MINISTRY OF SCIENCE AND HIGHER EDUCATION

eFinanse
Financial Internet Quarterly

MODELING STOCK MARKET INDEXES WITH COPULA FUNCTIONS

Jacek Leśkow, Justyna Mokrzycka, Krzysztof Krawiec[1]

Abstract

Contemporary financial risk management is significantly based on the analysis of time series of returns. One of the most significant errors frequently committed by analysts is the predominant use of normal distributions when it is clear that the returns are not normal. Copula models and models for non-normal multivariate distributions provide new tools to solve the problem because the obtained results are immediately applicable in portfolio management, option pricing and measuring risk without assuming normality. Therefore, both a theoretician and a practitioner are interested in multivariate models for returns and copula functions. The copula function models provide an effective and interesting technique of constructing multivariate distribution starting from marginal ones. Due to Sklar's result established in 1959, we can present any multivariate distribution with a help of corresponding marginal distributions and a selected copula function. In this work we present an application of copula function to construct multivariate conditional distributions of times series. In the last part of this paper dynamic models such as DCC-MVGARCH and conditional copula are analyzed. Moreover, we also present an application of bootstrap in the context of copula function. This work is appended by examples showing practical application of our work.

JEL classification: C22, C52
Key words: copula function, GARCH model, conditional copula, DCC-MVGARCH, dynamic conditional copula, bootstrap

Received: 14.03.2011 Accepted: 22.06.2011

Introduction

Copula functions have recently become a very popular tool in financial modeling and risk management as they liberate the analyst from the pitfalls of using the normal distribution in situations where it is evident that returns are not normal. Copula functions are essential as they allow to construct the multivariate model from the marginal distributions and a specific function, called a copula function. The standard approach to portfolio risk management is based on the assumption of multivariate normality since the multivariate normal distribution allows to model the behavior of portfolio via an analysis of individual positions and a covariance function. In some sense, therefore, the copula function approach is a natural generalization of multivariate normal theory. Within this approach we can model the individual positions using a variety of distributions and use the copula function to replace the covariance. This gives far more flexibility to copula functions and allows better modeling of dependences among individual positions of a portfolio. This immediately leads to better modeling of various risks related to portfolio, like liquidity loss and Value-at-Risk (Cherubini, Luciano and Vecchiato, 2003).

[1] Dr hab. Jacek Leśkow, Wyższa Szkoła Biznesu National-Louis University in Nowy Sącz, ul. Zielona 27, 33-300 Nowy Sącz, leskow@wsb-nlu.edu.pl, mgr Justyna Mokrzycka, mgr Krzysztof Krawiec.

To make the technically complex copula theory more comprehensible to the readers, in our paper we will analyze in detail a two-position portfolio containing stock indexes corresponding to 2-dimention copula. We show that our method is significantly better than traditional ones since it is not using normality assumption. We also show how to apply the DCC (dynamic conditional correlation) model and the dynamic conditional copula. Our model is shown to work for a portfolio containing WIG and DAX stock indexes.

General description of copulas

Sklar (1959) has developed the theory of copula functions by formulating a result, where it is possible to represent any multivariate distribution function with the help of its marginal distributions and an appropriately selected copula function. For continuous multivariate distribution functions the choice of copula is unique. To illuminate this fact, let us consider a 2-dimention distribution function $H(x_1, x_2)$ and let $F_1(x_1)$ and $F_2(x_2)$ be marginal distributions corresponding to H. Then, we can select such a copula function C that $\forall\, x_1, x_2 \in R$,

$$H(x_1, x_2) = C(F_1(x_1), F_2(x_2)). \quad (1)$$

Conversely, if C is a 2-dimensional copula function and F_1, F_2 are univariate distributions then we can find a bivariate distribution function H such that F_1, F_2 are marginals of H.

For the definition of the copula function the reader is referred to the monograph by Nelsen (1999). Below, we give examples of multivariate distributions and corresponding copulas.

Example 1. Normal marginals F_1, F_2 with non-normal bivariate distribution function H. Let F_1, F_2 be normal distribution functions and consider the following copula function

$$C(u,v) = u + v - 1 + (1-u)(1-v)e^{-\frac{1}{2}ln(1-u)ln(1-v)}. \quad (2)$$

Using Sklar's theorem we can construct joint distribution function with the following form

$$H(x,y) = F_1(x) + F_2(y) - 1 + (1-F_1(x))(1-F_2(y))e^{-\frac{1}{2}ln(1-F_1(x))ln(1-F_2(y))}. \quad (3)$$

Example 2. Student copula

Student copula, similar to normal copula, is constructed by an inversion method (Nelsen, 1999). Its formula is the following:

$$C_{\rho,\nu}(u_1,u_2) = \int_{-\infty}^{t_\nu^{-1}(u_1)} \int_{-\infty}^{t_\nu^{-1}(u_2)} \frac{\Gamma(\frac{\nu+2}{\nu})}{\Gamma(\frac{\nu}{2})\nu\pi\sqrt{1-\rho^2}} (1 + \frac{r^2 - 2\rho rs + s^2}{\nu(1-\rho^2)})^{-\frac{1}{2}(\nu+2)} drds \quad (4)$$

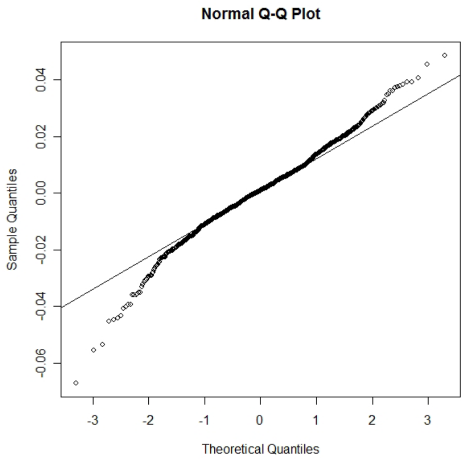
where ρ is a parameter of copula, ν is the number of degrees of freedom and Γ is the Gamma function.

