omega)`z_j$ in models from class H_0, i.e. in sampling models with no orthogonal component included. Isodensities are plotted on the basis of values of parameters equal to posterior means
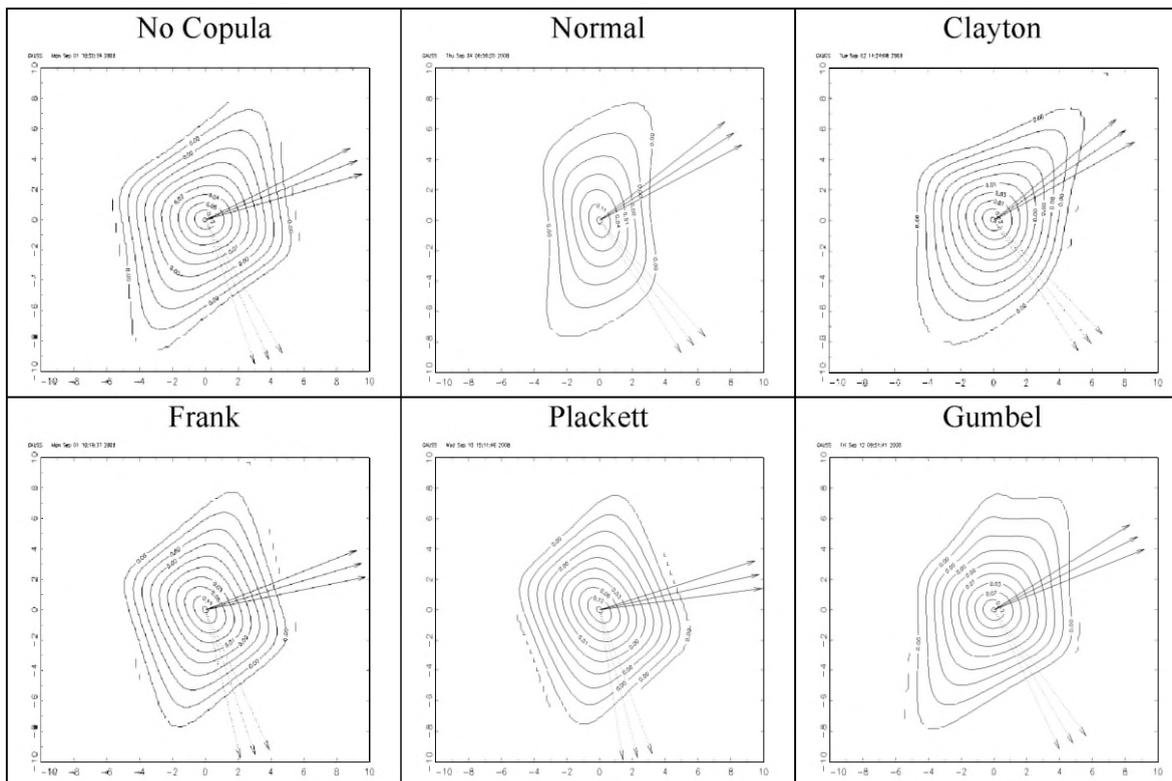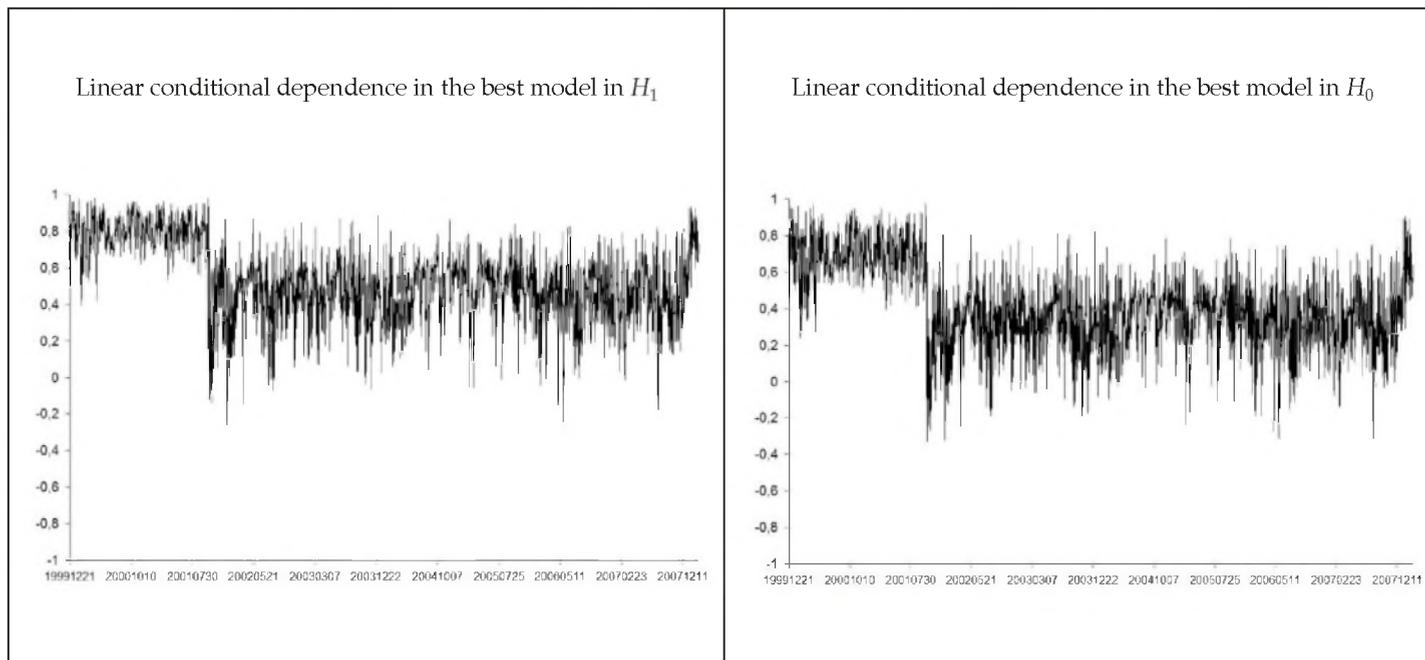
Table 5

