The Copula GARCH model for time varying betas in the banking sector.

INTERNATIONAL HELLENIC UNIVERSITY
Master of Science dissertation

by
Dimitris Nikolaidis

Dissertation Director: Dr. Markellos

August 2010

Contents

Abstract

The Copula GARCH model for time varying betas in the banking sector.

Dimitris Nikolaidis

2010

Copula functions have become an increasingly popular tool in finance when the distribution of asset returns is of extreme importance. The main features of copulas are that they seperate a multivariate distribution into the dependence structure and the margins, thus allowing two step estimation procedures for the distributional parameters that minimize the computational burden and also add flexibility to the distribution since the dependence governed by the copula and the margins do not have to belong to the same parametric family, unlike standard multivariate distributions. The aim of this study is twofold. In the first part, the statistical attributes of copulas are discussed in full detail while in the second part an empirical investigation of the evolution of stock betas during the modern global financial crisis period is conducted. In the empirical part, it is evident that copula models clearly outperform other, traditional models, in terms of both statistical validity and accuracy in risk calculations

Part I

The First Part - Statistical analysis of copulas

Chapter 1

Introduction to Copulas

The aim of this chapter is to provide to the reader a rigorous introduction to copulas. It contains the intuition behind copulas, the main theorems and definitions of the copula theory, discusses copula inference and presents an overview of the applications of copulas in finance. For more information, the interested reader is referred to the books of Joe (1996), Nelsen (2006) or Cherubinni et al. (2004).

1.1 Intuition behind copulas. Measures of dependence

In financial applications the usual measure of dependence is Pearson's linear correlation. For two random variables X and Y Pearson's correlation ϱ, is defined as:

$$\varrho\left(X,Y\right)=\frac{cov(X,Y)}{\sigma_X\sigma_Y}$$

However ϱ measures only linear dependence. It is the correct measure of dependence only in the Gaussian world, that is only if $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$. Linear correlation does not only incorporate information about the dependence of X and Y but also about their marginal behavior, that is why it is not invariant under strictly increasing transformations of the data. In other words the linear correlation of X and Y is different of that of $f(X)$ and $g(Y)$ for arbitrary strictly increasing functions f and g. To make this more clear consider the case of two independent standard normal variables X, Y and the transformation $f(x) = exp(x)$. Figure 1.1 contains the Scatterplots of 100 realizations of (X, Y) (left panel) and $(f(X), f(Y))$ (right panel). The distortion of the dependence structure caused by the nonlinear transformation is clear.

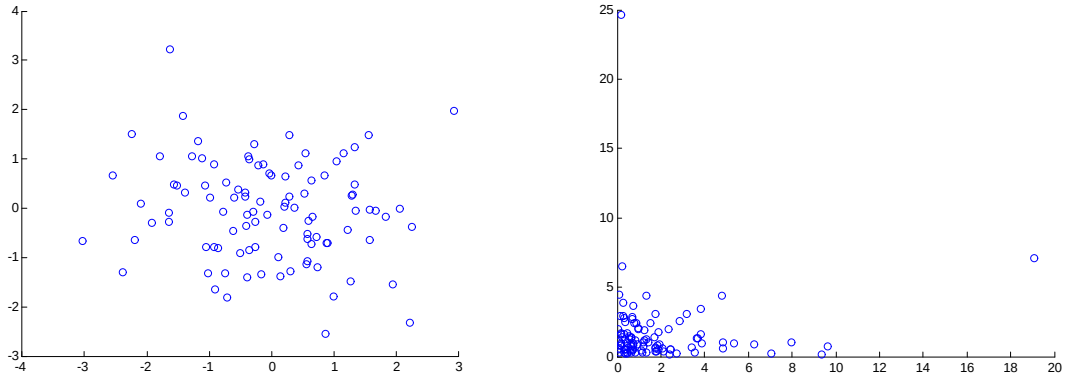


Figure 1.1: Scatterplot of 100 realizations of two independent standard normal variables (left panel) and of their transformation, created by $f(x) = \exp(x)$

Therefore a measure of dependence more appropriate that the linear correlation should not depend on X and Y and thus preserve the dependence for both pairs (X, Y) and $(f(X), g(Y))$. The remedy to this invariance problem is to use not the

data sample but the *ranks* associated with the sample. Let s_{Xi} and s_{Yi} denote the ranks of observations X_i and Y_i (*Note*: The rank of the observation X_i is the position of X_i if the sample was set in increasing order, that is the smallest value of X_i would have rank 1, the second smallest value of X_i would have rank 2, the biggest value of X_i, in a sample of size n, would have rank n and so on). But since the ranks depend on the sample size, instead of them the *standardized ranks*, S_{Xi}, S_{Yi} are used

$$S_{Xi} = \frac{s_{Xi}}{n+1}, \quad S_{Yi} = \frac{s_{Yi}}{n+1}$$

Obviously S_{Xi} and S_{Yi} belong in $(0, 1)$ where n is the sample size. The pairs (S_{Xi}, S_{Yi}) retain the greatest amount of information for the dependence of X and Y. This is illustrated in figure (1.2), where the scatterplots of the ranks of 100 realizations from X and Y (left panel) and $f(X)$ and $f(Y)$ (right panel) are drawn. We observe from figure two that the two scatterplots are identical, therefore both pairs of variables have exactly the same dependence structure.

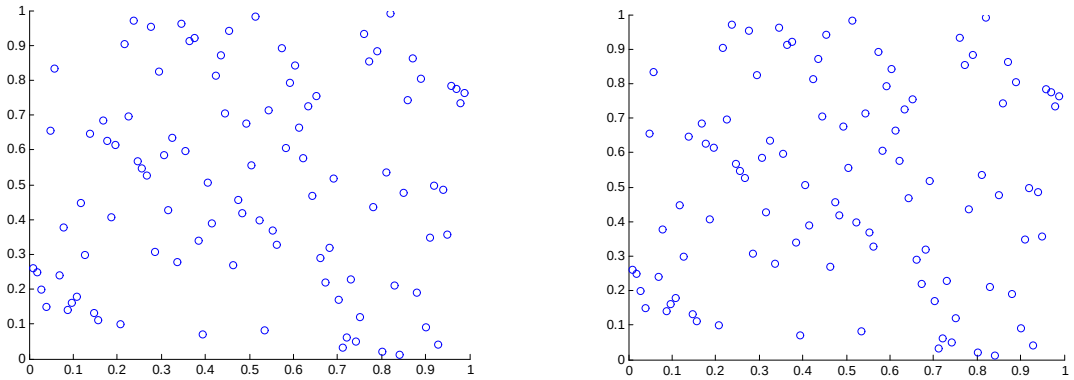


Figure 1.2: Scatterplots of ranks of 100 realizations of two standard normal variables (left panel) and of the ranks of the transformation of the variables based on $f(x) = \exp(x)$

Nevertheless, the dependence structure of two variables is fully characterized by the joint distribution $\boldsymbol{F}$, of the two variables and obviously the distribution function of the pair (X, Y) is not the same as the pair (S_X, S_Y). Hence a question arises: Is it possible to find (if it exists) a joint distribution, say $\boldsymbol{H}$, so as to $\boldsymbol{F}(X, Y) = \boldsymbol{H}(S_X, S_Y)$? If such a function exists, it would describe the dependence structure of X and Y in full, since both S_X and S_Y do not depend on the marginal behavior of X and Y. The answer to that question was provided by Sklar's theorem (it will be presented in the next section). According to it, if X and Y are continuous, such a function always exists and it is unique. This function is called the *copula* of X and Y, which is a Latin word that means *link*, because it couples the variables to define their dependence structure.

1.1.1 Rank based measures of dependence

There are many rank based measures of dependence proposed in the statistical literature. However in this study we will deal only with Srearman's rho and Kendall's tau because they are highly intuitive and they are directly connected to copulas, as it will become clear in later sections. For more information about other dependence measures the interested reader is referred to Genest and Verret (2005).

Spearman's rho

Spearman's rho is nothing more than Pearson's linear correlation calculated for the ranks of the data (*Note*: To what follows the word *ranks* will stand for *standard-*

ized ranks, unless otherwise noted)

$$r_n = \frac{\sum_{i=1}^{n}(S_{Xi} - \bar{S}_X)(S_{Yi} - \bar{S}_Y)}{\sqrt{\sum_{i=1}^{n}(S_{Xi} - \bar{S}_X)^2 \sum_{i=1}^{n}(S_{Yi} - \bar{S}_Y)^2}}$$

where $\bar{S}_X = \frac{n+1}{n}\sum_{i=1}^{n} S_{Xi} = \frac{1}{2} = \frac{n+1}{n}\sum_{i=1}^{n} S_{Yi} = \bar{S}_Y$.Spearman's rho has another, more convenient form

$$r_n = \frac{12}{n}\sum_{i=1}^{n} S_{Xi} \cdot S_{Yi} - 3\frac{n+1}{n-1}$$

Spearman's rho is theoretically superior than Pearson's correlation for the following reasons

1. It holds that $E(r_n) = 1$ if $Y = f(X)$ for *any* increasing function f, as opposed to the linear correlation where $E(\varrho_n) = 1$ if $Y = f(X)$ and f is an increasing *linear* function. Analogously $E(r_n) = -1$ if $Y = f(X)$ for *any* decreasing function f, as opposed to the linear correlation where $E(\varrho_n) = -1$ if $Y = f(X)$ and f is an decreasing *linear* function.

2. Spearman's rho always exists unlike linear correlation. For example if X or Y follow student's t distribution with degree of freedom parameter less than two, the corresponding variance and therefore the linear correlation does not exist.

3. It is invariant under strictly monotonic transformations; that is $r_n(X, Y) = r_n(g(X), g(Y))$ unlike linear correlation where $\varrho(X, Y) \neq \varrho(g(X), g(Y))$, for any strictly increasing function g.

Kendall's tau

Kendall's tau is based on the notion of *concordance*: Two pairs (X_i, Y_i) and (X_j, Y_j) will be called concordant when $(X_i, Y_i)(X_j, Y_j) > 0$.If the two pairs are not concordant they will be called discordant. The sample estimator of Kendall's tau is

$$\tau_n = \frac{P_n - Q_n}{\binom{n}{2}} = \frac{4}{n(n-1)} P_n - 1$$

where P_n and Q_n are the number of concordant and discordant pairs.τ_n is a rank based measure since $(X_i, Y_i)(X_j, Y_j) > 0$ if and only if $(S_{Xi} - S_{Xj})(S_{Yi} - S_{Yj}) > 0$. Kendall's tau shares all three characteristics of Spearman's rho, that make it a superior dependence measure when compared to the linear correlation and further it is directly linked to a class of copulas named Archimedean copulas, as I will demonstrate to the next section.

