

## Using Conditional Copula to Estimate Value at Risk

Helder Parra Palaro and Luiz Koodi Hotta  
*State University of Campinas*

*Abstract:* Value at Risk (VaR) plays a central role in risk management. There are several approaches for the estimation of VaR, such as historical simulation, the variance-covariance (also known as analytical), and the Monte Carlo approaches. Whereas the first approach does not assume any distribution, the last two approaches demand the joint distribution to be known, which in the analytical approach is frequently the normal distribution. The copula theory is a fundamental tool in modeling multivariate distributions. It allows the definition of the joint distribution through the marginal distributions and the dependence between the variables. Recently the copula theory has been extended to the conditional case, allowing the use of copulae to model dynamical structures. Time variation in the first and second conditional moments is widely discussed in the literature, so allowing the time variation in the conditional dependence seems to be natural. This work presents some concepts and properties of copula functions and an application of the copula theory in the estimation of VaR of a portfolio composed by Nasdaq and S&P500 stock indices.

*Key words:* Copula, multivariate distribution function, value-at-risk.

### 1. Introduction

Value at Risk (VaR) is probably the most popular risk measure, having a central role in risk management. Although VaR is a simple measure, it is not easily estimated. There are several approaches for the estimation of VaR, such as the variance-covariance (also known as analytical), the historical simulation and the Monte Carlo approaches. The analytical approach has been largely used after the publishing of the Riskmetrics methodology. This approach adopts the assumption of multivariate normality of the joint distribution of the assets returns. In this case, the covariance matrix is a natural measure of dependence between the assets and the variance is a good measure of risk. In finance the normality is rarely an adequate assumption. For instance, Longin and Solnick (2001) and Ang and Chen (2002) found evidence that asset returns are more highly correlated during volatile markets and during market downturns. The deviation from

normality could lead to an inadequate VaR estimate. In this case, the portfolio could be either riskier than desired or the financial institution unnecessarily conservative.

The theory of copulae is a very powerful tool for modeling joint distributions because it does not require the assumption of joint normality and allow us to decompose any  $n$ -dimensional joint distribution into its  $n$  marginal distributions and a copula function. Conversely, a copula produces a multivariate joint distribution combining the marginal distributions and the dependence between the variables. Copulae have been broadly used in the statistical literature. The books of Joe (1997) and Nelsen (1999) presented a good introduction to the copula theory. Although copulae have been only recently used in the financial area, there are already several applications in this area. The papers of Bouyé *et al.* (2000), Embrechts, McNeil and Straumann (2002) and Embrechts, Lindskog and McNeil (2003) provided general examples of applications of copulae in finance. There also several particular applications. For instance, Cherubini and Luciano (2001) estimated the VaR using the Archimedean copula family and the historical empirical distribution in the estimation of marginal distributions; Rockinger and Jondeau (2001) used the Plackett copula with GARCH process with innovations modeled by the Student-t asymmetrical generalized distribution of Hansen (1994) and proposed a new measure of conditional dependence; Georges (2001) used the normal copula to model options time of exercise and for derivative pricing; Meneguzzo and Vecchiato (2002) used copulae for modeling the risk of credit derivatives; and Fortin and Kuzmics (2002) used convex linear combinations of copulae for estimating the VaR of a portfolio composed by the FSTE and DAX stock indices; and Embrechts, McNeil and Straumann (2002) and Embrechts, Hoving and Juri (2003) used copulae to model extreme value and risk limits. These recently published papers show the wide range of copula applications in finance.

The recent extension of the unconditional copula theory to the conditional case has been used by Patton (2003a) to model time-varying conditional dependence. Time variation in the first and second conditional moments is widely discussed in the statistical literature, so allowing the temporal variation in the conditional dependence in time series seems to be natural.

In this paper we discuss the application of conditional copula in estimating the VaR of a portfolio with two assets. The paper is organized as follow. Section 2 defines copula and presents Sklar's theorem. The copulae families used in this work are presented in Section 3 while Section 4 discusses some inference methods for copulae, like estimation and model selection. Finally, Section 5 applies the method to a portfolio composed by two assets (the Nasdaq and the S&P500 stock indices). After modeling and estimating the parameters by the Inference Function for Margins method, we used out-of-sample simulation techniques to

test the accuracy of the VaR estimates. We compared the results obtained using traditional approaches to estimate VaR, like the Exponentially Weighted Moving Averages, the historical simulation, and univariate and bivariate GARCH models.

## 2. Introduction

According to Nelsen (1999, pg.1), copulae can be seen from two points of view: “From one point of view, copulas are functions that join or couple multivariate distribution functions to their one-dimensional marginal distribution functions. Alternatively, copulas are multivariate distribution functions whose one-dimensional margins are uniform on the interval  $(0,1)$ ”.

### 2.1 Definition

In this section we give the general definition of copulae and a equivalent definition for the random variable context.

For simplicity purposes, throughout this paper we treat only the bivariate case, but the extension for higher dimensions is straightforward (see, for instance, Nelsen (1999) and Patton (2003a)).

**Definition 2.1** A 2-dimensional copula function (or briefly a copula) is a function  $C$ , whose domains is  $[0, 1]^2$  and whose range is  $[0, 1]$  with the following properties:

1.  $C(\mathbf{x}) = 0$  for all  $\mathbf{x} \in [0, 1]^2$  when at least one element of  $\mathbf{x}$  is 0;
2.  $C(x_1, 1) = C(1, x_2) = 1$  for all  $(x_1, x_2) \in [0, 1]^2$ ;
3. for all  $(a_1, a_2), (b_1, b_2) \in [0, 1]^2$  with  $a_1 \leq b_1$  and  $a_2 \leq b_2$ , we have :

$$V_C([a, b]) = C(a_2, b_2) - C(a_1, b_2) - C(a_2, b_1) + C(a_1, b_1) \geq 0$$

Hence any bivariate distribution function whose margins are standard uniform distributions is a copula.

**Definition 2.2** The copula function  $C$  is a copula for the random vector  $\mathbf{X} = (X_1, X_2)^t$ , if it is the joint distribution function of the random vector  $\mathbf{U} = (U_1, U_2)^t$  where  $U_i = F_i(X_i)$ , and  $F_i$  are the marginal distribution functions of  $X_i$ ,  $i = 1, 2$ .

This imply that:

$$H(x_1, x_2) = C(F_1(x_1), F_2(x_2)),$$

where  $H$  is the joint distribution function of  $(X_1, X_2)$ . If  $F_1$  and  $F_2$  are continuous then the copula  $C$  is unique. Thus we can interpret a copula as a function which links the marginal distributions of a random vector to its joint distribution.

## 2.2 Sklar theorem

The next theorem is a key result in the theory of copulae :

**Theorem 2.1** (Sklar Theorem) Let  $H$  be a 2-dimensional joint distribution function with marginal distributions  $F_1, F_2$ . Then there exists a copula  $C$  such that for all  $\mathbf{x} \in \overline{\mathbb{R}}^n$ ,

$$H(x_1, x_2) = C(F_1(x_1), F_2(x_2)). \quad (2.1)$$

If  $F_1, F_2$  are continuous, then  $C$  is unique; otherwise,  $C$  is uniquely determined on  $Ran(F_1) \times Ran(F_2)$ . Conversely, if  $C$  is a copula and  $F_1, F_2$  are distribution functions, then the function  $H$  defined by (2.1) is a joint distribution function with margins  $F_1$  and  $F_2$ .