Financial time series models

Precise modeling of marginal distribution is a very important step while using copula functions in modeling behavior of the portfolio. Multivariate distributions are the best tools to model the portfolio and there is a need to extend the classical multivariate normal approach. While analyzing the empirical time series corresponding to returns e.g. for stock market indexes, we frequently encounter such phenomena as volatility clustering, that is the variability in time of the conditional variance. We also observe another phenomenon called fat tails phenomenon when the spread of returns is significantly larger than that corresponding to the normal distribution (see (Mandelbrot, 1963) or Fama (1965)). We also see the so-called financial leverage effect which emphasizes the asymmetric influence of positive and negative information on the whole variance of time series (Tsay, 2002).

It is important to emphasize once more that many financial time series of returns CANNOT be assumed to be normal. To illustrate that, we will use a popular tool called a q-q plot. Such a plot compares the quantiles of the normal distribution with the empirical quantiles corresponding to the analyzed returns. If returns are normal, then the picture should be showing a diagonal line. The picture below is done for the daily returns of the Warsaw Stock Index WIG20 in the period between January 2004 and February 2008.

Figure 1: Comparison of empirical quantiles and normal distribution quantiles for the returns of Warsaw Stock Index WIG20



The picture above shows that the normal distribution is a very bad model for the data. One should rather choose a distribution with fat tails, for example a t-student distribution. Very frequently, the models ARMA-GARCH with fat-tailed innovations are used to model such data (Tsay, 2002).

The key to appropriate and precise modeling of returns from portfolio, stock indexes and stock prices is the volatility modeling. The most popular traditional way to accomplish such task is using the sample variance as a measure of volatility. Again, such measure of volatility is appropriate when the data at hand are normal, which, as shown above, is far from being true e.g. for WIG20. The more precise approach uses the GARCH-type models, created by Engle (Engle, 1982). The advantage of the model is its ability to deal with time-variant conditional heteroskedasticity when the general model for the data has the unconditional time invariant variance. Let us have a closer look at GARCH(p,q) model.

The time series $\{\varepsilon_t\}_{t \in \mathbf{Z}}$ is said to follow the GARCH(p,q) model when

$$\varepsilon_t = \sqrt{h_t}\eta_t, \text{ where } h_t = \alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2 + \sum_{j=1}^{p}\beta_j h_{t-j}. \quad (5)$$

In the formula (5) the symbol η_t denotes the innovations.

The model is well defined when the parameters *p* and *q* are positive integers and $\alpha_0>0$, $\alpha_i >=0$, $i =1,2,\ldots,p$, $\beta_j >=0$, $j=1,2,\ldots,q$. It is important that in the GARCH(p.q) model the conditional variance described by the function h_t (5) depends on previous conditional variances calculated in times *t-1,…,t-q* and from previous values of time series. Model GARCH(p,q) is very important while describing the volatility clustering for residuals $\{e_t\}$ calculated for the ARMA model (Brockwell and Davis, 2002).

The effect of financial leverage is understood as asymmetric influence of good and bad news on the future variance of time series. Of course, good news corresponds to positive returns (profits) and bad news corresponds to negative returns (losses). In the case of stocks, the fall of prices increases the financial leverage and increases the level of risk. Such factors contribute to the increase in volatility, therefore while modeling returns it is very important to take that into consideration. In the literature, there are several ways to handle the above asymmetric influence phenomenon. The most popular is for example GJR-GARCH model presented in Glosten, Jagannathan and Runkleet (1993).

The model GJR-GARCH(p,q) is defined in the following way: the time series $\{\varepsilon_t\}_{t\in\mathbf{Z}}$ follows the GJRGARCH(p,q) model when

$$\varepsilon_t = \sqrt{h_t}\eta_t, \quad (6)$$

$$\text{where} \quad h_t = \alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2 + \sum_{j=1}^{p}\beta_j h_{t-j} + \sum_{i=1}^{q}\gamma_i \mathbf{I}_{\{\varepsilon_{t-i}<0\}}\varepsilon_{t-i}^2,$$

and $\{\eta_t\}$ are arbitrary innovations with mean zero and variance 1. In the model (6) the function $\mathbf{I}_{[a,b]}$ denotes the indicator function of the interval [a, b] that is $\mathbf{I}_{[a,b]}(x)=1$ while $x\in[a,b]$ and $\mathbf{I}_{[a,b]}(x)=0$ while $x\notin[a,b]$. Moreover, the parameters fulfill the conditions: $\alpha_0>0, \alpha_i\geq 0, i=1,..,q, \beta_1\geq 0, j=1,..,p, \gamma_i\geq 0, i=1,..,q.$

A good look at the model (6) entitles us to identify the parameter γ_i with the market asymmetry. When γ_i is bigger than zero then the negative returns (losses) have a stronger influence on volatility than the positive returns (gains). Therefore, parameters γ_i identify the sensitivity of volatility function h_t with respect to negative returns.

Another way to parameterize the lack of symmetry of the market reaction with respect to negative and positive returns is the Asymmetric Power ARCH model introduced by Z. Ding, C. Granger, R. Engle (1993). We give the description of the model below.

The time series $\{\varepsilon_t\}_{t\in\mathbf{Z}}$ is called APARCH(p,q), if for every *t* the following equation holds:

$$\varepsilon_t = \sigma_t\eta_t, \quad (7)$$

$$\text{where} \quad \sigma_t^\delta = \alpha_0 + \sum_{i=1}^{q}\alpha_i(|\varepsilon_{t-i}| - \gamma_i\varepsilon_{t-i})^\delta + \sum_{j=1}^{p}\beta_j\sigma_{t-j}^\delta. \quad (8)$$

In the equation (7) the symbol $\{\eta_t\}$ corresponds to innovations with mean zero and variance one. In order to make the APARCH model well defined the following conditions are needed:

$\alpha_0 > 0, \alpha_i \geq 0, i = 1,..,q, \beta_1 \geq 0, j = 1,..,p, -1 < \gamma_i < 1, i = 1,..,q$. The models like APARCH encompass the ARCH, GARCH, GJR-GARCH models as well as other models used in financial research (Wurtz, Chalabi and Luksan, 2002).