1.2 Definitions and fundamental properties

Definition 1 *A copula is a multivariate distribution with uniform margins.*

An equivalent, more mathematical definition is the following:

Definition 2 *A Copula is a function* $\boldsymbol{C} : [0,1]^d \longrightarrow [0,1]$, *that satisfies the following conditions:*

1. *For all* $(u_1, u_2, ..., u_d)$ *in* $[0,1]^d$, *if at least one component* u_i *is zero then* $C(u_1, u_2, ..., u_d) = 0$

2. *For every* $u_i \in [0,1]$, $\boldsymbol{C}(1,...,1,u_i,...1) = u_i$, *for* $i = 1,2,...,d$

3. *For all* $[u_{11},u_{12}]\times[u_{21},u_{22}]\times...\times[u_{d1},u_{d2}]$ *d - dimensional rectangles in* $[0,1]^d$ *,it holds that*

$$\sum ... \sum (-1)^{i_1+...+i_d} C(u_{1i_1},u_{2i_2},...,u_{di_d}) \geqslant 0$$

Copulas allow as to separate the marginal behavior of a multivariate distribution, from the dependence structure as it was proven by Sklar (1959).

Theorem 3 *(Sklar's theorem) Let* $\boldsymbol{F}$ *be a* $d-$*dimensional distribution function with univariate margins* $F_1,...,F_d$ *with domains* $D_1,...,D_d$ *respectively. Then there exists a unique function* $\boldsymbol{H}$, *defined on* $D_1 \times D_2 \times ... \times D_d$, *such that:*

$$\boldsymbol{F}(x_1,...,x_d) = \boldsymbol{H}(F_1(x_1),...,F_d(x_d)).$$

The extension of H *to* $[0,1]^d$ *is a copula* $\boldsymbol{C}$. *For such a function* $\boldsymbol{C}$ *we have:*

$$\boldsymbol{F}(x_1,...,x_d) = \boldsymbol{C}(F_1(x_1),...,F_d(x_d)). \tag{1.1}$$

If the marginal distributions $F_i, i = 1,2,...,d$ *are continuous, the function* $\boldsymbol{H}$ *coincides with the copula* $\boldsymbol{C}$, *which is then unique. Conversely, for given univariate distributions functions* $F_1,...,F_d$ *and a* $d-$ *dimensional copula* $\boldsymbol{C}$, *the function defined by:*

$$\boldsymbol{F}(x_1,...,x_d) = \boldsymbol{C}(F_1(x_1),...,F_d(x_d))$$

is a $d-$ *dimensional distribution function with univariate margins* $F_1,...,F_d$.

Remark 4 *To everything that follows, distributions will be denoted with capital letters whereas densities will be denoted with small letters. Multivariate functions will be denoted will bold letters. Further, all marginal distributions are considered continuous and strictly increasing, unless otherwise mentioned. Finally, the letter u will be strictly used for uniform variables.*

From Sklar's Theorem we see that a joint distribution is a function of the marginal distributions and the copula. Since the copula does not depend on the margins we can say that it represents the dependence between the variables, that is why a copula is referred to as the *dependence structure* in the international literature. Sklar's theorem provides the tools to construct a joint distribution from its marginal distributions and a copula but it doesn't say anything on how the copula distribution can be constructed, from the corresponding joint distribution. Before doing so we need the definition of the probability integral transformation.

Let $X_1, ..., X_d$ be random variables with distributions $F_1, ..., F_d$, respectively. The *probability integral transformation* is a function $T : \mathbb{R}^d \longrightarrow [0,1]^d$, such that: $(x_1, ..., x_d) \xrightarrow{T} (F_1(x_1), ..., F_d(x_d))$. The inverse of the probability integral transformation $T^{-1} : [0,1]^d \longrightarrow \mathbb{R}^d$ is called the quantile transformation and it is defined as: $(u_1, ..., u_d) \xrightarrow{T^{-1}} \left(F_1^{-1}(u_1), ..., F_d^{-1}(u_d)\right)$. Obviously the probability integral transformation of a random variable is a uniform variable, as depicted in figure 1.3, where a histogram of 1000 simulated values from the t_5 distribution is plotted (left panel) and its corresponding probability integral transformation.

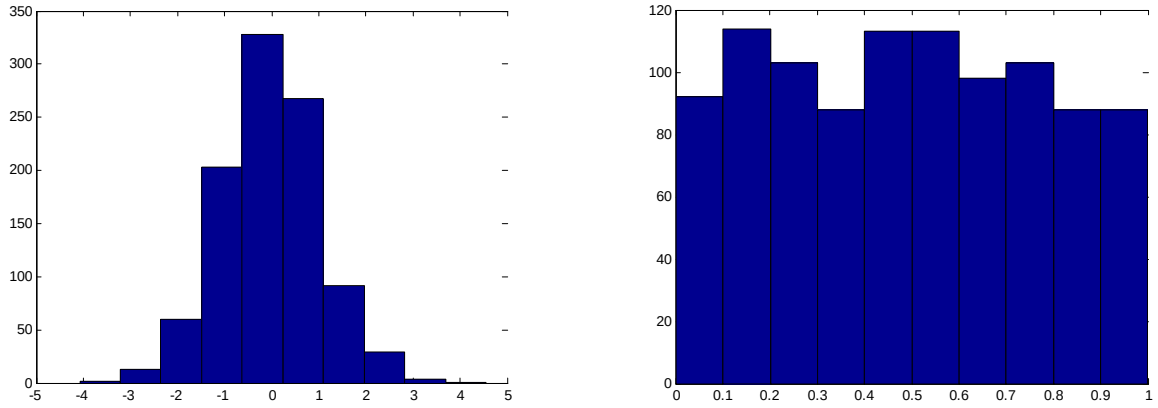


Figure 1.3: The probability integral transform of a random variable is a uniform variable

By applying the probability integral transform to Sklar's theorem the expression for the copula distribution function is derived:

Theorem 5 *(Inverse Sklar's theorem) Let* F_1^{-1},...,F_d^{-1} *denote the (quasi) inverse of the marginal distributions* $F_1, ..., F_d$. *Then for any* $\boldsymbol{u} \in [0,1]^d$, *it holds*

$$\boldsymbol{C}(u_1, ..., u_d) = \boldsymbol{F}\left(F_1^{-1}(u_1), ..., F_d^{-1}(u_d)\right) \tag{1.2}$$

where $u_i = F_i(x_i) \Leftrightarrow x_i = F_i^{-1}(u_i)$, $i = 1, 2, ..., d$

A very important attribute of copulas is that they remain invariant under strictly monotonic transformations of data.

Theorem 6 *(Copula Invariance) Let* $X_1, ..., X_d$ *be continuous random variables with copula* $\boldsymbol{C}$ *and let* $g_1, ..., g_n$ *be strictly increasing functions. Then the variables* $g(X_1), ..., g(X_d)$ *have exactly the same copula* $\boldsymbol{C}$.

From equations (1.1) and (1.2) the expression of the corresponding densities can be derived. By taking derivatives to equation (1.1) we have:

$$\begin{aligned} \boldsymbol{f}\left(x_1,...,x_d\right) &= \frac{\partial^d F\left(x_1,...,x_d\right)}{\partial x_1 \cdot ... \cdot \partial x_d} = \frac{\partial^d \boldsymbol{C}\left(F_1\left(x_1\right),...,F_d\left(x_d\right)\right)}{\partial x_1 \cdot ... \cdot \partial x_d} = \\ &= \frac{\partial^d \boldsymbol{C}\left(F_1\left(x_1\right),...,F_d\left(x_d\right)\right)}{\partial F_1\left(x_1\right) \cdot ... \cdot \partial F_d\left(x_d\right)} \prod_{i=1}^{d} \frac{dF_i\left(x_i\right)}{dx_i} = \\ &= \boldsymbol{c}\left(F_1\left(x_1\right),...,F_d\left(x_d\right)\right) \cdot \prod_{i=1}^{d} f_i\left(x_i\right) \end{aligned} \tag{1.3}$$

and, analogously, equation (1.2) becomes:

$$\boldsymbol{c}\left(u_1,...,u_d\right) = \frac{f\left(F_1^{-1}\left(u_1\right),...,F_d^{-1}\left(u_d\right)\right)}{\prod_{i=1}^{d} f_i\left(F^{-1}(u_i)\right)} \tag{1.4}$$

Equations (1.2) and (1.4) are of extreme importance because they provide a tool to create the copula density (and therefore the log - likelihood) that describes the dependence structure of some common multivariate densities.

Example 7 *(The Gaussian Copula) The Gaussian Copula describes the dependence structure of the multivariate Gaussian distribution. Let $\mathcal{X}$ be a $d-$dimensional random vector that follows the Gaussian distribution with zero mean and correlation matrix Σ. Then the random vector $\mathcal{U} = \left(\Phi\left(x_1\right),...,\Phi\left(x_d\right)\right)$ follows the Gaussian Copula, with distribution and density defined as:*

$$\boldsymbol{C}^G\left(u_1,...,u_d\right) = \boldsymbol{\Phi}_{\Sigma}\left(\Phi^{-1}\left(u_1\right),...,\Phi^{-1}\left(u_d\right)\right)$$

and

$$\boldsymbol{c}^{G}\left(u_1,...,u_d\right)=\frac{\boldsymbol{\phi}_{\Sigma}\left(\Phi^{-1}\left(u_1\right),...,\Phi^{-1}\left(u_d\right)\right)}{\prod_{i=1}^{d}\phi_i\left(\Phi^{-1}(u_i)\right)} \tag{1.5}$$

Example 8 *(The t - copula) The copula that corresponds to the multivariate t distribution with correlation matrix* Σ*, and degree of freedom parameter* v *is called the t - copula . Let* $\mathcal{X}$ *be a* $d-$*dimensional random vector that follows the t distribution with* v *degrees of freedom and correlation matrix* Σ*. Then the random vector* $\mathcal{U}=\left(t_v\left(x_1\right),...,t_v\left(x_d\right)\right)$ *follows the t copula and its distribution and density are:*

$$\boldsymbol{C}^{t}\left(u_1,...,u_d\right)=\mathbf{T}_{\Sigma,v}\left(t_v^{-1}\left(u_1\right),...,t_v^{-1}\left(u_d\right)\right)$$

and

$$\boldsymbol{c}^{t}\left(u_1,...,u_d\right)=\frac{\mathbf{t}\left(t_v^{-1}\left(u_1\right),...,t_v^{-1}\left(u_d\right)\right)}{\prod_{i=1}^{d}t_v\left(t_v^{-1}(u_i)\right)}, \tag{1.6}$$