The proof can be found in Nelsen (1999, p. 18).

It is the converse of the Sklar's theorem that is most interesting for modeling multivariate distributions in finance. It implies that we may link any group of  $n$  univariate distributions, of any type (not necessarily from the same family), with any copula and we will have defined a valid multivariate distribution. The usefulness of this result stems from the fact that while in economics and statistics literature we have a large set of flexible parametric univariate distributions available, the set of parametric multivariate distributions available is much smaller. According to Patton (2003a), referring to finance data, "Decomposing the multivariate distribution into the marginal distributions and the copula allows for the construction of better models of the individual variables than would be possible if we constrained ourselves to look only at existing multivariate distributions".

Patton (2003a) extended and proved the validity of the Sklar's theorem for the conditional case. In this extension the conditioning variable(s)  $\mathbf{W}$  must be the same for the marginal distributions and the copula.

For the definition of copulae for the general case ( $n > 2$ ) and for more details, look at Nelsen (1999), Bouyé *et al.* (2000) and Patton (2003a).

## 3. Some Families of Copulae

In this section we present the three families of copulae used in this work : Student- $t$ , Plackett, and symmetrized of Joe-Clayton copulae.

### 3.1 Student- $t$ copula

The bivariate Student- $t$  copula (or briefly  $t$  copula) is the function

$$C_{R_{12}, \nu}^t(u, v) = \int_{-\infty}^{t_{\nu}^{-1}(u)} \int_{-\infty}^{t_{\nu}^{-1}(v)} \frac{1}{2\pi(1-R_{12}^2)^{1/2}} \left\{ 1 + \frac{s^2 - 2R_{12}st + t^2}{\nu(1-R_{12}^2)} \right\}^{-(\nu+2)/2} ds dt.$$

where  $t_{\nu}^{-1}$  is the inverse of the univariate  $t$  distribution with  $\nu$  degrees of freedom.

If the marginal distributions  $F_1$  and  $F_2$  are two Student- $t$  distributions with (same)  $\nu$  degrees of freedom and  $C$  is a Student- $t$  copula with parameters  $\nu$  and  $R_{12}$ , then the bivariate distribution function  $H$  defined by  $H(x, y) = C(F_1(x), F_2(y))$  is the standardized bivariate  $t$  distribution, with  $\mu = 0$ , linear correlation coefficient  $R_{12}$  and  $\nu$  degrees of freedom. In this case the  $t$  copula is the copula function which join the marginal  $t$  distributions with same degrees of freedom to the bivariate  $t$  distribution. The  $t$  Student copula generalizes the bivariate  $t$  distribution because we can adopt any marginal distribution.

### 3.2 Plackett's copula

A measure of "association" or "dependence" in  $2 \times 2$  contingency tables is the cross product ratio, or odds ratio, which we will denote by  $\theta$ . For convenience, we have labelled the categories for each variable as 'low' and 'high'.

Table 1: A  $2 \times 2$  contingency table.

		Column variable		
		Low	High	
Row Variable	Low	$a$	$b$	$a + b$
	High	$c$	$d$	$c + d$
		$a + c$	$b + d$	$n$

If the observed count in the four categories are  $a, b, c, d$ , as shown in Table 1, then the cross product ratio is the positive real number  $\theta$  given by  $\theta = (ad)/(bc)$ . The value  $\theta = 1$  corresponds to independence, which implies that each 'observed' entry (for example  $a$ ) is equal to its 'expected value' under independence (here  $(a+b)(a+c)/n$ ), where  $n = a + b + c + d$ . When  $\theta > 1$ , the observations are more concentrated in the 'low-low' and 'high-high' cells; and when  $0 < \theta < 1$ , the observations are more concentrated in the 'low-high' and 'high-low' cells.

Plackett (1965) associated continuous marginal distributions to this table. Let  $X$  and  $Y$  be continuous random variables with joint distribution  $H$ , and margins

$F_1$  and  $F_2$ , respectively. Let  $(x, y)$  be any pair of real numbers and let the ‘low’ and ‘high’ categories for the column variable correspond to the events “ $X \leq x$ ” and “ $X > x$ ” respectively, for the variable  $X$ , “ $Y \leq y$ ” and “ $Y > y$ ” respectively for the variable  $Y$ . Then replacing the numbers  $a, b, c, d$  by their probabilities of occurrence, we have

$$\theta = \frac{H(x, y)[1 - F_1(x) - F_2(y) + H(x, y)]}{[F_1(x) - H(x, y)][F_2(x) - H(x, y)]}.$$

For most joint distributions,  $\theta$  will be a function of the point  $(x, y)$ . But there are joint continuous distributions which do not depend on  $(x, y)$ . Let  $u = F_1(x)$  and  $v = F_2(y)$ . Using Sklar’s Theorem 2.1, we have

$$\theta = \frac{C(u, v)[1 - u - v + C(u, v)]}{[u - C(u, v)][v - C(u, v)]},$$

where  $C$  is the copula of  $X$  and  $Y$ .

Solving for  $C$ , we have

$$C_\theta(u, v) = \frac{1}{2(\theta - 1)} [1 + (\theta - 1)(u + v) - ([1 + (\theta - 1)(u + v)]^2 - 4uv\theta(\theta - 1))^{1/2}]$$

for  $\theta \neq 1$ , and  $C_\theta(u, v) = uv$  for  $\theta = 1$ , which is defined for  $\theta > 0$ , and satisfies the copula conditions. Then  $H(x, y) = C_\theta(F(x), G(y))$  is the joint distribution function of  $X$  and  $Y$  and the function  $C_\theta$  is called Plackett’s copula.

Rockinger and Jondeau (2001) used the Plackett’s copula and a dependence measure to check whether the linear dependence varies with the time. They worked with returns of European stock market series, the S&P500 index and the Nikkei index. One disadvantage of the Plackett’s copula is that it cannot be easily extended for dimensions larger than two.

### 3.3 Symmetrized Joe-Clayton copula

Patton (2003a) used a modified form of Joe-Clayton copula to model the returns of Yen-US Dollar and Mark - US dollar exchange rates. Although this copula does not have the restriction of symmetrical dependence it includes the symmetric case as a particular case. The Joe-Clayton copula is given by:

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

where

$$\begin{aligned}\kappa &= 1/\log_2(2 - \tau_U) \\ \gamma &= -1/\log_2(\tau_L) \\ \tau_U &\in (0, 1), \tau_L \in (0, 1).\end{aligned}$$

This copula has two parameters,  $\tau_U$  and  $\tau_L$ , which are the coefficients of upper and low tail dependence, respectively (see Patton (2003a)). The Joe-Clayton copula still has a slight asymmetry when  $\tau_U = \tau_L$ , which is not convenient. In order to overcome this problem we have a modified form of the copula, known as symmetrized Joe-Clayton copula (SJC) which is given by:

$$\begin{aligned}C_{SJC}(u, v|\tau_U, \tau_L) &= \\ &0.5 C_{JC}(u, v|\tau_U, \tau_L) + 0.5 C_{JC}(1 - u, 1 - v|\tau_L, \tau_U) + u + v - 1,\end{aligned}$$

which is symmetric when  $\tau_U = \tau_L$ .