Copula functions in financial modeling

In the classical setup of analysis of returns of portfolio, the multivariate normality of returns is assumed. As we have seen in previous considerations, the assumption of normality is unrealistic. The contemporary approach to portfolio modeling is strongly based on copula functions (Haerdle, 2010). In the sequel to this section, to simplify, we will restrict ourselves to two-dimensional copula functions. Multidimensional generalization of two-dimensional case is an easy generalization and can be found e.g. in Haerdle (2010).

To start fitting the copula functions to time series data described via previously mentioned models (GJR, APARCH and the like) we need to clearly state, that our method of fitting should be applied to conditional copula functions, that is functions that allow to capture the time- dependence of the phenomenon. Then, to fit the copula function we will be using the IFM method of estimating the copula (Cherubini et al, 2003). The details of the method can be represented in the following algorithm:

STEP 1. For each of the positions in the portfolio we identify the appropriate model (e.g. GARCH, GJR etc), estimate the volatility function $\{h_t\}$ of the model and identify the conditional distribution of the $\{\varepsilon_i\}$. Taking, for example, a two-position portfolio, we will have two time series $\{x_t\}_{t\in\mathrm{T}}$ and $\{y_t\}_{t\in\mathrm{T}}$ for which two separate fitting procedures will be made. Each marginal fit identifies the parameters of marginal distribution with the help of AIC criterion. For details, see Mokrzycka (2008).

STEP 2. Using Step 1, for each time moment $t \in \{1,\dots,T\}$ we create independent, uniformly distributed on (0,1) random variables $\{u_t\}_{t\in\mathrm{T}}$ and $\{v_t\}_{t\in\mathrm{T}}$ that are transformations of the original data $\{x_t\}_{t\in\mathrm{T}}$ and $\{y_t\}_{t\in\mathrm{T}}$ via marginal distributions. For details see Cherubini et al (2003).

STEP 3. We are now using the time series $\{u_t\}_{t\in\mathrm{T}}$ and $\{v_t\}_{t\in\mathrm{T}}$ obtained in Step 2 to estimate the parameters of the copula function.

The estimation process described in Step 1 requires a very high precision since it defines the conditional distribution. Moreover, in Step 1, it is quite crucial to appropriately select the copula function. Such general problems are within the scope of ongoing research of the authors. To illuminate the method above, we will present below an example of fitting the bivariate copula corresponding to two stock market indexes: WIG20 and DAX. We have used the open source software R and also algorithms available on the web page of Professor A. Patton (see http://www.nuffield.ox.ac.uk/users/nielsen/mphileconometrics/index2008.htm).

Example 3. (Estimating copula for WIG20 and DAX, 2004-2008)

We are analyzing here the daily returns for bivariate portfolio based on two indexes: WIG20 and DAX observed daily from January 2, 2004 to February 27, 2008. For DAX index, the data set is larger since there are more holidays in Poland than in Germany. Therefore, for the subsequent analysis only working days in both countries are analyzed. We start with estimating the parameters of the marginal distribution for each of the indexes as specified in Step 1. For WIG 20 we have identified GARCH(1,1) model with the innovations coming from the skewed t-distribution. On the other hand, for DAX we have identified APARCH(1,1) with $\delta = 2$ also with skewed t-student distribution. Below, we present the results of the estimation. First, we present the results of Step 1 for DAX WIG20 data.

Table 1: Estimated coefficients of GARCH(1,1) model for WIG20

Parameters	Estimate	Std. Error	t value	Pr(> |t|)
α_0	1.827e-06	1.391e-06	1.314	0.188973
α_1	4.076e-02	1.331e-02	3.061	0.002203 **
β_1	9.505e-01	1.787e-02	53.174	< 2e-16 ***
LLF	3031.989			
AIC	-5.853544			

Table 2: Estimated parameters of APARCH(1,1) with $\delta = 2$ for DAX

Parameters	Estimate	Std. Error	t value	Pr(> |t|)
α_0	4.635e-06	1.515e-06	3.059	0.00222 **
α_1	3.889e-02	1.672e-02	2.326	0.02004 *
γ_1	0.99	3.720e-01	2.688	0.00719 **
β_1	8.703e-01	2.736e-02	31.812	<2e-16 ***
LLF	3400.805			
AIC	-6.566098			

For the DAX index data, the best model - APARCH(1,1) with the parameter $\delta = 2$ has resulted in selecting skewed t-student distribution. Moreover, the parameter γ_1 was close to one which indicated a strong asymmetry effect. In the above tables, the AIC corresponds to Akaike information criterion while identifying either GARCH or APARCH models. The values obtained are the smallest from all considered models. On the other hand, the acronym LLF corresponds to the log likelihood function. The models selected have maximized the LLF value among all considered models.

Now, to show the results of Step 2, we check whether the identification done in Step 1 was performed correctly, that is whether the time series $\{u_t\}_{t\in T}$ and $\{v_t\}_{t\in T}$ are indeed independent in time and uniformly distributed on (0,1).

Table 3: Results of tests

Transformed returns	p-value LM Test	p-value K-S Test
WIG20	0.8982	0.5611
DAX	0.1185	0.3428

For both WIG and DAX we have tested the independence using the LM test (Pesaran, Ullah and Yamagata, 2008) while the classical Kolmogorov-Smirnov test (K-S Test) was used to check the consistency with the uniform distribution. In both cases, p values are bigger than 5% therefore there is no evidence of lack of fit of models used in Step 1. Now we pass to the third Step in our algorithm, that is to identify the copula functions using $\{u_t\}_{t\in T}$ and $\{v_t\}_{t\in T}$.