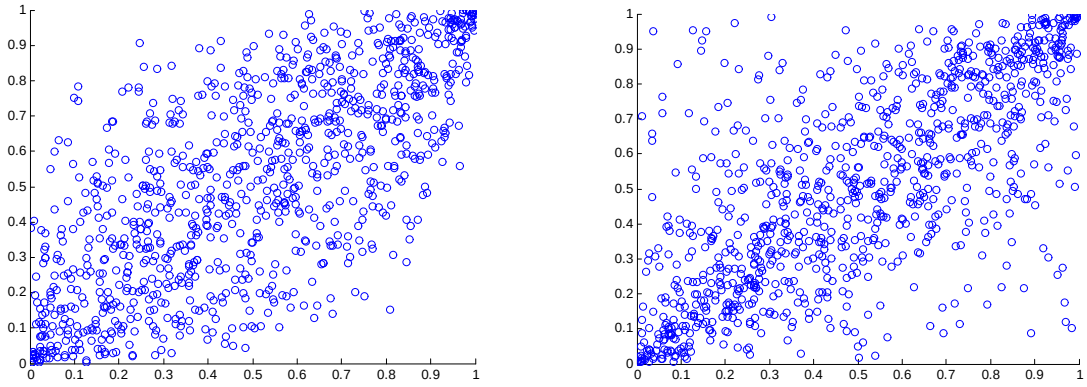


Figure 1.4: Scatterplots of 5000 simulated bivariate vectors from the Gaussian (left panel) and t copula with equal correlations

The Gaussian and t copula are constructed by applying Sklar's theorem to the multivariate Gaussian and multivariate t distribution, respectively. Another way of constructing copulas is by the use of a *copula generator function* ψ. The copulas constructed this way are called *Archimedean copulas*

Definition 9 *(Archimedean Copulas) An Archimedean copula is a function* $\boldsymbol{C}$ *from* $[0,1]^d \longrightarrow [0,1]$ *given by:*

$$\boldsymbol{C}(u_1, ..., u_d) = \psi^{[-1]}\left(\sum_{i=1}^{d} \psi(u_i)\right) \tag{1.7}$$

where ψ *(the generator function of* $\boldsymbol{C}$*) is a continuous, strictly decreasing convex function from* $[0,1]$ *to* $[0,+\infty)$ *such that* $\psi(1) = 0$ *and where* $\psi^{[-1]}$ *denotes the pseudo - inverse function of* ψ :

$$\psi^{[-1]}(t) = \begin{cases} \psi^{-1}(t) & for \;\; 0 \leqslant t < \psi(0) \\ 0 & for \;\; t \geqslant \psi(0) \end{cases}$$

When $\psi(0) = \infty$*,* ψ *and* $\boldsymbol{C}$ *are said to be strict (and* $\psi^{[-1]}(t) = \psi^{-1}(t)$*); when* $\psi(0) < \infty$ ψ *and* $\boldsymbol{C}$ *are non - strict. Furthermore,* $\boldsymbol{C}(u_1, ..., u_d) > 0$ *on* $(0,1]^d$ *if and only if* $\boldsymbol{C}$ *is strict.*

The term Archimedean copulas first appeared in Genest and Mackay (1995). It arises from the fact that the copulas constructed by the equation (1.7) follow the *Archimedean property* (Genest and Favre, 2007), which states that if u and v are real numbers in $(0,1)$ there exists an integer n such that $u^n < v$. Given a copula $\boldsymbol{C}$, the "$\boldsymbol{C} - multiplication$" $\circ$ is defined by $u \circ v = \boldsymbol{C}(u,v)$ .Then u^n is defined

recursively by $u^2 = u \circ u$ and $u^n = u^{n-1} \circ u$ [Note: Non Archimedean copulas also have this property, as long as $\boldsymbol{C}(u,u) < u$ for u in $(0,1)$.]

Example 10 *(Clayton Copula) If the generator function is* $\psi(t) = \frac{1}{\alpha}\left(t^{-a} - 1\right)$ *with* $a \in (0, +\infty)$ *the corresponding copula is called the Clayton copula with distribution*

$$\boldsymbol{C}^{Cl}\left(u_1, ..., u_d\right) = \max\left(\sum_{i=1}^{d} u_i^{-a} + 1, 0\right)$$

Example 11 *(Gumbel Copula) If the generator function is* $\psi(t) = (-\ln t)^a$ *with* $a \in (0, +\infty)$ *the corresponding copula is called the Gumbel copula with distribution*

$$\boldsymbol{C}^{Gu}\left(u_1, ..., u_d\right) = \exp\left(-\left(\sum_{i=1}^{d} (-\ln u_i)^a\right)^{1/a}\right)$$

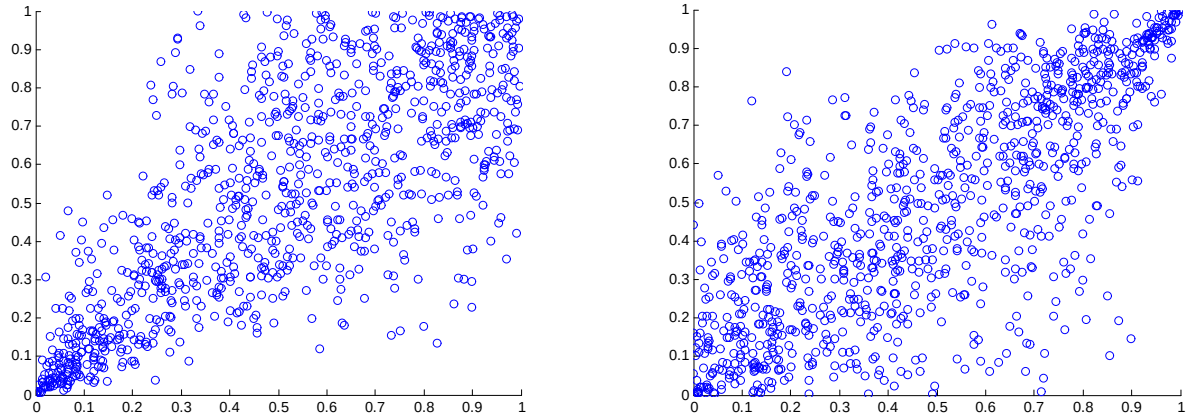


Figure 1.5: Scatterplots of 5000 simulated bivariate vectors from the Clayton (left panel) and Gumbel copula with equal association parameter

The parameter a, of the two aforementioned Archimedean copulas is called the

association parameter. For a bivariate random variable there is one to one correspondence between the copula and Kendall's tau τ.

$$\tau = 4\int_0^1\int_0^1 \boldsymbol{C}(u,v)\, dC(u,v) - 1 = 1 - 4\int_0^1\int_0^1 \frac{\partial \boldsymbol{C}(u,v)}{\partial u}\frac{\partial \boldsymbol{C}(u,v)}{\partial v} du dv$$

or

$$\tau = 1 + 4\int_0^1 \frac{\psi'(t)}{\psi(t)} dt$$

Further, the association parameter is directly linked to Kendall's tau. For the Clayton copula it holds that $\tau = \frac{\alpha}{\alpha+2}$ whereas for the Gumbel copula it holds that $\tau = \frac{a-1}{a}$. Since in both cases we observe that $\tau > 0$ these two copulas are suitable only for cases where there is positive dependence between the random variables. For the Gaussian and $t-$ copula there is one to one correspondence between Kendall's tau and linear correlation. If the bivariate vector (U, V) follows the Gaussian or $t-$ copula then:

$$\tau(U,V) = \frac{2}{\pi}\arcsin\left(\varrho(U,V)\right) \tag{1.8}$$

Equation (1.8) also holds in larger dimensions however there is no guaranty that the matrix $\mathcal{T} = (\tau_{ij})_{1\leqslant i,j\leqslant d}$ whose elements are obtained by Eq (1.8), is positive semi-definite. Archimedean copulas can describe many types of dependence nevertheless all of them are one or two parameter families, independent of the size of the random vector of interest. That is why in dimensions greater than two the association parameter has an unclear financial meaning and thus Archimedean copulas are used in

financial applications strictly for bivariate problems. Therefore in what follows we will consider only bivariate Archimedean copulas.

1.3 Tail dependence

In applications related to risk management, the tails of the returns distribution are of extreme importance, since the risk metric proposed by the Bassel accord for financial institutions (Value at Risk) is related to the quantile function of the returns. Analogously, in the multivariate setting, the accurate calculation of the probability of a joint extreme event (stock market crush or simultaneous default of the majority of loans, in a loan portfolio) is essential. A measure of this probability is the *tail dependence*. Let X and Y be two random variables with distributions F_X and F_Y. The limit (if it exists) of the conditional probability that Y is greater than the 100th - percentile of F_Y, given that X is greater than the 100th percentile of F_X is called *upper tail dependence* and is denoted as λ^U

$$\lambda^U = \lim_{u \longrightarrow 1^-} \Pr\left[Y > F_Y^{-1}(u) | X > F_X^{-1}(u)\right]$$

Similarly the limit of the conditional probability that Y is less than or equal to the 100th - percentile of F_Y, given that X is less than or equal to the 100th percentile of F_X is called *lower tail dependence* and is denoted as λ^L

$$\lambda^L = \lim_{u \longrightarrow 0^+} \Pr\left[Y \leqslant F_Y^{-1}(u) | X \leqslant F_X^{-1}(u)\right]$$

Theorem 12 *The upper and lower tail dependence coefficients depend only on the copula* $\boldsymbol{C}$*, of* X *and* Y *and not on the marginal distributions of* X *and* Y*. Further, it holds that:*

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

proof: (see Nelsen, 1999).

model	λ^U	λ^L
Gaussian Copula	0	0
t Copula	$2\bar{t}_{v+1}\left(\sqrt{(v+1)(1-\varrho)/(1+\varrho)}\right)$	λ^U
Clayton Copula	0	$2^{-1/a}$
Gumbel Copula	$2\text{-}2^{1/a}$	0

Table 1.1: Tail dependence coefficients for bivariate copula models

Table 1.1 summarizes the tail dependence coefficients of the copulas described in examples 7,8,10 and 11. For the Gaussian copula the correct expression for the tail dependence is: $\lambda^U = \lambda^L = \lim_{x \longrightarrow +\infty} 2\bar{\Phi}\left(x\sqrt{1-\varrho}/\sqrt{1+\varrho}\right)$ which equals zero for values of $\varrho < 1$.($\bar{\Phi} \equiv 1 - \Phi$, is the tail function of the standard normal distribution.) therefore the Gaussian copula (and the Gaussian distribution) assign zero probability to the occurrence of an extreme event. For the t copula the tail dependence coefficients are equal and depend on both the degree of freedom parameter v and the correlation coefficient ϱ. Table two contains values of the tail dependence coefficient for various values of v and ϱ. For the two Archimedean copulas, we observe that they are asymmetrical tail dependent ($\lambda^U \neq \lambda^L$) as opposed to the Gaussian and t copula that are symmetrically tail dependent. The Clayton copula exhibits only lower tail depen-

dence whereas the Gumbel copula exhibits only upper tail dependence. Figure four contains the plot of the tail dependence coefficient with respect to the association parameter, for the Clayton and the Gumbel copula.