#### 4. Statistical Inference of Copulae

Let  $(X_1, X_2)$ , be a vector of two random variables with joint distribution function  $H$  and marginal distribution functions  $F_1$  and  $F_2$  respectively. Each marginal distribution function depends only on the parameter  $\vartheta_i$ . Denote the unknown vector of parameters by  $\vartheta = (\vartheta_1, \vartheta_2, \theta)$ , where  $\theta$  is the vector of parameters of the n-dimensional copula  $\{C_\theta, \theta \in \Theta\}$  and  $C_\theta$  is completely known except for the parameter  $\theta$ . Suppose that  $\{(x_{1,t}, x_{2,t})\}_{t=1}^T$  is a sample of size T. Hence we have by Sklar's theorem:

$$H(x_1, x_2) = C(F_1(x_1; \vartheta_1), F_2(x_2; \vartheta_2); \theta). \quad (4.1)$$

Thus the joint distribution function  $H$  is completely specified by the parameter vector  $\vartheta = (\vartheta_1, \vartheta_2, \theta)$ .

Differentiating (4.1) with respect to all variables, we obtain the density function  $h$

$$h(x_1, x_2) = c(F_1(x_1), F_2(x_2))f_1(x_1)f_2(x_2), \quad (4.2)$$

where  $f_i$  is the density function associated to the marginal distribution  $F_i$  and  $c$  is the copula density, given by

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

#### 4.1 Estimation

The log-likelihood function is given by:

$$l(\vartheta) = \sum_{t=1}^T \ln c(F_1(x_1, t; \vartheta_1), F_2(x_2, t; \vartheta_2); \theta) + \sum_{t=1}^T \sum_{i=1}^2 \ln f_i(x_i, t; \vartheta_i). \quad (4.3)$$

Thus, the maximum likelihood estimate  $\hat{\vartheta}$  maximizes (4.3), i.e., it is given by:

$$\hat{\vartheta} = \arg \max_{\vartheta} l(\vartheta).$$

This method can be computationally cumbersome when the dimensional is large because the marginal and copula parameters must be estimated jointly. For this reason it is not usually used in practice. One alternative is given by the Inference Function for Margins (IFM) Method, where the parameters are estimated in two stages and it is computationally simpler than the maximum likelihood method. In the first stage we estimate the parameters of the marginal distributions  $F_i$ , and in the second stage we estimate the copula parameters conditioned to the previous marginal distributions estimates. In each stage we use the maximum likelihood method. In an unpublished Ph.D. dissertation Xu (1996) used simulation to compare the exact maximum likelihood and the IFM estimators and found that the ratio of the mean square errors are approximately equal to one.

Under regularity condition Patton (2003b) showed that the IFM estimator is consistent and asymptotically normal.

#### 4.2 Conditional case

The conditional likelihood is given by the conditional version of the Sklar's Theorem. Let  $F_i$  be the conditional distribution of  $X_i|\mathbf{W}$  for  $i = 1, 2$  and let  $H$  be the joint (absolutely continuous) conditional distribution of  $\mathbf{X}|\mathbf{W}$ , where  $\mathbf{X} = (X_1, X_2)$  has conditional copula function  $C$ . Then, we have :

$$H(x_1, x_2|\mathbf{w}) = C(F_1(x_1|\mathbf{w}), F_2(x_2|\mathbf{w})|\mathbf{w}).$$

Thus an expression equivalent to (4.2) is given by:

$$c(F_1(x_1|\mathbf{w}), F_2(x_2|\mathbf{w})|\mathbf{w})f_1(x_1|\mathbf{w})f_2(x_2|\mathbf{w}),$$

where  $f_i(x_i|\mathbf{w})$  is the conditional density of  $X_i|\mathbf{W} = \mathbf{w}$  and

$$c(u_1, u_2|\mathbf{w}) = \frac{\partial^2 C(u_1, u_2|\mathbf{w})}{\partial u_1 \partial u_2}.$$



The log-likelihood expression is equivalent to (4.3) :

$$l(\vartheta) = \sum_{t=1}^T \ln c(F_1(x_{1,t}; \vartheta_1 | \mathbf{w}_t), \dots, F_n(x_{n,t}; \vartheta_n | \mathbf{w}_t); \theta | \mathbf{w}_t) + \sum_{t=1}^T \sum_{i=1}^2 \ln f_i(x_{i,t}; \vartheta_i | \mathbf{w}_t),$$

and we can use all the previous methods to estimate the parameters.

### 4.3 Empirical Copula

The empirical copula  $\hat{C}$  is defined as:

$$\hat{C}\left(\frac{t_1}{T}, \frac{t_2}{T}\right) = \frac{1}{T} \sum_{t=1}^T \mathbf{1}_{[x_{1,t} \leq x_{1,(t_1)}, x_{2,t} \leq x_{2,(t_2)}]}, \quad (4.4)$$

where  $\mathbf{1}$  is the indicator function,  $x_{i,(t_j)}$ ,  $i = 1, 2$ ,  $j = 1, 2$  are the  $t_j$ -th order statistics of the  $i$ -th variable and  $t_1, t_2 \in \{1, \dots, T\}$ . Therefore the empirical copula is the proportion of elements from the sample that satisfies  $x_{1,t} \leq x_{1,(t_1)}$  and  $x_{2,t} \leq x_{2,(t_2)}$ .

### 4.4 Selection of the copula function

The quadratic distance between two copulae  $C_1$  and  $C_2$  in a (finite) set of bivariate points  $\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m\}$  is defined as:

$$\bar{d}(C_1, C_2) = \left[ \sum_{i=1}^m (C_1(\mathbf{a}_i) - C_2(\mathbf{a}_i))^2 \right]^{1/2}. \quad (4.5)$$

Let  $\{C_k\}_{1 \leq k \leq K}$  be the set of copulae under consideration. One criterion is to select the copula  $C_k$  which minimizes the quadratic distance between  $C_k$ , the estimated copula, and the empirical copula  $\hat{C}$ , defined by (4.4), in the region of interest. For instance, when selecting a model in order to estimate the VaR, the region of interest should be the lower tail. We will use this region in the Subsection 5.5.

Another suggestion is to use a criterion like Akaike's information criterion (AIC) (Akaike, 1973) which is defined as

$$AIC(M) = -2 \log\text{-likelihood}(\hat{\theta}, \hat{\vartheta}) + 2M,$$

where  $M$  is the number of parameters being estimated and  $\hat{\cdot}$  denotes the maximum likelihood estimates.  $M$  and the parameters estimated depend on the selection approach. When we are selecting the copula function,  $M$  is the number of the copula parameters if the models for the marginal distributions are considered as known. Smith(2003) used this criterion for copula selection.

## 5. Application

### 5.1 Value at risk

In 1994 the American bank JP Morgan published a risk control method known as Riskmetrics, based mainly on a parameter named Value at Risk.