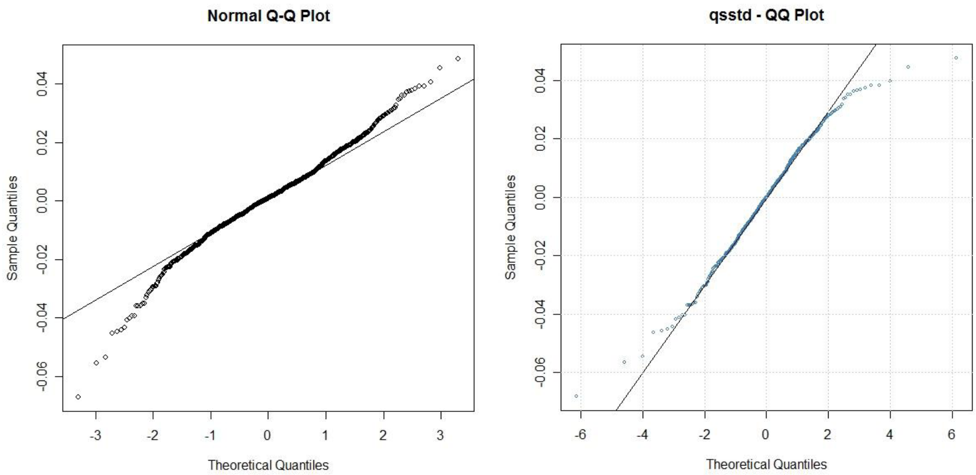
Table 4: Values of LLF and AIC for estimated copula

Copula	LLF	AIC
Normal	120.765	-0.2346
Sym. Joe-Clayton	127.7099	-0.2499
Gumbel	102.3056	-0.1990
t-student	128.6105	-0.2517

The biggest value of log likelihood function LLF and the smallest value of Akaike information criterion AIC is obtained for t-student copula. The estimated number of degrees of freedom for this copula is equal to 8.1969 and estimated correlation coefficient is 0.4583. For details of the parameters of such copula see formula (4).

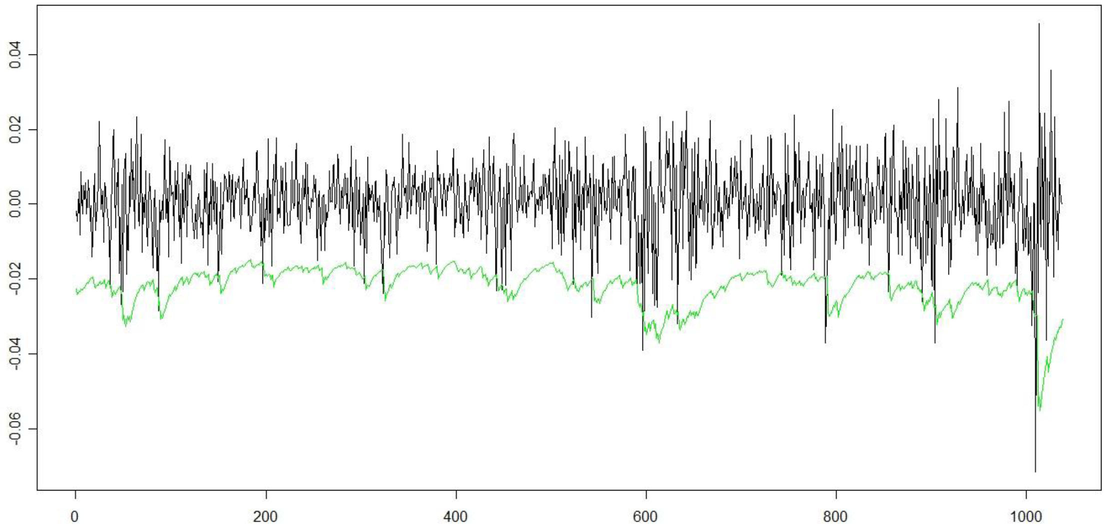
While summarizing this example, we would like to emphasize that for our simple two-position (bivariate) portfolio we have obtained significantly different models for each marginal position. Returns in both cases are highly non-normal and with fat tails.

Figure 2: Comparison of empirical quantiles and normal distribution quantiles for the returns of Warsaw Stock Index WIG20 (on the left side) and for the standarized residuals form GARCH model(on the right side)



Moreover, the copula function is of t-student type, so clearly this portfolio CANNOT be modeled with the multivariate normal distribution. We have clearly shown that copula function approach is far better than the multivariate normal approach. Using our advantageous modeling, we can approximate a very popular risk measure for portfolio, such as Value-at-Risk (VaR). A more formal approach to this problem will be presented in the subsequent research of the Authors, here we will show only graphical illustration copula-based approach. The portfolio selected is based in 50% on WIG20 and in other 50% on DAX in the analyzed period.

Figure 3: Return of portfolio and 95% VaR (green line)



The picture above shows the ability of our models to capture the behavior of Value-at-Risk even for clusters of highly volatile observations which, on Figure 3, occur at the end of the observation interval.

Dynamic models

In the previous section we have shown the advantages of using copula-based model for describing the behavior of portfolio composed of positions generating non-normal returns. The approach previously presented was based on the assumption that the dependence among the positions can be modeled with the one copula function that is not changing in time. In reality, however, we have dynamic changes between market positions – for example correlations can be time-dependent. To have a clearer insight into that phenomenon, let us consider DCC-MVGARCH (dynamical conditional correlation - MVGARCH) model and a dynamical conditional copula model. See Engle et Shephard (2001) for the former model and Patton (2001) for the latter.