$v\backslash\varrho$	-0.75	-0.25	0	0.3	0.8	1
2	0.020	0.111	0.182	0.293	0.604	1
4	0.002	0.034	0.076	0.162	0.490	1
9	0	0.002	0.010	0.043	0.317	1
20	0	0	0	0.003	0.141	1
50	0	0	0	0	0.021	1
200	0	0	0	0	0	1

Table 1.2: Tail dependence coefficient of the t copula for various values of v and ϱ

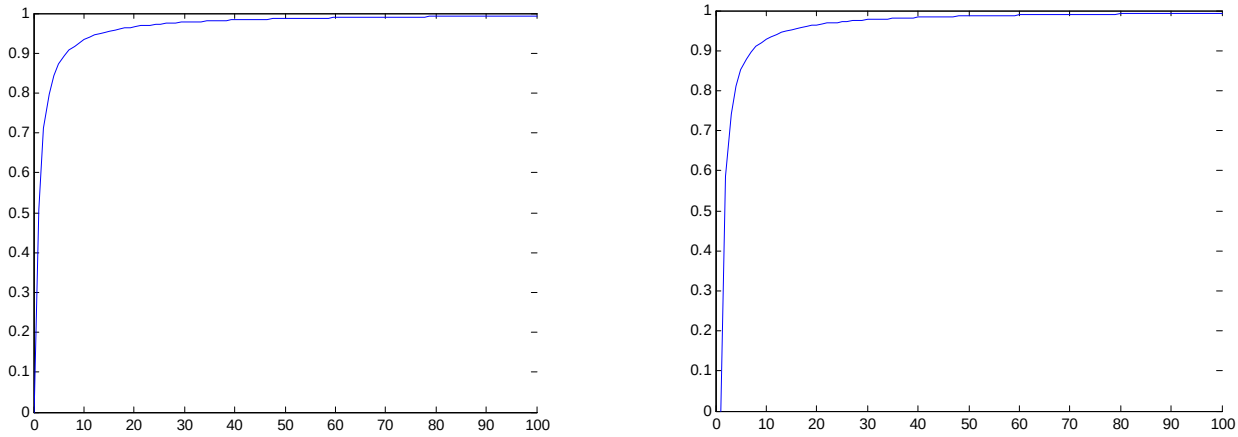


Figure 1.6: Tail dependence coefficient for the Clayton copula (left panel) and Gumbel copula (right panel) as a function of the association parameter, a.

Example 13 *(Joe Clayton and Symmetrized Joe Clayton Copulas) All copulas presented so far exhibit tail dependence symmetry or they exhibit tail independence in*

the upper or (and) the lower tail. A copula that is tail dependent in both upper and lower tail with $\lambda^U \neq \lambda^L$ *is the symmetrized Joe Clayton Copula with distribution*

$$\boldsymbol{C}^{SGC}(u,v) = \frac{1}{2}\left(\boldsymbol{C}^{JC}\left(u,v;\lambda^U,\lambda^L\right) + \boldsymbol{C}^{JC}\left(1-u,1-v;\lambda^L,\lambda^U\right) + u + v - 1\right)$$

where $\boldsymbol{C}^{JC}$ *is the distribution of the Joe Clayton copula defined as*

$$\boldsymbol{C}^{JC}(u,v) = 1 - \left(1 - \frac{1}{\left(\frac{1}{\left(1-(1-u)^k\right)^\gamma} + \frac{1}{\left(1-(1-v)^k\right)^\gamma} - 1\right)^{1/\gamma}}\right)^{1/k}$$

where the parameters k *and* γ *are related to* λ^U *and* λ^L *according to the following equations*

$$k = \frac{1}{\log_2\left(2-\lambda^U\right)} \text{ and } \gamma = -\frac{1}{\log_2 \lambda^L}$$

1.4 Inference on copulas

Let $\boldsymbol{X} = (X_1, ..., X_d)$ be the multivariate random vector of interest and $\boldsymbol{x}_t = \{(x_{1,t}, ..., x_{d,t})\}_{t=1}^n$ be a sample of n *iid* realizations of $\boldsymbol{X}$**.** The logarithm of equation (1.3)

$$\begin{aligned} LL_t(\boldsymbol{\alpha}_i, \boldsymbol{\vartheta}; \boldsymbol{x}_t) &= \log c\left(F_1\left(x_{1,t};\boldsymbol{a}_1\right), ..., F_d\left(x_{d,t};\boldsymbol{a}_d\right);\boldsymbol{\vartheta}\right) + \sum_{i=1}^{d} \log\left(f_i\left(x_{i,t};\boldsymbol{a}_i\right)\right) \\ LL_t(\boldsymbol{\alpha}_i, \boldsymbol{\vartheta}; \boldsymbol{x}_t) &= L_{c,t}\left(\boldsymbol{\alpha}_i, \boldsymbol{\vartheta}; \boldsymbol{x}_t\right) + L_{m,t}\left(\boldsymbol{\alpha}_i, \boldsymbol{x}_t\right) \end{aligned} \tag{1.9}$$

is the log - likelihood function for the $t-th$ observation of a copula based model, where $\boldsymbol{\alpha}_i$ is the vector of parameters of the $i-th$ margin and $\boldsymbol{\vartheta}$ is the vector of copula parameters. To estimate the parameters vector $\boldsymbol{\phi} = (\boldsymbol{a}_1', ..., \boldsymbol{a}_d', \boldsymbol{\vartheta}')'$ three methods have been proposed in the statistical literature, maximum likelihood (MLE), Inference function for margins (IFM) and the pseudo likelihood method (PML).

1.4.1 Maximum likelihood method

According to the maximum likelihood method the vector $\boldsymbol{\phi}$ is estimated by the function

$$\begin{aligned}
\mathcal{LL}(\boldsymbol{\alpha}_i, \boldsymbol{\vartheta}; \boldsymbol{x}_t) &= \sum_{t=1}^{n} LL_t(\boldsymbol{\alpha}_i, \boldsymbol{\vartheta}; \boldsymbol{x}_t) \qquad (1.10)\\
&= \sum_{t=1}^{n}\sum_{i=1}^{d} \log\left(f_i\left(x_{i,t}; \boldsymbol{a}_i\right)\right) + \sum_{t=1}^{n} \log \boldsymbol{c}\left(F_1\left(x_{1,t}; \boldsymbol{a}_1\right), ..., F_d\left(x_{d,t}; \boldsymbol{a}_d\right); \boldsymbol{\vartheta}\right)\\
&= \sum_{i=1}^{d} \mathcal{LL}_i\left(\boldsymbol{\alpha}_i; x_{t,i}\right) + \mathcal{LL}_c(\boldsymbol{\alpha}_i, \boldsymbol{\vartheta}; \boldsymbol{x}_t)\\
\hat{\boldsymbol{\phi}}_{ML} &= \arg\max_{\phi\in\Phi} \mathcal{LL}(\boldsymbol{\phi}; \boldsymbol{x}_t)
\end{aligned}$$

where f_i denotes the probability density function of the $i-$margin and $\boldsymbol{c}$ is the copula density. Under the usual regularity conditions of the asymptotic maximum likelihood theory, $\hat{\boldsymbol{\phi}}_{ML}$ is normally distributed and

$$\sqrt{n}\left(\hat{\boldsymbol{\phi}}_{ML} - \boldsymbol{\phi}\right) \longrightarrow N\left(0, \mathcal{I}^{-1}\right) \qquad (1.11)$$

where $\boldsymbol{\phi}$ is the true parameters vector and $\mathcal{I}$ denotes the *Fisher's information matrix*, which is the negative Hessian matrix of the function $\mathcal{LL}$. From equation (1.10) standard errors of the estimates can be constructed. If the size of the column vector $\boldsymbol{\phi}$ is $k \times 1$, the standard error of the $j-th$ element of $\hat{\boldsymbol{\phi}}_{ML}$, based on the finite sample of size n, is the $j-th$ element on the main diagonal of the Fisher's information matrix, divided by $\sqrt{n}$

$$st.Error(\hat{\boldsymbol{\phi}}_{ML}(j)) = \frac{1}{\sqrt{n}}\mathcal{I}^{-1}(j,j)$$

Let s_i be the size of the marginal parameters vector $\boldsymbol{\alpha}_i$ and s_ϑ be the size of the copula parameters vector [Note: All parameter vectors in this study are assumed to be column vectors therefore s_i and s_ϑ are actually the number of rows of the vectors $\boldsymbol{\alpha}_i$ and ϑ, respectively]. To estimate the vector $\boldsymbol{\phi}$ we need to solve the systems of equations of the log likelihood *score vectors*

$$\frac{\partial \mathcal{LL}}{\partial \boldsymbol{\phi}} = \mathbf{0} \text{ or } \quad vec\left(\frac{\partial \mathcal{LL}_1}{\partial \boldsymbol{a}_1}, \frac{\partial \mathcal{LL}_2}{\partial \boldsymbol{a}_2}, ..., \frac{\partial \mathcal{LL}_c}{\partial \boldsymbol{\phi}}\right) = \mathbf{0} \tag{1.12}$$

where

$$\frac{\partial \mathcal{LL}_i(\boldsymbol{\alpha}_i; x_{t,i})}{\partial \boldsymbol{a}_i} = \left(\frac{\partial \mathcal{LL}_i}{\partial a_{1i}}, \frac{\partial \mathcal{LL}_i}{\partial a_{2i}}, ..., \frac{\partial \mathcal{LL}_i}{\partial a_{s_i i}}\right)' \tag{1.13}$$

$$\frac{\partial \mathcal{LL}_c(\boldsymbol{\alpha}_i, \boldsymbol{\vartheta}; \boldsymbol{x}_t)}{\partial \boldsymbol{\phi}} = \left(\frac{\partial \mathcal{LL}_c}{\partial a_{11}}, ..., \frac{\partial \mathcal{LL}_c}{\partial a_{s_d d}}, \frac{\partial \mathcal{LL}_c}{\partial \vartheta_1}, ..., \frac{\partial \mathcal{LL}_c}{\partial \vartheta_{s_\vartheta}}\right)' \tag{1.14}$$

and vec is the operator that transforms a matrix to a column vector, for example if $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ then $vecA = \begin{bmatrix} 1 & 3 & 2 & 4 \end{bmatrix}'$. Therefore there are $2(s_1 + s_2 + s_d) + s_\vartheta$

equations based on the first derivatives of the log likelihood function that need to be solved. It is obvious that the number of equations grows with d, making very difficult to estimate such models in large dimensions. Further, the calculations of the analytical scores of the log likelihood can be very tedious, that is why solving the equations defined in (1.12) is a method to estimate $\boldsymbol{\phi}$ that is never used in practice. Instead numerical methods are usually employed. The term "numerical methods" covers all methods where the optimum vector is not obtained by solving systems of equations but by an iterative process based on numerical approximations of the log likelihood gradient vector and Hessian matrix. A hybrid method of the two maintains the iterative scheme of the numerical methods but uses analytical expressions for the gradient vector and/or the Hessian matrix of the log likelihood. The analytical derivatives based method is more efficient when compared to numerically based derivatives method, in both accuracy and speed of convergence, as it was advocated in Hafner and Herwartz (2007) or in Diamantopoulos and Vrontos (2010), in a similar context. In what follows, when we refer to numerical optimization we mean the iterative process that uses numerical approximations of the gradient and Hessian, as opposed to the analytical optimization method where again an iterative process is employed but with analytical expressions for the log likelihood gradient and the Hessian. The iterative process that is usually employed is the so called BFGS algorithm (Goldfarb (1970) and Shanno (1970)), according to which the parameters vector $\hat{\boldsymbol{\phi}}^p$ obtained after $p-$ iterations of the algorithm, equals