The Value at Risk is a forecast of a given percentile, usually in the lower tail, of the distribution of returns on a portfolio over some period  $\Delta_t$ . The VaR of a portfolio at time  $t$ , with confidence level  $1 - \alpha$ , where  $\alpha \in (0, 1)$  is defined as:

$$VaR_t(\alpha) = \inf\{s : F_{p,t}(s) \geq \alpha\},$$

where  $F_{p,t}$  is the distribution function of the portfolio return  $X_{p,t}$  at time  $t$  (return from  $t - \Delta_t$  to  $t$ ). Equivalently, we have  $P(X_{p,t} \leq VaR_t(\alpha)) = \alpha$  at time  $t$ . This means that we are  $100(1 - \alpha)\%$  confident that the loss in the period  $\Delta_t$  will not be larger than the VaR.

VaR is being used for several needs; risk reporting, risk limits, regulatory capital, internal capital allocation and performance measurement. An inadequate VaR estimation can lead to a underestimation of the risk incurred. On the other hand, a conservative position due to overestimation of the VaR would, for instance, not consider certain risk controlled positions.

We are going to work with one-day period VaR. Consider the case of a portfolio composed by two assets, with returns at  $t - th$  day, denoted as  $X_{1,t}$  and  $X_{2,t}$  respectively. The portfolio return, denoted as  $X_{p,t}$  is approximately equal to  $\omega_1 X_{1,t} + \omega_2 X_{2,t}$ , where  $\omega_1$  and  $\omega_2$  are the portfolio weights of assets 1 and 2. For the VaR estimation, we could study the distribution of the univariate portfolio return series, or the bivariate distribution of the vector  $(X_{1,t}, X_{2,t})$ .

In this paper we use copula theory to model the vector  $(X_{1,t}, X_{2,t})$ . The modeling is done in the following sequence :

- An exploratory data analysis is done in Subsection 5.2.
- The general model is presented in Subsection 5.3. The model selection and estimation is done in two steps :

ARMA-GARCH models are fitted for each return series in Subsection 5.4.

The bivariate distribution of the ARMA-GARCH models is modeled by three copulae families in Subsection 5.5. They are fitted for the innovations estimated in Subsection 5.4.

- The fitted models are used to estimate the VaR. A backtesting is used to test the VaR estimates. This is done in Subsection 5.6.
- The best copula model is compared to traditional VaR estimation methods in Subsection 5.7.

## 5.2 Data description

The theory presented is applied to a portfolio composed by Nasdaq and S&P500 stock indices. The database contains 2972 daily closing prices, from January 2nd 1992 to October 01st 2003. We denote the log-returns of Nasdaq as variable 1 and the log-returns of S&P500 as variable 2. Figure 1 presents the plots of both series and Table 2 contains descriptive statistics.

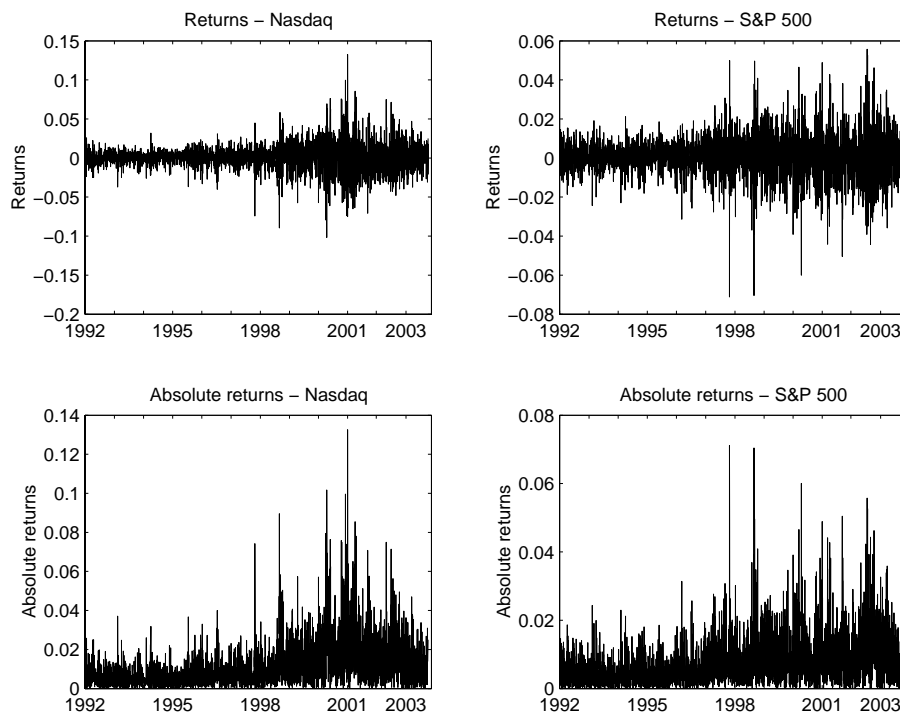


Figure 1: Daily returns and absolute returns of Nasdaq and S&P500 stock indices.

Table 2: Descriptive statistics of daily log-returns of Nasdaq and S&amp;P500 stock indices.

Statistics	Nasdaq	S&P500
Mean	0.00038	0.00030
Mean (annualized)	10.141%	7.857%
Standard Deviation	0.01694	0.01076
Minimum	-0.10168	-0.07113
Median	0.00122	0.00028
Maximum	0.13255	0.05574
Excess of Kurtosis	4.91481	3.78088
Asymmetry	0.01490	-0.10267

In Figure 1 we can see the evidence of the stylized fact known as volatility clustering, in which large absolute returns tend to follow large absolute returns and the same for small returns. Table 2 shows that the annualized means of both series are positive. Both return series distributions are nearly symmetric and have large kurtosis, with the Nasdaq presenting the larger one. We do not present the autocorrelation functions of the series, but for the Nasdaq returns, only the autocorrelations of lag 12 and 13 are significant at the 5% level ( $t$  statistic equals to 3.68 and 4.48 respectively). There is no significant correlation for the S&P500 returns at the 5% level.

### 5.3 The model

In specifying the bivariate model we must specify the two models for the marginal variables and the model for the conditional copula. The models for the univariate variables must take into account the characteristics of the variables. Return series have been successfully modeled by ARMA-GARCH models by many authors and will be used here. For instance, for an AR(1)xGARCH(1,1) the models for the margins are given by:

$$\begin{aligned}
 X_{i,t} &= \mu_i + \phi_i X_{i,t-1} + \epsilon_{i,t}; \\
 \epsilon_{i,t} &= \sigma_{i,t} \eta_{i,t}; \\
 \sigma_{i,t}^2 &= \alpha_i + \beta_i \epsilon_{i,t-1}^2 + \gamma_i \sigma_{i,t-1}^2,
 \end{aligned}
 \tag{5.1}$$

where  $i = 1, 2$ ,  $\{\eta_{1,t}\}$  and  $\{\eta_{2,t}\}$  are white noise processes with zero mean and unit variance,  $\alpha_i, \beta_i, \gamma_i$  follow Nelson and Cao (1992) restrictions; and  $\beta_i + \gamma_i < 1$ , for  $i = 1, 2$ . The conditional distribution of the standardized innovations

$$\eta_{i,t} = \frac{\epsilon_{i,t}}{\sigma_{i,t}} | \mathcal{F}_{i,t-1}, \quad i = 1, 2$$

was modeled by standard normal and standard t distributions. We denoted these models, respectively by GARCH-N, GARCH-t. The distribution of the innovation vector  $\eta_t = (\eta_{1,t}, \eta_{2,t})$  is modeled by copula. This approach was applied, for instance, by Dias and Embrechts (2003) and Patton (2003a). The ARMA  $\times$  GARCH models work as a filter in order to have innovation processes, which are serially independent.