Suppose that we have a k-dimensional vector r_t of returns in time moments $t = 1,...,T$. We can represent this vector via the vector autoregression (VAR) model as:

$$A(L)r_t = \varepsilon_t \text{, where } \varepsilon_t|F_{t-1} \sim N_k(0,H_t), \tag{9}$$

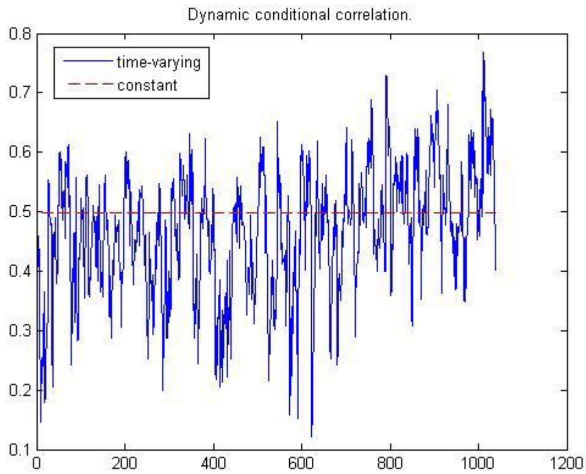
A(L) corresponds to the matrix of the lag operator L ($Lr_{1,t} = r_{1;t-1}$), ε_t corresponds to the errors in model VAR with the covariance matrix $H_t = \{h_{ij}\}_t, i, j = 1,2,...,k$ corresponding to the conditional distribution $N_k(0;H_t)$. Using the approach of Engle and Sheppard (2001) we can have the following representation:

$$H_t = D_t R_t D_t,$$

where D_t is the $k \ x \ k$ diagonal matrix containing the standard deviations of the univariate conditional distributions of the corresponding GARCH models and also R_t is a time-variant correlation matrix. Such representation of the covariance matrix significantly simplifies the process of estimating parameters. The estimation procedure is done using the maximum likelihood method.

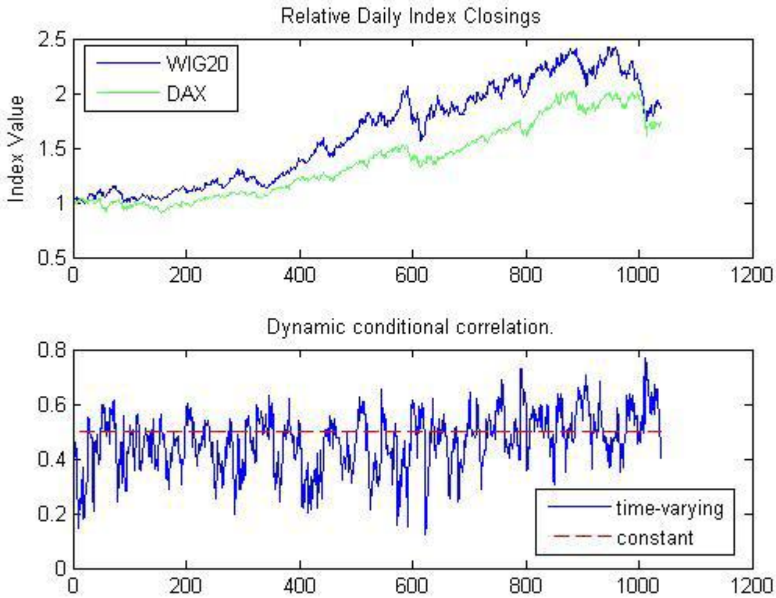
Example 4. Estimating DCC-MVGARCH model for the DAX and WIG20 time series
We are analyzing the WIG20 and DAX data previously considered. We are using the DCC-MVGARCH model directly to the returns. We start with the test of constancy of correlation described in Engle (2001). According to this test, we have rejected the model with the fixed correlation in favor of the model with the time-variant correlation with lag 4. For the consecutive lags (from order 1 to order 4) the p-values were equal to *pvalue1= 0.9208, pvalue2=0.9504, pvalue3=0.2994, pvalue4=0.0047*. For the sake of comparison, the fixed estimated value of the correlation coefficient was equal to 0.4985. We have the following figure to substantiate our time-varying correlation model:

Figure 4: Estimated correlation in DCC(1,1) model



The value of the log likelihood function with the time varying correlation is equal to LLF = 6496.8. It estimates a multivariate GARCH model using Bollerslev's constant correlation estimator LLF = 6489.6 (see Bollerslev (1986)). Let us now compare the dynamics of the time varying correlation in our DCC-MVGARCH model and the dynamics of returns. The picture above represents rescaled indexes, that is the rescaled time series $\{x_t^R\}_{t\in T}$ and $\{y_t^R\}_{t\in T}$, where $x_t^R = \frac{x_t}{x_1}$ and $y_t^R = \frac{y_t}{y_1}$, where x_t denotes the WIG index.

Figure 5: Rescaled indexes and time varying correlation



Observation of the joint dynamics of the time varying correlation and the relative movement of $\{x_t^R\}_{t\in T}$ and $\{y_t^R\}_{t\in T}$ leads to several conclusions. First, the simultaneous drops of indexes on both markets (around the time point 1000) generate a high value of the correlation. When there is a growing tendency on both markets, the correlation is not very high. The falling tendency generates higher values of correlation. This is another example of the lack of symmetry in simultaneous movements of the markets. Again, bad news is spreading more rapidly...

Dynamic conditional copula model

The previous example showing lack of symmetry in dependence between two market indexes motivates us to seek better models of asymmetric movements of the markets. We will start our search for better models by considering a normal copula with the time dependent correlation coefficient with some specific form of time dependence. Then we will move to a different concept of correlation - that is to tail dependence index and the Joy-Clayton copula.
Let us consider first the normal copula. It has the following form

$$C_\rho(u,v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} exp\{\frac{-(r^2-2\rho rs+s^2)}{2(1-\rho^2)}\}drds, -1<\rho<1. \qquad (10)$$

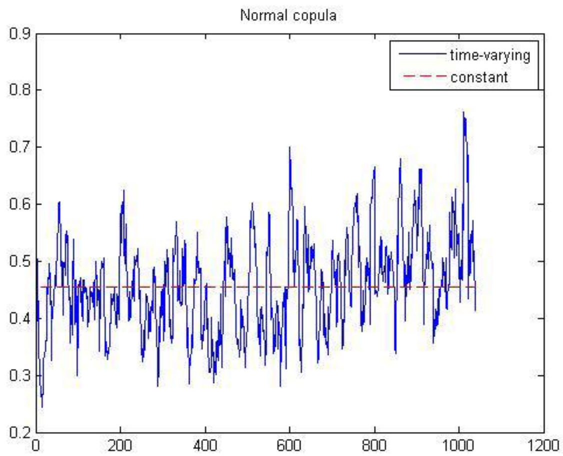
The research of Patton (2006) provides a following representation for the time-dependent correlation:

$$\rho_t = \Lambda(\omega_\rho + \beta_\rho \rho_{t-1} + \alpha_\rho \frac{1}{10}\sum_{j=1}^{10} \phi^{-1}(u_{t-j})\phi^{-1}(v_{t-j})$$

where $\Lambda(x)=\frac{1-e^{-x}}{1+e^{-x}}$ is the logistic transformation to keep the correlation coefficient in the interval (-1,1).We will now run the estimation procedure using the above model for our two index portfolio.