Posterior inference about linear conditional dependence obtained on the basis of the elements of matrix $H_j(\boldsymbol{\beta}, \boldsymbol{\psi}_{j-1})$
in case of the best copula function (Placket). All parameters assumed to be equal to posterior means

A very important question concerning discussed empirical analysis involves possible conclusions about changes of the linear dependence between modelled univariate series, when orthogonal component and copula function is incorporated. In Table 5 we present plots of posterior expectations of conditional correlations between returns of spot and futures quotations of WIG20. Since the results are practically the same in case of all pairs of models, we focus our attention on the best models in H_1 and H_0 respectively, both based on Plackett copula function. In case of the best model from the set H_1 the variability of the conditional correlation coefficient seems to be only slightly less variable during the whole time interval covering modelled time series.

Existence of orthogonal mechanism in sampling model does not seem to influence the dynamics of conditional linear dependence strongly. Both series of posterior expectations exhibit the same dynamic pattern, with strong variability around value 0.4, starting from August the 1st 2001, when Warsaw Stock Exchange quoted WIG20 index officially for the first time.

6. CONCLUDING REMARKS

The main goal of this paper was to check the empirical importance of some generalisations of the conditional distribution in M-GARCH case. We considered copula M-GARCH model with coordinate free conditional distribution. We continue research concerning specification of the conditional distribution in multivariate volatility models started by Pipień (2007, 2010). The main advantage of the proposed family of probability distributions is that the coordinate axes, along which heavy tails and symmetry can be modelled, are subject to statistical inference. Along a set of specified coordinates both, linear and nonlinear dependence can be expressed in formal and composed form.

In the empirical part of the paper we considered a problem of modelling the dynamics of the returns on the spot and future quotations of the WIG20 index from the Warsaw Stock Exchange. On the basis of the posterior odds ratio we checked the data support of considered generalisation, comparing it with BEKK model with the conditional distribution simply constructed as a product of the univariate skewed components.

Our example clearly showed the empirical importance of the proposed class of the coordinate free conditional distributions. Both, orthogonal component, and copula function, are necessary in proper modelling of the conditional distribution of the vector financial returns. The existence of the orthogonal transformation of coordinates in observation space receives decisive data support invariantly with respect to the existence copula function in the sampling model and to the type of specified copula. The dataset support much different orientation in the sample space along which heavy tails, asymmetry and dependence between

coordinates, can be discovered. Among the class of copula function Plackett one received the greatest data support. Generally, presented in the empirical part of the paper noticeable flexibility of the class in directional modelling of the tails and asymmetry suggests that possible applications, concerning futures hedging or Value-at-Risk calculation, are very promising.

REFERENCES

Azzalini A. (1985), A Class of Distributions which Includes the Normal Ones, *Scandinavian Journal of Statistics* 12, 171–178.

Balkema G., Nolde N., Embrechts P (2012), The shape of asymptotic dependence, [in:] *Prokhorov and Contemporary Probability Theory* [eds.:] A. Shiryaev, S. Varadhan, and E. Presman, Springer, Berlin.

Bauwens L., Laurent S. (2005), A New Class of Multivariate Skew Densities with Application to Generalised Autoregressive Conditional Heteroscedasticity Models, *Journal of Business and Economic Statistics* 23, 346–254.