In order to apply the copula models we need to specify the conditional marginal distributions. Additionally to the normal and t distributions we are also going to use the empirical distribution of the residuals (estimates of the standardized innovations). According to the IFM method the selected copula functions will be fitted to these residuals series. The estimation of the models is done in the following subsections.

#### 5.4 Modeling the marginal distributions

We fitted two AR(1)-GARCH(1,1) models for the series  $X_1$  and  $X_2$  as initial models with normal and t distributions. Table 3 presents maximum likelihood estimates obtained without constraints, using the MATLAB maximization function. We considered other ARMA models, such as AR(2) and ARMA(1,1) for the level of returns, but the additional parameters estimated were not significant (at the 5% level) for both series.

Table 3: Parameter estimates of GARCH models and standard errors.

Parameter	GARCH-N	GARCH-t
$\mu_1$	$8.25 \times 10^{-4}$ ( $1.85 \times 10^{-4}$ )	$9.43 \times 10^{-4}$ ( $1.79 \times 10^{-4}$ )
$\phi_1$	0.0846 (0.0199)	0.0884 (0.0188)
$\alpha_1$	$1.44 \times 10^{-6}$ ( $4.06 \times 10^{-7}$ )	$6.13 \times 10^{-7}$ ( $2.69 \times 10^{-7}$ )
$\beta_1$	0.1022 (0.0145)	0.0652 (0.0111)
$\gamma_1$	0.8958 (0.0140)	0.9180 (0.0134)
$\nu_1$		9.6859 (0.0773)
$AIC_1$	$-1.738 \times 10^4$	$-1.745 \times 10^4$
$\mu_2$	$5.18 \times 10^{-4}$ ( $1.42 \times 10^{-4}$ )	$5.66 \times 10^{-4}$ ( $1.35 \times 10^{-4}$ )
$\phi_2$	0.0156 (0.0194)	-0.0021 (0.0188)
$\alpha_2$	$5.76 \times 10^{-6}$ ( $1.87 \times 10^{-7}$ )	$2.55 \times 10^{-7}$ ( $1.20 \times 10^{-7}$ )
$\beta_2$	0.0701 (0.0097)	0.0430 (0.0069)
$\gamma_2$	0.9273 (0.0097)	0.9397 (0.0093)
$\nu_2$		7.2459 (0.1516)
$AIC_2$	$-1.939 \times 10^4$	$-1.950 \times 10^4$

Figure 2 presents the normal and t probability plots of the transformed series:

$u_t = F_{1,t}(x_{1,t}|\mathcal{F}_{t-1})$  and  $v_t = F_{2,t}(x_{2,t}|\mathcal{F}_{t-1})$ , where  $F_{1,t}$  and  $F_{2,t}$  are marginal distributions conditioned to  $\mathcal{F}_{t-1}$ , the information available up to time  $t - 1$ . If the models were correctly specified then both series will be standard uniform series. The fit seems quite good.

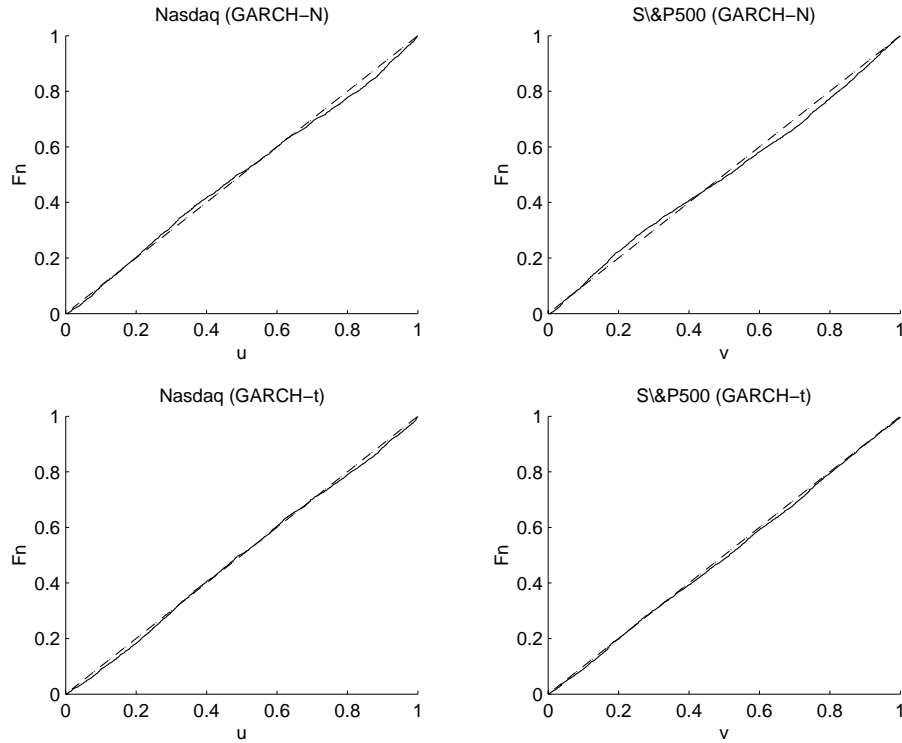


Figure 2: Empirical distribution of transformed series  $u_t$  and  $v_t$ .

The Ljung-Box test applied to the residuals of the GARCH-N and GARCH-t models does not reject the null hypothesis of null autocorrelations from lag 1 to 10 for the residuals for both series at the 5% significance level. The p-values are 0.320 and 0.087 for the GARCH-N and GARCH-t models respectively, for the first series, and 0.185 and 0.083 for the second series. The Ljung-box test also does not reject the null hypothesis from lag 1 to 10 for the square of the residuals series at the 5% significance level. The p-values are 0.959 and 0.759 for the square of residuals of the first series and 0.968 and 0.636 for the second. Therefore, we consider that the models are adequate.

## 5.5 Modeling the dependence between series

Figure 3 shows the plots of the bivariate residual series for both GARCH-N and GARCH-t models. We can see that there is a positive dependence between the two series.

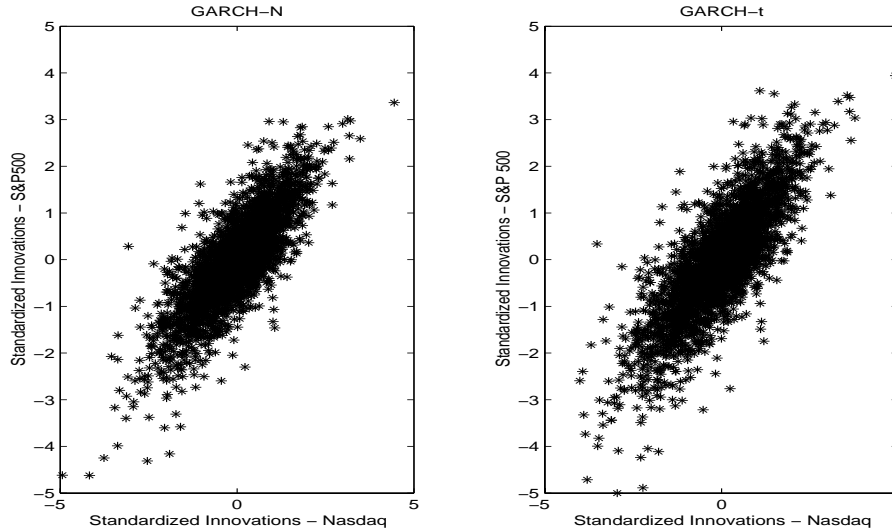


Figure 3: Scatterplot of the Nasdaq and S&P500 dollar standardized innovations (GARCH-N and GARCH-t models).