$$\hat{\boldsymbol{\phi}}^p = \hat{\boldsymbol{\phi}}^{p-1} - \lambda H^{-1} \frac{\partial \mathcal{LL}}{\partial \boldsymbol{\phi}} \tag{1.15}$$

where λ is a scalar that controls the step size, H is some approximation of the Hessian, computed at $\hat{\boldsymbol{\phi}}^{p-1}$and defines the direction of the search. In case where the actual Hessian (numerical or analytical) is used, the BFGS method is known as *Newton - Raphson* (NR) method, while if H is approximated by the outer product of gradients

$$H = \left(\frac{\partial \mathcal{LL}}{\partial \boldsymbol{\phi}} \right)' \frac{\partial \mathcal{LL}}{\partial \boldsymbol{\phi}}$$

the method is known as *BHHH* method, from Berndt, Hall, Hall and Hausman (1970). For more information on optimization methods the interested reader is referred to Fletcher and Roger (1987) or Avriel and Mordecai (2003). Nevertheless, numerical methods (we are not aware of any attempt to estimate copulas with analytical methods, except of Liu and Luger (2009) who used analytical derivatives to test a novel estimation method called maximization by parts. However they used the analytical method to the simplest case possible, Gaussian copula with Gaussian margins.) speed and accuracy depend heavily on the size of the parameters vector. Therefore in empirical applications the maximum likelihood method is rarely used. Instead, two step methods, like the IFM or the PML methods are employed.

1.4.2 Inference functions for margins (IFM)

The IFM method was proposed by Joe (1996) and is a fully parametric method that consists of two steps. Its basic idea is instead of optimizing the log likelihood function ($\mathcal{LL}$, in equation (1.10)), to optimize $\mathcal{LL}_i$, $i = 1, 2, ..d$ independently from each other, in order to obtain estimates for the marginal parameters vectors $\tilde{\alpha}_i$ at

the first step and at the second step the function $\mathcal{LL}_c$ is optimized, conditioned upon the results from the first step. Therefore the score vector of the first step for the $i-$ margin is

$$\frac{\partial \mathcal{LL}_i\left(\boldsymbol{\alpha}_i; x_{t,i}\right)}{\partial \boldsymbol{a}_i} = \left(\frac{\partial \mathcal{LL}_i}{\partial a_{1i}}, \frac{\partial \mathcal{LL}_i}{\partial a_{2i}}, ..., \frac{\partial \mathcal{LL}_i}{\partial a_{s_i i}}\right)' \quad i = 1, 2, ..., d.$$

Note that the above expression is identical to equation(1.13) thus the marginal estimates of both methods are identical. At the second step the conditional score vector equations obtained by the copula log likelihood, are used to estimate the copula parameters

$$\frac{\partial \mathcal{LL}_c(\boldsymbol{\vartheta}; \boldsymbol{\alpha}_i, \boldsymbol{x}_t)}{\partial \boldsymbol{\vartheta}} = \left(\frac{\partial \mathcal{LL}_c}{\partial \vartheta_1}, ..., \frac{\partial \mathcal{LL}_c}{\partial \vartheta_{s_\vartheta}}\right)' \tag{1.16}$$

In other words, IFM method assumes working independence between the marginal parameters and the copula log - likelihood. The procedure that is used to estimate the parameters, with the IFM, is similar to that of the MLE method. One can solve the systems of equations defined by the log likelihood scores or use an iterative algorithm, as it is always done in practice. Let $\mathbf{g}_{k,n}$ be the score vectors defined in (1.13) and (1.16)

$$\mathbf{g}_{k,n} = vec\left(\frac{\partial \mathcal{LL}_i\left(\boldsymbol{\alpha}_i; x_{t,i}\right)}{\partial \boldsymbol{a}_i}, \frac{\partial \mathcal{LL}_c(\boldsymbol{\vartheta}; \boldsymbol{\alpha}_i, \boldsymbol{x}_t)}{\partial \boldsymbol{\vartheta}}\right)$$

Joe proved that the variance - covariance matrix $\tilde{V}_n$, of the parameters vector $\tilde{\boldsymbol{\phi}}$**,** based on the finite sample of size n, is given by the *Godambe Information matrix*,

defined as

$$\tilde{V}_n = H_n^{-1} J_n H_n^{-1}$$

where $H_n = \frac{\partial \mathbf{g}_{k,n}}{\partial \phi'} = \begin{bmatrix} \frac{\partial^2 \mathcal{LL}_i}{\partial \boldsymbol{a}_i \partial \boldsymbol{a}_j'} & \mathbf{0} \\ \mathbf{0} & \frac{\partial^2 \mathcal{LL}_c}{\partial \boldsymbol{\vartheta} \partial \boldsymbol{\vartheta}'} \end{bmatrix}$ and $J_n = \mathbf{g}_{k,n} \cdot \mathbf{g}_{k,n}'$.Van der Vaart (1998) argued that instead of the Godambe information matrix , the variance covariance matrix of the parameters can be obtained by applying one step of the Newton - Raphson algorithm to the full likelihood, using the IFM estimators. An alternative method to calculate the variance covariance matrix under the IFM framework is the *jackknife* method (Dias, 2004).

Proposition 14 *(Joe, 1996, page 302) Let* $\tilde{\boldsymbol{\phi}}^{(-t)}$, $t = 1, 2, ..., n$ *be the IFM estimate of* $\boldsymbol{\phi}$ *obtained from the observed sample, with the* t *- th observation excluded. The jackknife estimate of* $\tilde{V}_n$ *is*

$$\tilde{V}_n = n \sum_{t=1}^{n} \left(\tilde{\boldsymbol{\phi}}^{(-t)} - \tilde{\boldsymbol{\phi}} \right) \left(\tilde{\boldsymbol{\phi}}^{(-t)} - \tilde{\boldsymbol{\phi}} \right)'$$

Further, if h *is a real valued function, the standard errors of* $h\left(\tilde{\boldsymbol{\phi}}\right)$ *are given by*

$$\left(\sum_{t=1}^{n} \left(h\left(\tilde{\boldsymbol{\phi}}^{(-t)}\right) - h\left(\tilde{\boldsymbol{\phi}}\right) \right) \right)^{1/2}.$$

1.4.3 The pseudo likelihood method

As it was pointed out in subsection 1.1 the copula of a $d-$ dimensional random vector $\boldsymbol{X}$ is the joint distribution of the ranks associated with $\boldsymbol{X}$**.** In both MLE and IFM

methods the ranks of $\boldsymbol{X}$ are obtained for the probability integral transformation of $\boldsymbol{X}$**,** that is by assuming a marginal distribution F_i for X_i and calculating the ranks of X_i, for a sample of size n by $F_i\left(x_{i,t}\right)$, $i=1,..,d$ $t=1,...,n$. Therefore the correct choice of the marginal distribution F_i is crucial for the accuracy of the method, as noted for example in Kim et al. (2007). To avoid this potential pitfall, Genest and Rivest (1995) proposed to calculate the ranks of the data not with the parametric probability integral transformation but with the non - parametric empirical CDF function

$$F_n\left(x\right)=\frac{1}{n+1}\sum_{j=1}^{n}\mathcal{I}_{\{y\in\mathbb{R},y\leqslant x\}}\left(x_{ij}\right)$$

Thus, according to the pseudo likelihood method, the n observed $d-$ dimensional vectors $\boldsymbol{x}_i=\left(x_{1i},...,x_{di}\right)_{i=1}^{n}$ are transformed to ranks

$$\left(F_n\left(x_{1i}\right),...,F_n\left(x_{di}\right)\right)$$

and the copula parameters are estimated by optimizing the corresponding copula log likelihood

$$\mathcal{LL}_c(\boldsymbol{\vartheta};\boldsymbol{x}_t)=\sum_{t=1}^{n}\log c\left(F_n\left(x_{1,t}\right),...,F_d\left(x_{d,t}\right);\boldsymbol{\vartheta}\right) \tag{1.17}$$

Genest and Rivest proved the asymptotic normality of the parameters vector $\bar{\boldsymbol{\vartheta}}_{PL}$ obtained by optimizing Eq (1.17).

$$\bar{\boldsymbol{\vartheta}}_{PL}\sim N(\boldsymbol{\vartheta},\frac{v^2}{n})$$

A consistent estimator of the variance - covariance matrix of $\bar{\boldsymbol{\vartheta}}_{PL}$ is

$$\frac{\bar{v}_n^2}{n} = \frac{\bar{\sigma}_n^2}{n\bar{\beta}_n^2}$$

where

$$\bar{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}\left(M_i - \bar{M}\right)^2$$

and

$$\bar{\beta}_n^2 = \frac{1}{n}\sum_{i=1}^{n}\left(N_i - \bar{N}\right)^2$$

are sample variances computed from two sets of pseudo - observations. These pseudo observations are computed from a procedure described in Genest and Favre (2007) for a bivariate data set, as follows:

- **Step one**: Relabel the original data $(X_1, Y_1), ..., (X_n, Y_n)$, so as to $X_1 < .... < X_n$

- **Step two**: Write $\mathcal{LL}_c(\boldsymbol{\vartheta}; \boldsymbol{x}_t)$ as $L(\boldsymbol{\vartheta}, u_1, u_2)$ and calculate the derivatives of eq(1.17) with respect to $\boldsymbol{\vartheta}$, $u_1 = F_n(x_{1,t})$ and $u_2 = F_n(x_{2,t})$, denoted as $L_{\boldsymbol{\vartheta}}$, $L_{\boldsymbol{u}_1}$ and $L_{\boldsymbol{u}_2}$, respectively.