Bauwens L., Laurent S., Rombouts J. (2006), Multivariate GARCH Models: A Survey, *Journal of Applied Econometrics* 21, 79–109.

Bradley B. O., Taqqu M.S. (2004), Framework for Analyzing Spatial Contagion between Financial Markets, *Finance Letters* 2, 8–16.

Chavez-Demoulin V, Embrechts P (2010), Revisiting the edge, ten years on, *Communications in Statistics — Theory and Methods* 39, 1674–1688.

Doman R. (2008), Modelling Conditional Dependencies between Polish Financial Returns with Markov Switching Copula Models, *Dynamic Econometric Models* 8, 21–28.

Doman R., Doman M. (2010), Copula Based Impulse Respone Analysis of Linkages between Stock Markets, *Risk Management eJournal* 05/2010; DOI:10.2139/ssrn.1615108.

Embrechts P, McNeil A., Straumann, D. (2002), Correlation and Dependence in Risk Management: Properties and Pitfalls [in:] *Risk Management: Value at Risk and Beyond*, [ed.] Dempster M.A.H., Cambridge University Press, Cambridge, 176–223.

Fang K.-T, Kotz S., Ng K.-W. (1990), *Symmetric Multivariate and Related Distributions*, Chapman and Hall, New York.

Ferreira J. T. A. S, Steel M. F. J. (2006), A Constructive Representation of Univariate Skewed Distributions, *Journal of the American Statistical Association* 101, 823–839.

Ferreira J. T A. S, Steel M. F. J. (2007), Model Comparison for Coordinate-Free Multivariate Skewed Distributions with an Application to Stochastic Frontiers, *Journal of Econometrics* 137, 641–673.

Genton M. G. (2004), *Skew Elliptical Distributions and Their Applications: A Journey Beyond Normality*, Chapman & Hall, London.

Golub G.H., Van Loan C.F. (1983), *Matrix Computations*, John Hopkins University Press, Baltimore.

Jaworski P, Pitera M. (2012), On Spatial Contagion and mGARCH Models, *Institute of Mathematics, University of Warsaw Working Paper.*

Joe H. (1997), *Multivariate Models and Dependence Concepts*, Chapman&Hall, London.

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.

Kwiatkowski Ł. (2010), Markov Switching In Mean Effect. Bayesian Analysis in Stochastic Volatility Framework, *Central European Journal of Economic Modelling and Econometrics* 2, 59–94.

Nelsen R. B. (2006), *An Introduction to Copulas*, Springer, Berlin.

Osiewalski J., Osiewalski K. (2011), Modele hybrydowe MSV-MGARCH z trzema procesami ukrytymi w badaniu zmienności cen na różnych rynkach, *Folia Oeconomica Cracoviensia* 52, 71–85.

Osiewalski J., Osiewalski K. (2012), Modele hybrydowe z dwoma procesami ukrytymi, *Zeszyty Naukowe UEK, Seria Finanse* 895, (w druku).

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.

Osiewalski J., Pajor A. (2010), Bayesian Value-at-Risk for a Portfolio: Multi- and Univariate Approaches using MSF–SBEKK Models, *Central European Journal of Economic Modelling and Econometrics* 2 (2010), 253–277.

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.

Patton A. (2009), Copula-Based Models for Financial Time Series. [in:] Andersen T.G., Davis R.A., Kreiss J. P, Mikosch T., [eds], *Handbook of Financial Time Series*, Springer, Berlin, 767–785.

Patton A. J. (2001), *Applications of Copula Theory in Financial Econometrics*, Unpublished Ph.D. dissertation, University of California, San Diego.

Pipień M. (2006), Bayesian Comparison of GARCH Processes with Skewness Mechanism in Conditional Distributions, *Acta Physica Polonica* B 37, 3105–3121.

Pipień M. (2007), An Approach to Measuring the Relation between Risk and Return. Bayesian Analysis for WIG Data, *Folia Oeconomica Cracoviensia* 48, 97–119.

Pipień M. (2010), A Coordinate Free Conditional Distributions in Multivariate GARCH Models, [in:] *Financial Markets. Principles of Modelling Forecasting and Decision Making*, FindEcon Conference Monograph Series No. 8, Łódź University Press, Łódź, 99–111.

Rodriguez J. C. (2007), Measuring Financial Contagion: a Copula Approach. *Journal of Empirical Finance* 14, 401–423.

Sklar A. (1959), Fonctions de répartition à n dimensions et leurs marges, *Publications de l' Institut Statistique de l.Universite´ de Paris* 8, 229–231.

Steward G. W. (1980), The Efficient Generation of Random Orthogonal Matrices with an Application to Condition Estimators, *SIAM Journal of Numerical Analysis* 17, 403–409.