We are going to use three copula functions earlier presented in this paper and which have been applied by other authors: SJC, t and Plackett copulae with normal, t and empirical marginal distributions. The normal and t distributions are used with the residuals of the adjustment of the GARCH-N and GARCH-t, respectively. The empirical marginal distribution was used with the residuals of the adjustment of the GARCH-N model. We could consider it as model GARCH-E estimated by quasi-maximum likelihood method. Table 4 presents the IFM estimates of the copula function parameters.

The quality of the adjustment can be assessed by the quadratic distance defined in (4.5) between the estimated and the empirical copulae. Since our main interest relies on the extreme losses it is interesting to measure the distance only in the lower tail. Table 5 presents the quadratic distance in (4.5) only in the square region  $0 \leq u, v \leq \gamma$  for  $\gamma = 0.05, 0.1$  and  $0.2$ . In this table “model Plackett + GARCH-E”, for instance, means model GARCH for the margins estimated with normal innovations and Plackett copula with empirical distribution fitted to the residuals series. The analysis of the results in Table 5 shows that the t copula provided the best adjustment. The results obtained using the SJC copula

Table 4: Parameter estimates and standard errors for SJC, Plackett and Student- $t$  copulae

Copula	Parameter	GARCH-N	GARCH-t	GARCH-E
SJC	$\tau_L$	0.6706 (0.0092)	0.6887 (0.0080)	0.7124 (0.0072)
	$\tau_U$	0.5937 (0.0151)	0.5858 (0.0152)	0.5286 (0.0188)
	(AIC)	$-3.9918 \times 10^4$	$-4.0208 \times 10^4$	
Plackett	$\theta$	21.6216 (0.9248)	20.0555 (0.8699)	19.9550 (0.8630)
	(AIC)	$-3.9939 \times 10^4$	$-4.0047 \times 10^4$	
Student- $t$	$\nu$	9.5305 (0.6648)	6.9129 (0.5210)	7.5432 (0.5865)
	$R_{12}$	0.8248 (0.0052)	0.8174 (0.0063)	0.8175 (0.0056)
	(AIC)	$-4.0117 \times 10^4$	$-4.0291 \times 10^4$	

Table 5: Quadratic distance between the estimated and empirical copulae in the left tail (square  $[0, \gamma]^2$ ).

Model	$\gamma = 0.20$	$\gamma = 0.10$	$\gamma = 0.05$
SJC + GARCH-N	4.8896	0.7971	0.4328
SJC + GARCH-t	3.9492	0.8381	0.5181
SJC + GARCH-E	<b>2.8348</b>	1.0524	0.6273
Plackett + GARCH-N	6.7868	2.1017	0.5102
Plackett + GARCH-t	7.4747	2.2761	0.5530
Plackett + GARCH-E	7.5159	2.2880	0.5559
$t$ -Student + GARCH-N	3.6668	<b>0.6846</b>	0.2483
$t$ -Student + GARCH-t	3.8616	0.7011	0.2553
$t$ -Student + GARCH-E	3.8996	0.7102	<b>0.2467</b>

were also reasonable (specially with GARCH-E margins for the  $[0; 0.20]^2$  region), except for the  $[0; 0.05]^2$  region.

## 5.6 Estimating the value at risk

We arbitrarily consider a portfolio with equal weights for both indices, but this is not a constraint and they can vary freely. Hence, considering that the returns are small, we have approximately  $X_{p,t} \approx (1/2)X_{1,t} + (1/2)X_{2,t}$ .



In order to assess the accuracy of the VaR estimates we backtested the method at 95%, 99% and 99.5 % confidence level by the following procedure. Since at time  $t$  you only know the data up to this point the VaR must be evaluated with the model (AR(1) - GARCH(1,1) + copula) estimated using only information up to this time. After estimating the total model we can simulate from the estimated copula to have an estimate of the joint distribution of the vector of innovations  $\eta_{t+1}$ . Using model (5.1) we can have an estimate of the portfolio distribution and estimate the VaR. Since estimating the model and simulating from the copula could be computationally too cumbersome we estimated the model only once at every 50 observations because we did not expect to have large difference in the estimated models when modifying a fraction of the observations.

We initially estimated the parameters model using data from  $t = 1$  to  $t = 750$ . Then we simulated 5000 values of the innovations for  $(\eta_{1,751}, \eta_{2,751})$ , evaluated the VaR value for  $\alpha = 0.005, 0.01$  and  $0.05$  and tested whether  $x_{p,751}$  was below these estimated VaR values. For the next 49 observations we used this same estimated model. This means that we used the same copula simulation for each 50 observations, but at each new observation we up-to-dated the VaR estimates because the conditional level and variance estimates was up-to-dated according to model (5.1). At observations  $t = 800, 850, \dots, 2950$  we re-estimated the model and repeated the whole process. We always used a window of 750 observations. Since we have 2971 observations we had a total of 2220 tests for VaR at each level. The results are presented in Table 6 and Figure 4.

Table 6: Proportion of observations (number of observations in brackets), for  $t = 751$  to  $2971$ , where the portfolio loss exceeded the estimated VaR for  $\alpha = 0.005, 0.01$  and  $0.05$ .

Copula	GARCH-N	GARCH- $t$	GARCH-E
$\alpha = 0.05$ (111)			
SJC	0.0675 (150)	0.0698 (155)	<b>0.0558 (124)</b>
Plackett	0.0684 (152)	0.0707 (157)	0.0576 (128)
$t$ -Student	0.0716 (159)	0.0761 (169)	0.0626 (139)
$\alpha = 0.01$ (22)			
SJC	0.0185 (31)	0.0149 (33)	<b>0.0104 (23)</b>
Plackett	0.0225 (50)	0.0176 (39)	0.0131 (29)
$t$ -Student	0.0243 (54)	0.0189 (42)	0.0126 (28)
$\alpha = 0.005$ (11)			
SJC	0.0126 (28)	0.0090 (20)	<b>0.0041 (9)</b>
Plackett	0.0162 (36)	0.0122 (27)	0.0077 (17)
$t$ -Student	0.0171 (38)	0.0126 (28)	0.0077 (17)

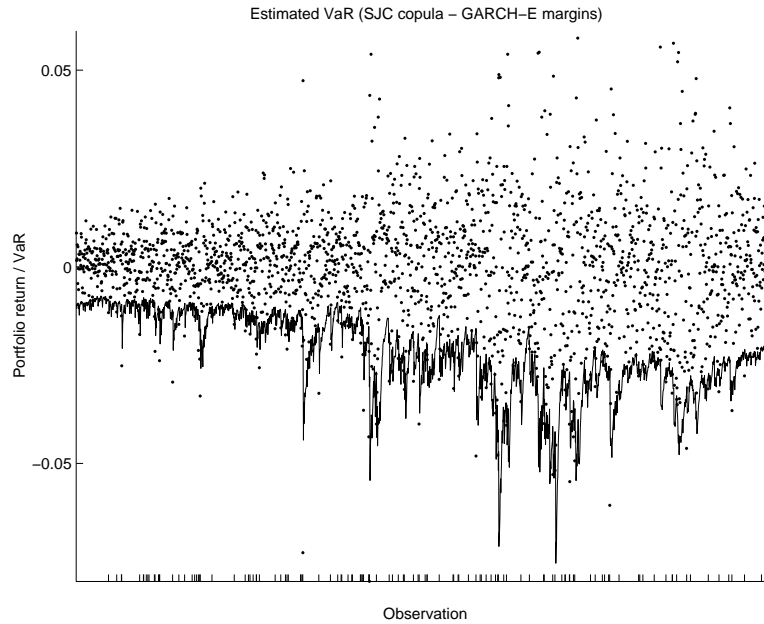


Figure 4: Out of sample one step ahead estimated 5% portfolio VaR (SJC copula and GARCH-E margins) and observed returns. In the x-axis, observations where the observed portfolio loss was larger than the estimated VaR.