- **Step three**: For $i \in \{1, 2, ..., n\}$, set

$$N_i = L_{\boldsymbol{\vartheta}}\left(\bar{\boldsymbol{\vartheta}}_{PL}; \frac{i}{n+1}, S_{Yi}\right)$$

- **Step four**: For $i \in \{1, 2, ..., n\}$, define

$$\begin{aligned} M_i &= N_i - \frac{1}{n}\sum_{j=1}^{n} L_{\boldsymbol{\vartheta}}\left(\bar{\boldsymbol{\vartheta}}_{PL}; \frac{j}{n+1}, S_{Yj}\right) L_{\boldsymbol{u}_1}\left(\bar{\boldsymbol{\vartheta}}_{PL}; \frac{j}{n+1}, S_{Yj}\right) \\ &\quad -\frac{1}{n}\sum_{S_{Yj} \geqslant S_{Yi}} L_{\boldsymbol{\vartheta}}\left(\bar{\boldsymbol{\vartheta}}_{PL}; \frac{j}{n+1}, S_{Yj}\right) L_{\boldsymbol{u}_1}\left(\bar{\boldsymbol{\vartheta}}_{PL}; \frac{j}{n+1}, S_{Yj}\right) \end{aligned}$$

1.5 Goodness of fit tests for copulas

Let $\boldsymbol{U} = (U_1, ..., U_d)$ be a random vector with uniform $[0, 1]$ margins and suppose we have fitted a copula $\boldsymbol{C}_{\vartheta_n}$ to $\boldsymbol{U}$, based on a finite sample $\boldsymbol{u}_t = \{(u_{1,t}, ..., u_{d,t})\}_{t=1}^{n}$ of $\boldsymbol{U}$. A natural question that arises is how good is the fit of $\boldsymbol{C}_{\vartheta_n}$ to the data. To answer this section we will describe three goodness of fit tests for copulas; a graphical inspection method, applicable only in the bivariate case, a test for goodness of fit for the Gaussian and t copula, based on the squared radius of the corresponding joint distribution, that was developed by Kole et al. (2007) and a test suitable for Archimedean copulas that is based on the bivariate probability integral transform of the Archimedean copula, which was originally developed by Wang and Wells (2000) and generalized by Genest et al. (2006). For a more comprehensive presentation of goodness of fit tests for copula based models the interested reader is referred to Genest et al. (2009) who provides a review and a comparison of the various test used in the copula context.

1.5.1 Graphical inspection method

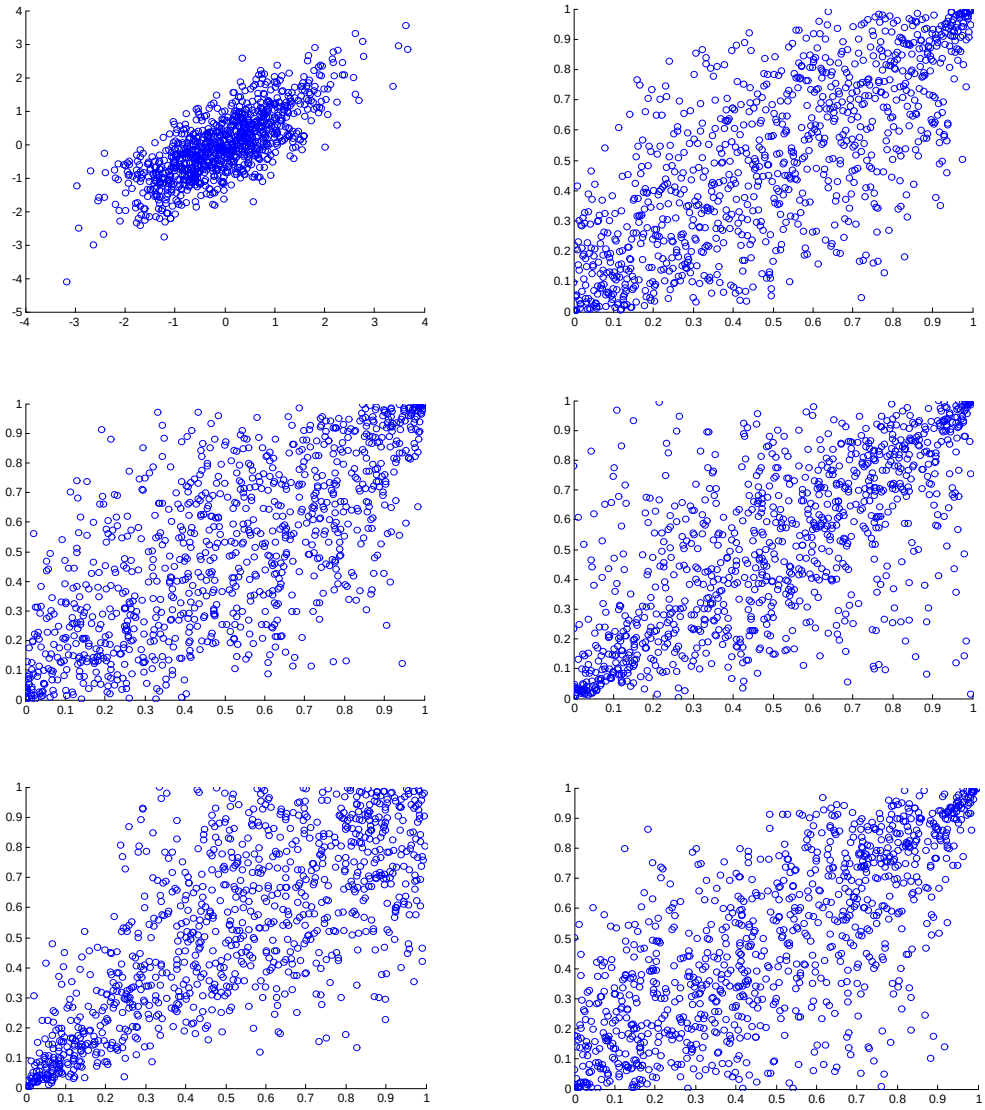


Figure 1.7: (From upper left to bottom right) Scatterplots of A) 1000 simulated vectors from a Gaussian distribution with zero mean and correlation $\varrho = 0.75$ B) (upper right panel) the ranks corresponding to the simulated data set C) 1000 simulated vectors from a Gaussian copula with $\varrho = 0.75$ D) 1000 simulated vectors from a t copula with $\varrho = 0.75$ and $v = 3$ E) 1000 simulated

vectors from a Clayton copula with $a = 2.35$ and F) 1000 simulated vectors from a Gumbel copula with $a = 2.17$

Let $\boldsymbol{x}_t = \{(x_{1,t}, x_{2,t})\}_{t=1}^n$ and $\boldsymbol{u}_t = \{(u_{1,t}, u_{2,t})\}_{t=1}^n$ be a sample of the initial random vector $\boldsymbol{X}$ and its corresponding ranks, calculated by the probability integral transform or by the empirical CDF function. The graphical inspection method compares the scatterplots of $\boldsymbol{u}_t$ and $\tilde{\boldsymbol{u}}_t$, where $\tilde{\boldsymbol{u}}_t$ is a sample of size n, simulated by the copula model assumed for $\boldsymbol{u}_t$. The copula whose simulated values have scatterplot that best resembles the scatterplot of $\boldsymbol{u}_t$ is the copula with the best fit. In figure (1.7) the graphical inspection method is illustrated. At the top left panel there is the scatterplot of 1000 bivariate vectors drawn from a multivariate Gaussian distribution with correlation coefficient $\varrho = 0.75$ and at the top right panel the scatterplot of the ranks of the simulated data set is drawn. The other four scatterplots of figure seven are scatterplots from a) A Gaussian copula with $\varrho = 0.75$, b) a t copula with $\varrho = 0.75$ and degree of freedom parameter $v = 3$, c) A Clayton copula with association parameter $a = 2.35$ and d) A Gumbel copula with association parameter $a = 2.17$ (Note: the values of α, for the Clayton and Gumbel copula were chosen to resemble the case of $\varrho = 0.75$). It is clear that the scatterplot that best resembles the one of the ranks (upper right) is the middle left that correspond to the Gaussian copula. The shortcomings of the graphical inspection are obvious; it can be used only for bivariate data sets and it does not quantify the goodness of fit with a valid statistical test but only visualizes the dependence. Further, since this method is based on simulations from a given parametric copula, it depends on whether the

inverse copula distribution exists in closed form

1.5.2 Squared radius method

This goodness of fit test can be seen as a variaton of the Kolmogorov - Smirnoff and of the Anderson Darling test, specifically designed to examine the Gaussian and t copula fit. Let $\boldsymbol{u}_t$ be a sample of size n, drawn from a random vector $\boldsymbol{U}$ with uniform margins. If we denote as $\boldsymbol{C}_H$ the hypothesized copula and as $\boldsymbol{C}_{E,n}$ the empirical copula of $\boldsymbol{U}$ (Deheuvels, 1979)

$$C_{E,n}\left(u_1,...,u_d\right)=\frac{1}{n}\sum_{t=1}^{n}\mathcal{I}\left(S_{Xi}\leqslant u_1,...,S_{Xd}\leqslant u_d\right)$$

where I is the indicator function defined as

$$I\left(\text{expression}\right)=\begin{cases}1, & \text{if expression is true}\\ 0, & \text{otherwise}\end{cases}$$

then the two tests take the form

$$\begin{aligned}D_{KS} &= \max_t\left|C_{E,n}\left(\boldsymbol{u}_t\right)-C_H\left(\boldsymbol{u}_t\right)\right|\\ D_{AD} &= \max_t\frac{\left|C_{E,n}\left(\boldsymbol{u}_t\right)-C_H\left(\boldsymbol{u}_t\right)\right|}{\sqrt{C_H\left(\boldsymbol{u}_t\right)\left(1-C_H\left(\boldsymbol{u}_t\right)\right)}}\end{aligned}$$

or, instead of these expressions that measure the largest distance (deviation) in the sample, one can use their averages as well

$$\begin{aligned} D_{KS}^{a} &= \int_{u} \left| C_{E,n}\left(\boldsymbol{u}_t\right) - C_H\left(\boldsymbol{u}_t\right) \right| dC_H\left(\boldsymbol{u}_t\right) \\ D_{AD}^{a} &= \int_{u} \frac{\left| C_{E,n}\left(\boldsymbol{u}_t\right) - C_H\left(\boldsymbol{u}_t\right) \right|}{\sqrt{C_H\left(\boldsymbol{u}_t\right)\left(1 - C_H\left(\boldsymbol{u}_t\right)\right)}} dC_H\left(\boldsymbol{u}_t\right) \end{aligned}$$

The distributions of the two statistics for the copula based model, under the null hypothesis that the specific copula accurately fits the corresponding observations, is non - standard, therefore costly computational simulations are needed to derive the critical values of the tests. Moreover as the dimension of the random vector increases, the calculation of the corresponding probabilities is a computationally demanding task. Instead, the authors suggest to transform the multivariate vectors of observations to their univariate squared radii, and thus transform the multivariate problem to an equivalent univariate one. The squared radii of the Gaussian and t copula are defined as follows.