Example 5. Dynamic normal copula model
The data, as before, concern the mutual dynamics of WIG20 and DAX model. The fit provides a dynamic correlation coefficient very similar to previously obtained dynamic coefficient in the model DCC-MVGARCH.

Figure 6: Estimated dynamic correlation in the dynamic copula model



The value of the log-likelihood function in this case is equal to 128.2976 and is bigger than in all the previously obtained models that have used a copula approach.
We will now pass to the Joy-Clayton copula model, where the copula function is defined as:

$$C_{JC}(u,v\,|\,\tau_U,\tau_L)=1-(1-\{[1-(1-u)^{\kappa}]^{-\gamma}+[1-(1-v)^{\kappa}]^{-\gamma}-1\}^{-1/\gamma})^{1/\kappa}, \qquad (11)$$

where $\kappa=1/\log_2(1-\tau^U)$, $\gamma=-1/\log_2(\tau^L)$, $\tau^U\in(0,1)$, $\tau^L\in(0,1)$.

It is important to emphasize that the parameters of the copula, τ^L and τ^U express the tail dependence of the distributions. To make it more technically clear, we give the definitions of τ^L and τ^U below (Patton, 2006). For random variable X_1 and X_2 with continuous distributions F_1 and F_2 respectively the upper tail dependence index τ^U is defined as

$$\tau^U=\lim_{u\to 1}Pr[X_2>F_2^{-1}(u)\,|\,X_1>F_1^{-1}(u)]=\lim_{u\to 1}Pr[X_1>F_1^{-1}(u)\,|\,X_2>F_2^{-1}(u)]. \qquad (12)$$

On the other hand, the lower tail dependence index τ^L is defined as

$$\tau^L=\lim_{u\to 0}Pr[X_2\le F_2^{-1}(u)\,|\,X_1\le F_1^{-1}(u)]=\lim_{u\to 0}Pr[X_1\le F_1^{-1}(u)\,|\,X_2\le F_2^{-1}(u)]. \qquad (13)$$

The tail dependence indexes reflect the interactions of two random variables for extreme events. It is, therefore, quite natural to introduce such parameters while studying mutual dynamics of two or more positions of the portfolio. In our case one may wonder how to parameterize the mutual dynamics of DAX and WIG20 market indexes. The considerations below will shed some light on that problem.
We will now introduce the symmetric Joy-Clayton copula. Define

$$C_{SJC}(u,v\,|\,\tau^U,\tau^L)=0.5(C_{JC}(u,v\,|\,\tau^U,\tau^L)+C_{JC}(1-u,1-v\,|\,\tau^U,\tau^L)+u+v-1). \quad (14)$$

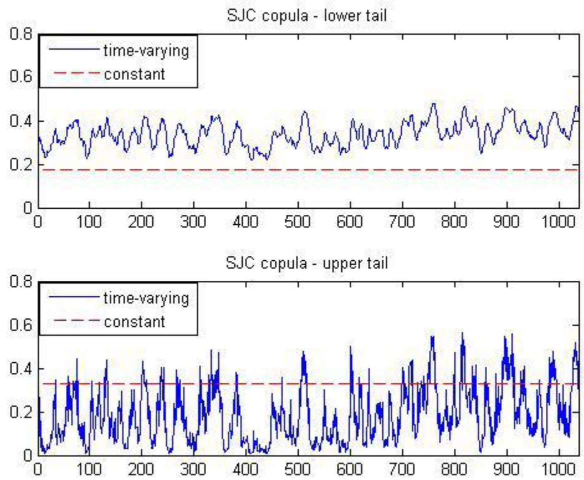
Observe that when $\tau^U=\tau^L$ then the copula is symmetric. We can parameterize the tail dependence by the following equations

$$\tau_t^U=\Lambda(\omega_U+\beta_U\tau_{t-1}^U+\alpha_U\frac{1}{p}\sum_{j=1}^{p}|u_{t-j}-v_{t-j}|),\ \tau_t^L=\Lambda(\omega_L+\beta_L\tau_{t-1}^U+\alpha_L\frac{1}{p}\sum_{j=1}^{p}|u_{t-j}-v_{t-j}|),$$

where $\Lambda(x)\equiv(1+e^{-x})^{-1}$ is the logistic transformation to keep the correlation coefficient within (-1,1) interval. In applications, the integer p is moderate and we assume it is equal 10 (Patton, 2006). We will now fit the symmetric Joy-Clayton dynamic copula to WIG20 and DAX data.

Example 7. Fitting symmetric dynamic Joy-Clayton copula model
We will now show the process of estimation of dynamic Joy-Clayton copula (SJC copula) for the returns based on indexes WIG20 and DAX. The range of data is as before. After the estimation process the value of the log likelihood function is equal LLFSJC = 136.7780. Below, we present a picture showing the time varying parameters.

Figure 7: Time varying parameters of SJC copula



On the picture above the dotted line represents the constant value of the estimated parameters in the nondynamic approach. The likelihood function is maximized for such copula function, since the dynamic symmetric Joy-Clayton copula is the best while modeling the dependence between returns for WIG20 and DAX market indexes.