The exercise results showed that the GARCH-E model provided the best performance for VaR estimation. The model with SJC copula and GARCH-E margins was the best for all  $\alpha$  levels considered. The t copula was the model with smallest distance to the empirical copula (see Table 5) for the regions  $[0, 0.05]^2$  and  $[0, 0.10]^2$ , but the SJC copula provided the best performance in the backtest.

### 5.7 Comparison of the value at risk estimates

We estimated the VaR using different approaches for benchmark purposes. We worked directly with the univariate portfolio return series, estimating the VaR using univariate GARCH model, historical simulation and Exponential Weighted Moving Average (EWMA) Methods. We also used Bivariate GARCH (BEKK and DCC) models and bivariate EWMA Method. We are going to present these methods briefly in the following.

We fitted GARCH(1,1) model directly to the portfolio returns series  $X_{p,t} = (1/2)X_{1,t} + (1/2)X_{2,t}$ . The GARCH(1,1) model is given by:

$$X_{p,t} = \mu + \sigma_t \epsilon_t ;$$

$$\sigma_t^2 = \alpha + \beta \epsilon_{t-1}^2 + \gamma \sigma_{t-1}^2 ,$$

where  $\{\epsilon_t\}$  is a white noise process with zero mean and unit variance and  $\alpha, \beta$  and  $\gamma$  follow the same restrictions given in model (5.1). We considered the Normal and t distributions for the innovations.

The EWMA method is generally used in the Riskmetrics methodology. Let  $\sigma_{p,t}^2$  be the variance of the portfolio return in time  $t$ . The estimated variance, using data up to time  $t - 1$ , is given by :

$$\sigma_{p,t/t-1}^2 = (1 - \lambda)x_{p,t-1}^2 + \lambda\sigma_{p,t-1/t-2}^2,$$

where  $\sigma_{p,(t+1)/t}^2$  is the smoothed variance considering data up to time  $t$ . Considering the normal distribution, we have  $x_{p,t} \sim \mathcal{N}(0, \sigma_{p,t}^2)$ . The parameter  $\lambda$  is re-estimated on each day  $t$ , minimizing the quantity

$$\sum_{i=50}^{t-1} (\sigma_{p,(i+1)/i}^2 - x_{p,i+1}^2)^2.$$

We can also use the EWMA method in the bivariate case. In this case the estimated variances and covariance are given by :

$$\begin{aligned} \sigma_{1,t/t-1}^2 &= (1 - \lambda_1)x_{1,t-1}^2 + \lambda_1\sigma_{1,t-1/t-2}^2; \\ \sigma_{2,t/t-1}^2 &= (1 - \lambda_2)x_{2,t-1}^2 + \lambda_2\sigma_{2,t-1/t-2}^2; \\ \sigma_{12,t/t-1} &= (1 - \lambda_{12})x_{1,t-1}x_{2,t-1} + \lambda_{12}\sigma_{12,t-1/t-2}, \end{aligned}$$

where the optimal parameters  $\lambda_1, \lambda_2$  and  $\lambda_{12}$  are obtained at every time  $t$  minimizing the quantities

$$\begin{aligned} &\sum_{i=50}^{t-1} (\sigma_{1,(i+1)/i}^2 - x_{1,i+1}^2)^2, \\ &\sum_{i=50}^{t-1} (\sigma_{2,(i+1)/i}^2 - x_{2,i+1}^2)^2, \quad \text{and} \\ &\sum_{i=50}^{t-1} (\sigma_{12,(i+1)/i} - x_{1,i+1}x_{2,i+1})^2 \end{aligned}$$

respectively.  $\sigma_{1,t/t-1}^2, \sigma_{2,t/t-1}^2$  and  $\sigma_{12,t/t-1}$  are estimates of the bivariate covariance matrix and interpreted similarly to the univariate case.

We also considered an exponential smoothing for the returns mean, but the results obtained were very close, so we maintained the model with zero means.

We also used the historical simulation approach. For this approach, consider some day  $t_0$ . The estimated VaR for the day  $t_0 + 1$  with confidence level  $1 - \alpha$

is given by the  $\alpha t_0$ -th order statistics of the sample of portfolio returns for  $t = 1, \dots, t_0$ .

The first class of bivariate GARCH models used is the BEKK model (Engle and Kroner, 1995). Let

$$\mathbf{x}_t = \begin{pmatrix} x_{1,t} \\ x_{2,t} \end{pmatrix}, \quad \epsilon_t = \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix},$$

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}, \quad C = \begin{pmatrix} c_{11} & c_{12} \\ 0 & c_{22} \end{pmatrix} \quad \text{and let}$$

$$\Sigma_t = \begin{pmatrix} \sigma_{1,t}^2 & \sigma_{12,t} \\ \sigma_{12,t} & \sigma_{2,t}^2 \end{pmatrix} = \text{Cov}[\epsilon_t | \mathfrak{S}_{t-1}].$$

The BEKK model is given by :

$$\begin{cases} \mathbf{x}_t = \mu + \epsilon_t ; \\ \Sigma_t = C' C + A' \epsilon_{t-1} \epsilon_{t-1}' A + B' \Sigma_{t-1} B . \end{cases}$$

We assumed that the innovations vector  $\epsilon_t$  has conditionally a normal distribution  $\mathcal{N}(0, \Sigma_t)$ . The program used for the estimation was the SAS 8.2, proc VARMAX.