For the random vector $\boldsymbol{U}$ whose dependence structure is described by a Gaussian Copula, with correlation matrix $\boldsymbol{\Sigma}_{\Phi}$ the squared radius is

$$Z^{\Phi} = \tilde{\boldsymbol{U}} \cdot \boldsymbol{\Sigma}_{\Phi}^{-1} \tilde{\boldsymbol{U}}'$$

where $\tilde{\boldsymbol{U}} = \left(\Phi^{-1}\left(U_1\right), ..., \Phi^{-1}\left(U_d\right)\right)$. The squared radius Z^{Φ} is a sum of d squared *iid* standard normal variables therefore follows the chi - square distribution with $d-$ degrees of freedom, $Z^{\Phi} \sim \chi_d^2$. If the dependence structure of $\boldsymbol{U}$ is described by a t

copula with correlation matrix $\mathbf{\Sigma}_t$ and degree of freedom parameter v, the squared radius equals

$$Z^{st} = \frac{\tilde{\boldsymbol{U}} \cdot \mathbf{\Sigma}_t^{-1} \tilde{\boldsymbol{U}}'}{d}$$

where $\tilde{\boldsymbol{U}} = \left(t_v^{-1}(U_1), ..., t_v^{-1}(U_d)\right)$. The variable Z^{st} follows the F distribution with d and v degrees of freedom. Thus in both cases the sample $\boldsymbol{u}_t$ of $\boldsymbol{U}$ of size $n \times d$ is transformed to a sample z_t of squared radii, of size $n \times 1$, with known hypothesized univariate distributions. The distance measures now become

$$D_{KS} = \max_t \left| F_{E,n}\left(z_t^i\right) - F_H\left(z_t^i\right) \right| \tag{1.18}$$

$$D_{AD} = \max_t \frac{\left| F_{E,n}\left(z_t^i\right) - C_H\left(z_t^i\right) \right|}{\sqrt{F_H\left(z_t^i\right)\left(1 - F_H\left(z_t^i\right)\right)}} \tag{1.19}$$

$$D_{KS}^a = \int_z \left| F_{E,n}\left(z_t^i\right) - F_H\left(z_t^i\right) \right| dF_H\left(z_t^i\right) \tag{1.20}$$

$$D_{AD}^a = \int_z \frac{\left| F_{E,n}\left(z_t^i\right) - F_H\left(z_t^i\right) \right|}{\sqrt{F_H\left(z_t^i\right)\left(1 - F_H\left(z_t^i\right)\right)}} dC_H\left(z_t^i\right) \tag{1.21}$$

where $i = \Phi$ or St, $F_{E,n}$ is the univariate empirical cdf and F_H is the chi square distribution with d degrees of freedom if $i = \Phi$ or F_H is the F distribution with d and v degrees of freedom, for the case $i = St$. Nevertheless, as in the previous case, the test statistics under the null do not follow a standard distribution therefore the critical values of the test are derived via simulations. For example if one want's to derive the critical values for the Gaussian copula he has to simulate N samples of size $n \times d$ with uniform margins and dependence structure described by the Gaussian

copula, transform each of the N samples to their corresponding squared radii, and calculate the distances defined in equations (1.18) to (1.21), N times. The critical values of the test are the empirical quantiles of the N simulated values of the distance measure of choice. For illustrative purposes the histograms of the distributions of the four aforementioned test statics, for the Gaussian copula, derived by N = 500 replications are presented in figure 1.8. The 2.5% and 97.5% empirical quantiles of these distributions are presented in table 1.3.

test\quantile	0.025	0.975
KS test	0.1727	0.2073
KSa test	0.1085	0.1541
AD test	0.4566	8.3222
ADa test	0,2831	0,4364

Table 1.3: Empirical 2.5% and 97.5% quantiles of the distributions of the four test statistics, for a bivariate Gaussian copula with correlation coefficient $\varrho = 0.75$ and sample size $n = 500$. Based on $N = 500$ replications

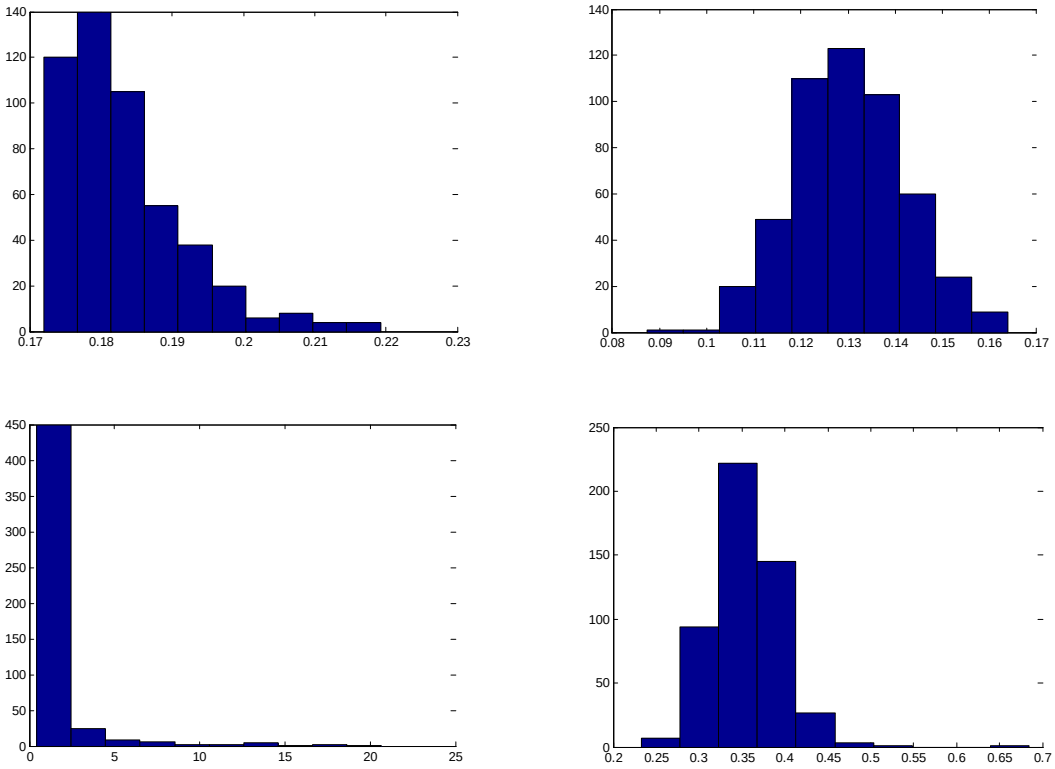


Figure 1.8: The empirical distribution of the Kolmogorov Smirnoff (upper left panel), average Kolmogorov Smirnoff (upper right panel), Anderson Darling (lower left panel) and average Anderson darling, for a bivariate Gaussian Copula with correlation coefficient $\varrho = 0.75$ and sample size $n = 500$. Based on $N = 500$ replications.

1.5.3 Bivariate probability integral transform method

Let $\boldsymbol{u}_t$ be a sample of size n, of the bivariate vector $\boldsymbol{U} = (U_1, U_2)$ and define the variables

$$W_i = \frac{1}{n}\sum_{j=1}^{n} I_{ij} \text{ where } I_{ij} = \begin{cases} 1, & \text{if } u_{1j} < u_{1i}, u_{2j} < u_{2i} \\ 0, & \text{otherwise} \end{cases}$$

and

$$W = \boldsymbol{C}_{\vartheta_n}(U, V)$$

where $\boldsymbol{C}_{\vartheta_n}$ is a copula from the Archimedean family and the pair (U, V) is drawn from $\boldsymbol{C}_{\vartheta_n}$.Further, define K_n as the empirical distribution of the variables W_i and K_{ϑ_n} as the theoretical distribution of W. Genest and Rivest called the function K_{ϑ_n} the *bivariate probability integral transform* and proved that $K_{\vartheta_n}(w) = w - \frac{\psi(w)}{\psi'(w)}$ where ψ is the generator function of the corresponding Archimedean copula. The first derivative of $K_{\vartheta_n}(w)$ with respect to w is denoted as $k_{\vartheta_n}(w)$. Genest et al.(2006) proposed two statistics $\mathcal{S}_n$ and $\mathcal{T}_n$to test if $\boldsymbol{C}_{\vartheta_n}$ fits the data well, where

$$\mathcal{S}_n = \int_0^1 n \left| K_n(w) - K_{\vartheta_n}(w) \right|^2 k_{\vartheta_n}(w)\, dw$$

and

$$\mathcal{T}_n = \sup_{0 \leqslant w \leqslant 1} \left| \sqrt{n} \left(K_n(w) - K_{\vartheta_n}(w) \right) \right|$$

or, the equivalent, simpler to calculate forms

$$\begin{aligned}\mathcal{S}_n &= \frac{n}{3} + n\sum_{j=1}^{n-1} K_n^2\left(\frac{j}{n}\right)\left[K_{\vartheta_n}\left(\frac{j+1}{n}\right) - K_{\vartheta_n}\left(\frac{j}{n}\right)\right] \\ &\quad -n\sum_{j=1}^{n-1} K_n\left(\frac{j}{n}\right)\left[K_{\vartheta_n}^2\left(\frac{j+1}{n}\right) - K_{\vartheta_n}^2\left(\frac{j}{n}\right)\right] \end{aligned} \tag{1.22}$$

and

$$\mathcal{T}_n = \sqrt{n}\max_{0\leqslant j\leqslant n-1}\left\{\left|K_n\left(\frac{j}{n}\right) - K_{\vartheta_n}\left(\frac{j+1}{n}\right)\right|, \left|K_n\left(\frac{j}{n}\right) - K_{\vartheta_n}\left(\frac{j}{n}\right)\right|\right\} \tag{1.23}$$

In table 1.4, the expressions of the functions K_ϑ and k_ϑ for the Clayton and Gumbel copulas, are illustrated.

copula\function	$K_\vartheta(w)$	$k_\vartheta(w) = dK_\vartheta(w)/dw$
Clayton copula	$w + \frac{w\left(1-w^\vartheta\right)}{\vartheta}$	$1 + \frac{1-w^\vartheta - w^{\vartheta+1}\log\vartheta}{\vartheta}$
Gumbel copula	$w - (1-\vartheta)\, w \log w$	1-$(1-\vartheta)(1+\log\vartheta)$

Table 1.4: Expressions for the bivariate probability integral transform and its derivative for the Clayton and Gumbel copulas.

Again, both tests do not follow any standard distribution and the critical values associated with these tests are derived by a bootstrap procedure, described in the following steps

- Step one: Estimate the parameter ϑ_n of the copula, based on the sample of size $n \times 2$

- Step two: Simulate N random samples drawn from $\boldsymbol{C}$, of size $n \times 2$.For each sample calculate the corresponding value of the test statistic of interest

- Place the N values of the test statistic, derived in step two in ascending order. The value of the test statistic that is located at the position $[(1-a)N]$ is the $a\%$ critical value.