Bootstrap confidence intervals for VaR using copula functions

In this section we apply the aforementioned technique of copula functions to get better idea on calculating Value-at-Risk for portfolio. We will be showing a method together with the example. Additionally, a bootstrap method will be used to generate confidence intervals for Value-at-Risk. In such a way, we will be getting some results concerning the accuracy of our method.
To start, consider a portfolio composed of four assets coming from the Polish stock market. We take TP SA - the Polish Telecom, PKO BP - popular Polish Bank, the fuel conglomerate -

PKN ORLEN and , finally, the media company TVN. Calculation of Value-at-Risk for a profit and loss function L was conducted for logarithmic returns modeled with the GARCH(1,1) model. Subsequently, the residuals from the GARCH(1,1) model were described with the t-Student distribution and the joint distribution of four-dimensional residuals was modeled with the copula C_θ. This copula function C_θ had parameters estimated via the IFM method described in previous sections. To this fit with the copula function C_θ we have applied our dynamic conditional correlation model which has shown that there is essentially no dynamic component for correlation. The table below provides the empirical evidence for that fact.

Table 5: p-value of test for presence of dynamic correlation

Lag	p-value
1	0.1095
2	0.1796
3	0.3006
4	0.2589
5	0.3346
6	0.3949
7	0.3367

The data used to estimate parameters of the chosen t-Student copula correspond to data in the period January 2007 to June 2010. Once copula function was estimated for such data, the Monte Carlo simulation was applied to get the forecast from the profit and loss function L. This, finally, was used to get the Value-at-Risk as a α-level percentile from the predicted profit and loss function L.

The accuracy of this method was checked with the application of the nonparametric bootstrap method for the residuals (see e.g. Lahiri, 2003, page 200). The input for the bootstrap procedure was provided by the residuals ($\varepsilon_1,\ldots,\varepsilon_n$) generated from the joint distribution modeled by the copula function C_θ. We were verifying the independence of the residuals with the usual ACF/PACF diagnostics. In such a way, using the bootstrapped residuals ($\varepsilon_1^P,\ldots,\varepsilon_n^P$) we were able to get the bootstrapped versions of the returns using the formula:

$$r_t^b = \mu + \sigma_t \varepsilon_t^b \tag{15}$$

The empirical quantile obtained for the profit and loss function for the bootstrap data was considered an estimate of the original empirical quantile. Therefore, the confidence interval for $\widehat{VaR(\alpha)}$ at the level $1-\beta$ is defined via $(u_{\frac{\beta}{2}}, u_{1-\frac{\beta}{2}})$, where $u_{\frac{\beta}{2}}$ and $u_{1-\frac{\beta}{2}}$ are empirical quantiles from the estimates $\widehat{VaR(\alpha)}^b$, of the order $\frac{\beta}{2}$ and $1-\frac{\beta}{2}$, respectively, obtained from (15).

Figure 8: 95 percent confidence intervals for 5 percent VaR

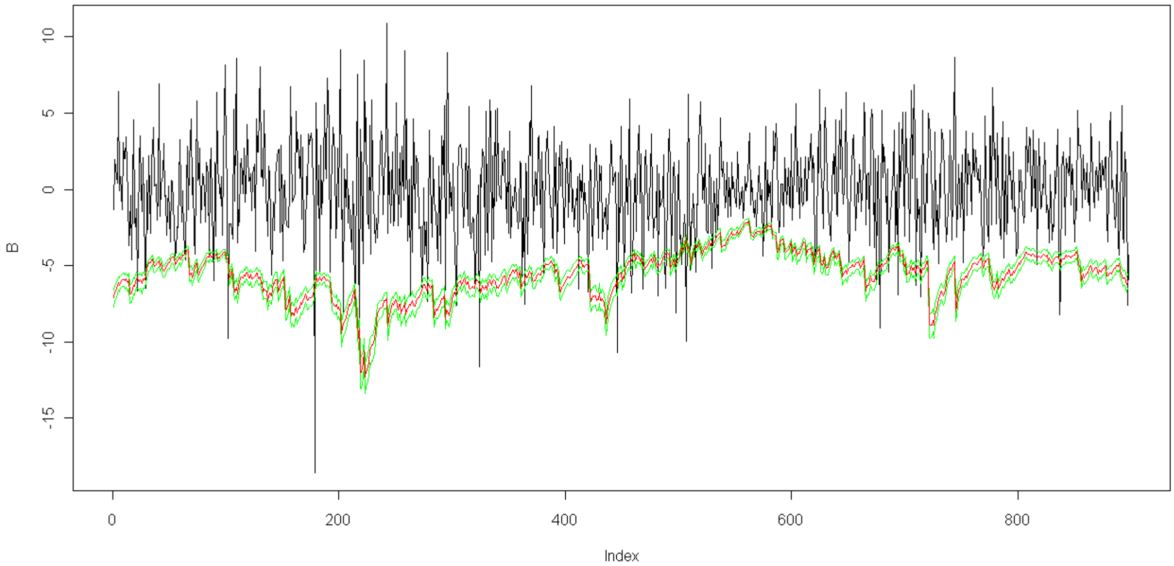
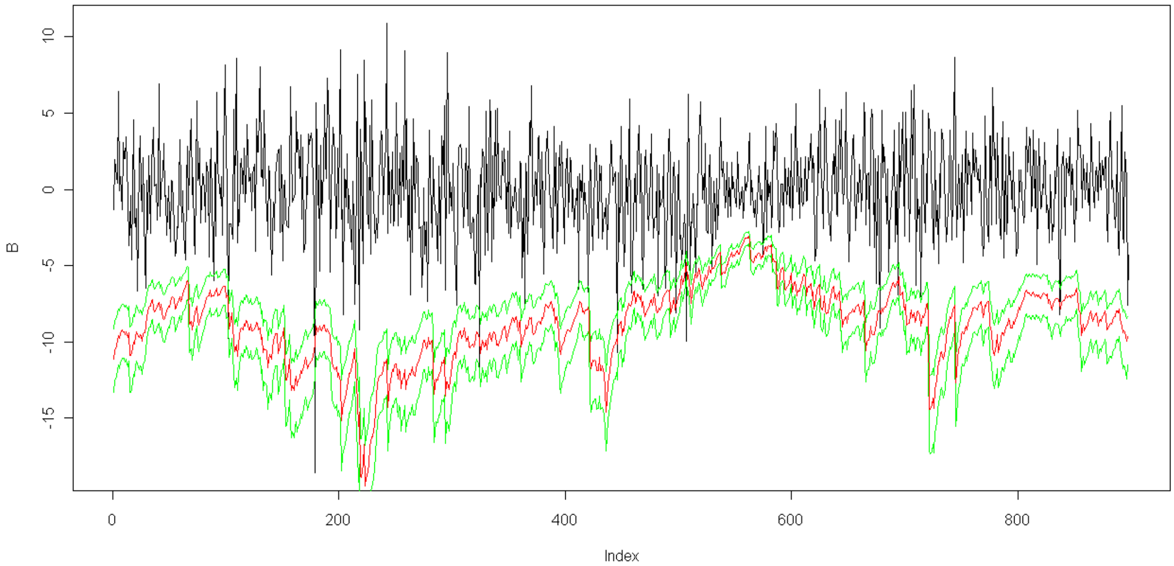