The second class of bivariate GARCH models used is the DCC one, proposed by Engle (2002). Let  $x_i$  be the returns with mean equals zero, for  $i=1, \dots, n$ . The conditional correlation and variances are defined by :

$$\sigma_{i,j,t} = \text{E}[x_{i,t} x_{j,t} | \mathfrak{S}_{t-1}] / \sqrt{\text{E}[x_{i,t}^2 | \mathfrak{S}_{t-1}] \text{E}[x_{j,t}^2 | \mathfrak{S}_{t-1}]}.$$

Let  $\sigma_{i,t}^2 = \text{E}[x_{i,t}^2 | \mathfrak{S}_{t-1}]$  and  $\eta_{i,t} = x_{i,t} / \sigma_{i,t}$ . So the correlation can be written as  $\sigma_{i,j,t} = \text{E}[\eta_{i,t} \eta_{j,t} | \mathfrak{S}_{t-1}]$ . Engle (2002) suggests estimating the GARCH processes

$$q_{i,j,t} = \bar{\sigma}_{i,j} + \alpha(\eta_{i,t-1} \eta_{j,t-1} - \bar{\sigma}_{i,j}) + \beta(q_{i,j,t-1} - \bar{\sigma}_{i,j})$$

for  $i, j = 1, \dots, n$  and obtaining  $\sigma_{i,j,t} = q_{i,j,t} / \sqrt{q_{i,i,t} q_{j,j,t}}$ . We can interpret  $\bar{\sigma}_{i,j}$  as the unconditional correlation between the innovations of the univariate GARCH models  $\eta_{i,t} \eta_{j,t}$ . Therefore the correlations and variances are modeled as GARCH processes with common parameters  $\alpha$  and  $\beta$ , and with different unconditional expected values  $\bar{\sigma}_{i,j}$ .

We began the VaR estimation using data from  $t = 1$  to  $t = 750$ . In the case of the bivariate and univariate GARCH models the model parameters were re-estimated after each 50 returns, like before for copulae. For the EWMA and Historical Simulation methods, the estimates were updated day by day from  $t = 750$  on.

Table 7 compares the benchmark models and the best model with copula (SJC copula with GARCH-E margins).

Table 7: Proportion of observations (number of observations in brackets), for  $t = 751$  to 2971, where the portfolio loss exceeded the estimated VaR for  $\alpha = 0.005, 0.01$  and  $0.05$

Model	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.005$
SJC Copula + GARCH-E	0.0558 (124)	<b>0.0104 (23)</b>	<b>0.0041 (9)</b>
Bivariate GARCH (BEKK)	0.0819 (182)	0.0338 (75)	0.0248 (55)
Bivariate GARCH (DCC)	0.0432 (96)	0.0140 (31)	0.0113 (25)
EWMA (Bivariate)	0.0387 (86)	0.0144 (32)	0.0104 (23)
GARCH-N (Portfolio)	0.0666 (148)	0.0207 (46)	0.0144 (32)
GARCH- $t$ (Portfolio)	0.0693 (154)	0.0131 (29)	0.0104 (23)
EWMA (Portfolio)	<b>0.0527 (117)</b>	0.0135 (30)	0.0099 (22)
Historical Simulation (Portfolio)	0.1220 (271)	0.0293 (65)	0.0144 (32)

We can see that for  $\alpha = 0.01$  or  $\alpha = 0.005$  the copula model provided a far better result in the VaR estimation. For  $\alpha = 0.05$  the copula model was the second best model, but close to the best model.

## 5.8 Concluding remarks

This work showed how conditional copula theory can be a very powerful tool in estimating the VaR. We used different copulae and marginal distributions for the GARCH innovations. We compared the results obtained with traditional methods of VaR estimation. The SJC copula which allows for different dependence in the tails produced the best result and produced reliable VaR limits. Further researches must be done in order to extend this type of application to higher dimensions.

## Acknowledgments

We acknowledge grants from FAPESP and from CNPq. We also thank the Referee and the Editor for valuable suggestions.

---

**References**

- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. *Second International Symposium on Information Theory*, 267-281, Akademiai Kiado, Budapest.
- Ang, A. and J. Chen (2002). Asymmetric Correlations of Equity Portfolios, *Journal of Financial Economics* **63**, 443-494.
- Bouyé, E., Durrleman, V., Nikeghbali, A., Riboulet, G. and Roncalli, T. (2000). Copulas for finance, a reading guide and some applications. Working paper, Financial Econometrics Research Center, City University, London.
- Cherubini, U. and Luciano, E. (2001). Value at risk trade-off and capital allocation with copulas. *Economic Notes* **30**, 235-256.
- Dias, A., and Embrechts, P. (2003). Dynamic copula models for multivariate high-frequency data in finance. Working Paper, ETH Zurich: Department of Mathematics.
- Embrechts, P. and Hoing, A., Juri, A. (2003). Using copulae to bound the value-at-risk for functions of dependent risks. *Finance and Stochastics* **7**, 145-167.
- Embrechts, P., Lindskog, F. and McNeil, A.J. (2003). Modelling dependence with copulas and applications to risk management. In *Handbook of Heavy Tailed Distributions in Finance* (Edited by S. T. Rachev), 329-384 Elsevier
- Embrechts, P., McNeil, A. and Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls. In *Risk Management Value at Risk and Beyond* (Edited by M. Dempster), 176-223. Cambridge University Press.
- Engle, R. (2002). Dynamic conditional correlation - a simple class of multivariate GARCH. *Journal of Business and Economic Statistics* **20**, 339-350.
- Engle, R., Kroner, K. F. (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory* **11**, 122-150.
- Fortin, I. and Kuzmics, C. (2002). Tail dependence in stock return pairs. *International Journal of Intelligent Systems in Accounting, Finance & Management* **11**, 89-107.
- Georges, P., Lamy, A.G., Nicolas, E., Quibel, G. and Roncalli, T. (2001). Multivariate survival modelling: a unified approach with copulas. Working paper, Credit Lyonnais, Paris.
- Hansen, B. (1994). Autoregressive conditional density estimation. *International Economic Review* **35**, 705-730.
- Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman and Hall.
- Kimberling, H. C. (1974). A probabilistic interpretation of complete monotonicity. *Aequationes Mathematicae* **10**, 152-164.

- Longin, F. and Solnik, B. (2001). Extreme correlation of international equity markets. *Journal of Finance* **56**, 649-676.
- Mashal, R. and Zeevi, A. (2002). Beyond correlation : Extreme co-movements between financial assets. Working paper, Columbia University.
- Meneguzzo, D. and Vecchiato, W. (2002). Copulas sensitivity in collateralized debt obligations and basket defaults swaps pricing and risk monitoring. Working paper, Veneto Banca.
- Nelsen, R. B. (1999). *Introduction to Copulas*, Springer Verlag.
- Nelson, D. B. and Cao, C. Q. (1992). Inequality constraints in the univariate GARCH model. *Journal of Business and Economic Statistics* **10**, 229-235.
- Patton, A. (2003a). Modelling asymmetric exchange rate dependence. Working paper, University of California, San Diego.
- Patton, A. (2003b). Estimation of multivariate models for time series of possibly different lengths. Working paper, University of California at San Diego.
- Plackett, R. L. (1965). A class of bivariate distributions. *Journal of American Statistical Association* **60**, 516-522.
- Rockinger, M. and Jondeau, E. (2001). Conditional dependency of financial series: an application of copulas. Working paper NER # 82, Banque de France. Paris.
- Smith, M. (2003) Modelling sample selection using archimedean copulas. *Econometrics Journal*, **6**, 99-123.
- Xu, J. J. (1996). Statistical modelling and inference for multivariate and longitudinal discrete response data. PhD thesis, Statistics Department, University of British Columbia.

Received April 16, 2004; accepted September 27, 2004.

Helder Parra Palaro  
University of Campinas  
Department of Statistics, CP 6065  
13083-970, Campinas, SP, Brazil.  
helder1977@yahoo.com.br

Luiz Koodi Hotta  
University of Campinas  
Department of Statistics, CP 6065  
13083-970, Campinas, SP, Brazil.  
hotta@ime.unicamp.br.