Figure 9: 99 percent confidence intervals for 1 percent VaR



Figures 8 and 9 represent the ability of the bootstrap method to construct the confidence interval around the Value-at-Risk. Observe that 99% confidence intervals for VaR are wider than those for 95% VaR.

Conclusions

In our paper we were considering the applications of the copula function theory in modeling returns. As a result, we presented an approach that is far more accurate and flexible than that based on the normal distribution theory. Moreover, using the copula functions we can more flexibly model the dependence between various elements in the portfolio. It is now increasingly evident that serious and contemporary returns analysis has to be based on the copula theory.

In particular, we would like to stress the importance of the dynamic copula approach to assets modeling. We are putting together some important statistics relevant to that problem in the table below.

Table 6: Comparison of the log likelihood function

	t-stud.	DCC(1,1)	dyn. SJC	dyn. normal
LLF_{WIG20}	3031.989		3031.989	3031.989
LLF_{DAX}	3400.805		3400.805	3400.805
LLF_{Copula}	128.6105		136.7780	128.2976
LLF	6561.4045	6496.8	6569.572	6561.0916

According to the analysis of the likelihood function above, the dynamic SJC model is the best and has a significant advantage over the DCC(1,1) model.

We believe that the copula theory and its applications will provide an important benefit to practitioners. As an example, at the end of the paper, we provided a copula-based calculation of Value-at-Risk with the use of bootstrap method to get better control over the error rate of the calculation. We strongly recommend using copula-based techniques to all practitioners working in the area of predicting Value-at-Risk since these methods are shown here to be highly superior than techniques based on multivariate normal distribution.

References

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity, *Journal of Econometrics 31*, pp. 307-327.

Brockwell, P. J., Davis, R. A. (2002). *Introduction to Time Series and Forecasting*, Springer-Verlag, New York.

Cherubini, U., Luciano, E., Vecchiato W. (2003). *Copula Methods in Finance*, Wiley Series in Finance.

Ding, Z., Granger C. W. J., Engle R. F. (1993). A Long Memory Property of Stock Market Returns and a New Model, *Journal of Empirical Finance*, pp. 83-106.

Engle, R. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Ination, *Econometrica 50* (4), pp. 987-1007.

Engle, R., Sheppard, R. (2001). Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH, University of California at San Diego, *Economics Working Paper Series*.

Fama, E. G. (1965). The Behavior of Stock-Market Prices, *The Journal of Business*, Vol. 38, Nr 1. (Jan., 1965), pp. 34-105.

van den Goorbergh, R. W. J. (2004). A Copula-Based Autoregressive Conditional Dependence Model of International Stock Markets, http://ideas.repec.org/p/dnb/dnbwpp/022.html.

Glosten, L., Jagannathan, R., Runkle, D. (1993). On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *Journal of Finance 48*, pp. 1179-1801.

Haerdle, W., Durante, F., Jaworski, P. and Rychlik, T. (2010). *Copula Theory and Its Applications*, Lecture Notes in Statistics, Springer.

Lahiri S. N. (2003). *Resampling Methods for Dependent Data*, Springer Series in Statistics.

Leśkow, J., Napolitano, A. (2002). Quantile Prediction for Time Series in the Fraction-of-time Probability Framework, *Signal Processing 82*, pp. 1727-1741.

Mandelbrot, B. (1963). The Variation of Certain Speculative Prices, *The Journal of Business*, Vol. 36, Nr 4, pp. 394-419.

Mokrzycka, J. (2008). Applications of Copula Functions to Analysis of Characteristics of Time Series, MSc theses (in Polish), Academy of Mining and Metallurgy, Cracow.

Nelsen R. B. (1999). *An Introduction to Copulas*, Springer-Verlag.

Nelson, D. B., (1991). Conditional Heteroscedasticity in Assets Returns: A New Approach, *Econometrica 59* (2), pp. 347-370.

Patton, A. J., (2001). Modeling Time-Varying Exchange Rate Dependence Using the Conditional Copula, *Discussion Paper 2001-09*, University of California, San Diego.

Patton, A. J., (2006). Modeling Asymmetric Exchange Rate Dependence, *International Economic Review 47* (2).

Pesaran, M. H., Ullah, A., Yamagata, T. (2008). A Bias-adjusted LM Test of Error Cross-section Independence, *Econometric Journal 11* (1), pp. 105-127.

Sklar, A. (1959). Fonctions de Répartition à n Dimensions et Leurs Marges, *Publications de l'Institut de Statistique de L'Université de Paris* 8, pp. 229–231.

Tsay, R.,(2002). *Analysis of Financial Time Series*, Wiley and Sons, Chicago.

Wurtz, D., Chalabi, Y., Luksan, L. (2002). Parameter Estimation of ARMA Models with GARCH/ARCH Errors. An R and SPlus Software Implementation, *Journal of Statistical Software*.