1.6 Conditional Copulas

Until so far, the dependence parameter of the copula was assumed to remain the same over the whole sample period. However it has become a stylized fact that dependence (usually measured by Pearson's correlation in the standard financial setting) is known to change over time, since, in general, economic variables depend on their past observations. The time varying nature of correlations was first documented by Longin and Solnik (1995) where it was proved that correlation of the excess returns of seven major markets are time varying and also that correlations tend to increase in high volatility periods. This gave birth to a new class of multivariate GARCH models, known as *correlation models*, where the correlations assumed to be time varying, like the DCC model of Engle (2002) or the TVC model of Tse and Tsui (2002). Initially, correlation models were estimated under the joint normality assumption, that fails to account for some of the characteristics of the financial time series, like tail dependence. Therefore these models were extended to copula based models, where the dependency parameter is allowed to change over time, conditioned

upon the set of past information. This extension is due to Patton (2006) who proved that Sklar's theorem, its inverse and the invariance property hold also in the case that each margin is *conditionally* uniformly distributed

$$\boldsymbol{F}_t\left(\boldsymbol{x}|\mathcal{F}_{t-1}\right)=\boldsymbol{C}_t\left(F_{1,t}\left(x_1|\mathcal{F}_{t-1}\right),...,F_{d,t}\left(x_d|\mathcal{F}_{t-1}\right)|\mathcal{F}_{t-1}\right) \tag{1.24}$$

where $X_i|\mathcal{F}_{t-1}\sim F_{i,t}$ and $\boldsymbol{C}_t$ is the conditional copula of $\boldsymbol{X}_t$ given $\mathcal{F}_{t-1}$. Note that in equation (1.24), which is the extension of Sklar's theorem to conditional variables, the information set $\mathcal{F}_{t-1}$ is the same for the margins and the copula. Fermanian and Wegkamp (2004) and Fermanian and Scaillet (2005) considered the implications of a failure to use the same information set and concluded that in this case, the function $\boldsymbol{F}_t$ might not be a valid joint distribution. In financial applications however, it is a common practice to assume an information subset $\mathcal{F}_{i,t-1}$ such that $X_i|\mathcal{F}_{i,t-1}\overset{\mathcal{D}}{=}X_i|\mathcal{F}_{t-1}$, and thus being able to use $\mathcal{F}_{i,t-1}$ for the $i-$margin and $\mathcal{F}_{t-1}$ for the copula. Panchenko (2005) describes a method on how to reduce the size of the information set and avoid this potential pitfall. The rest of this section aims to present the most important specifications for the time varying dependency parameter, introduced in the financial literature.

1.6.1 Specifications for the dependence parameters

When the time varying parameter is the correlation ϱ_t like in a bivariate (note: Patton's specification can be used only in the bivariate case) Gaussian, Clayton or t copula, Patton (2006) proposed a specification that is a function of lagged data and

autoregressive correlation terms

$$\varrho_t = \Lambda_1 \left(\omega + \alpha \frac{1}{m} \sum_{i=1}^{m} \Phi^{-1}\left(u_{1,t-1}\right) \cdot \Phi^{-1}\left(u_{2,t-1}\right) + \beta \Lambda_1^{-1}\left(\varrho_{t-1}\right) \right)$$

where $\Lambda_1(x) = \frac{1-\exp(-x)}{1+\exp(-x)}$ is an increasing function that holds ϱ_t in $(-1,1)$, Φ^{-1} is the standard Gaussian quantile function and m is an arbitrary positive integer. The rational of the model is that when the variables move together (positive dependence) then the $z_i = \Phi^{-1}\left(u_{i,t-1}\right)$ have the same sign and if $\alpha > 0$, the value of ϱ_t increases. Therefore in the case of positive dependence α should be positive. In the Archimedean copula case, when the association parameter is not linked to correlations, Patton proposed a slightly different equation for the dependence parameter ϑ_t

$$\vartheta_t = \Lambda_1 \left(\omega + \alpha \frac{1}{m} \sum_{i=1}^{m} \left| u_{1,t-1} - u_{2,t-1} \right| + \beta \Lambda_1^{-1}\left(\varrho_{t-1}\right) \right)$$

where $\Lambda_2(x) = (1+\exp(-x))^{-1}$ when ϑ_t is the tail dependence parameter, like in the SJC copula, or $\Lambda_2(x) = \exp(x)$ for the Clayton copula parameter and $\Lambda_2(x) = (1+\exp(x))$ for the Gumbel copula. In a similar context, Creal et al. (2008) proposed a unifying framework for the evolution of the time varying parameter ϑ_t

$$\vartheta_t = \omega + \sum_{j=1}^{q} \beta_j \vartheta_{t-j} + \sum_{i=0}^{p-1} \alpha_i \dot{f}_{t-i}$$

where $\dot{f}_t = S_{t-1} \cdot \nabla_{\vartheta_t} \mathcal{LL}$ is the scaled gradient ($\nabla_{\vartheta_t} \mathcal{LL}$) of the corresponding log likelihood with respect to its parameters, times a scaling matrix $.S_{t-1}$, which is ap-

proximated by the Fisher's information matrix. When compared to Patton's specifications, the model proposed by Creal et al is more sensitive to extreme values, it can be used in models with dimension $d > 2$ and the estimated parameters are closer to the true parameter values but with higher variance. Other specifications for the time varying correlations are based directly on the specifications used in the multivariate GARCH framework. For example Jondeau and Rockinger (2006) proposed that the time varying correlations of a Gaussian or a t - copula follow the TVC model of Tse and Tsui (2002)

$$\varrho_t = (1 - a - \beta)\,\rho + a\xi_{t-1} + \beta\varrho_{t-1}$$

where ρ is the sample correlation and ξ_t is the sample correlation of a moving window of arbitrary size m of the variables z_i, defined as $z_i = \Phi^{-1}(u_{i,t-1})$ for the gaussian copula or $z_i = t_v^{-1}(u_{i,t-1})$ for the t copula, with Φ^{-1} and t_v^{-1} be the quantile functions of the standard normal and student t distribution with v degrees of freedom, respectively. In a similar context, Serban et al (2007) used the DCC equation proposed by Engle (2002)

$$\begin{aligned} \varrho_t &= diagQ_t^{-1/2} \cdot Q_t \cdot diagQ_t^{-1/2} \\ Q_t &= Q(1 - a - \beta) + a\boldsymbol{z}_{t-1} \cdot \boldsymbol{z}'_{t-1} + \beta Q_{t-1} \end{aligned}$$

where $\boldsymbol{z}_{t-1} = (z_{1,t}, ..., z_{d,t})$ defined earlier and Q is the sample covariance of $\boldsymbol{z}_t$. Both specifications ensure the positive definiteness of the corresponding correlation matrix and can be used in arbitrary dimensions. Other specifications include the stochastic

autoregressive copulas, proposed by Hafner and Manner (2008) where the dependence parameter ϑ_t is a function of a latent stochastic process λ_t, similar to the spirit of the stochastic volatility models

$$\begin{aligned} \vartheta_t &= g(\lambda_t) \\ \lambda_t &= \omega + \beta\lambda_{t-1} + \sigma\eta_t, \quad \eta_t \overset{iid}{\sim} N(0,1) \end{aligned}$$

where g is an appropriate function to keep the parameter ϑ_t in the desired domain. Finally, some authors, instead of assuming that the copula parameter ϑ_t is time varying they assumed that there are k states in the world and the dependence at each stage is different, by assuming that the dependence parameter is constant within states but changes from a state to another, as in Pelletier (2006) or by assuming a copula mixture where the mixture parameter is state dependent, as in Rodriguez (2007), or by assuming the functional form of the copula is state dependent and therefore the dependence structure at each state, is describe by a different copula as in Chollete et al (2008). All the aforementioned models assume that the latent variable k_t taking values $1, 2, .., k$ dependent on the current state, follows a Markov chain of order one with $\pi_{ij,t}$ the probability of moving to regime j in period t, conditional on being in the i regime at $t-1$.For the case when $k = 2$, the transition matrix is

$$P = \begin{bmatrix} \pi_{11} & 1-\pi_{11} \\ 1-\pi_{22} & \pi_{22} \end{bmatrix}$$

now let $\hat{\xi}_{t|t}$ be the vector of probabilities of being in each state in the $t-$th period, given the information at time t. Then

$$\begin{aligned}
\hat{\xi}_{t|t} &= \frac{\hat{\xi}_{t|t-1} \odot \eta_t}{\mathbf{1}'\hat{\xi}_{t|t-1} \odot \eta_t} \\
\hat{\xi}_{t+1|t} &= P'\hat{\xi}_{t|t} \\
\mathcal{LL}_C &= \sum_{t=1}^{n} \log\left(\mathbf{1}'\hat{\xi}_{t|t-1} \odot \eta_t\right)
\end{aligned}$$

where n is the sample size, $\mathbf{1}'$ is a vector of ones and η_t is the regime dependent quantity of interest. The estimation procedure proposed for the above class of models is called the expectation maximization (EM) algorithm and it is described in Hamilton (1994). There are also other approaches proposed in the literature like ones that assume structural breaks as in Dias and Embrechts (2004), the Local change point method of Mercurrio and Spokoiny (2004) or the Adaptive estimation method of Giacomini et al. For a review of these approaches the interested reader is referred to Manner and Reznikova (2009).

Part II

The second part - Empirical application

Chapter 2

Introduction

In this section I will try to quantify the dependence structure between some large companys from the financial sector, listed in the S&P500 index and the index itself. This is a fundamental problem in various areas of finance like risk management and derivatives pricing since, both the risk (measured by the VaR of the CVaR of the portfolio) and the price of a derivative written on a portfolio of assets, depend on the distribution of the portfolio. Furthermore, my model can be seen as a time varying, non normal version of the CAPM. The second part of the thesis is organized as follows: Section (2.1) presents a literature review of the use of copulas and tries to explain why copulas have become so popular in finance, section (2.2) describes the proposed model in full detail and sections (2.3) to (2.5) compare the proposed with other models and discuss the results and implications of the proposed model from financial prespective. Section (2.6) concludes.

2.1 Literature review

In many financial applications, like portfolio selection, derivatives pricing and risk management, the majority of the models used in practice, assume that the returns distribution is multivariate normal and that the correlation among assets, which is the measure of dependence in the multivariate normality framework, is constant through time. For example, three of the most prominent financial applications, Markowitz's (1952) portfolio theory, portfolio Value at Risk (Morgan Stanley 1995) and Black Merton and Scholes (1973) derivatives pricing theory, assume multivariate normal returns and constant correlation. For more information the interested reader is referred to Dowd (2001) or Jorion (1997).

However these two assumptions gained much criticism over the last years, since they have failed to find any empirical support. Univariate financial time series share some common characteristics, like heteroscedasticity, asymmetry and fat tails. The univariate volatility structure is adequately described by the generalized autoregressive conditional heteroscedasticity (GARCH) models of Bollerslev (1986) and its extensions like the EGARCH model of Nelson (1991) and the asymmetric GARCH model (GJR) of Glosten et al (1993), especially when these models are combined with flexible distributions like the Skew-T distribution of Hansen (1994), as in Jondeau and Rockinger (2006) or the SU-Normal distribution of Johnson's (1949) system as in Choi and Nam (2008). The interested reader in univariate GARCH models and their extensions is referred to Terasvirta (2006) or Tsay (2002) and the references therein. In the multivariate framework financial time series share some common

characteristics, like the tail dependence. Two series X and Y with distributions F_X and F_Y, are said to be tail dependent if the probability of an extreme co movement is greater than zero. The measure of tail dependence, is the tail dependnece coefficient

$$\lambda^L = \lim_{u \longrightarrow 0^+} \Pr\left[Y \leqslant F_Y^{-1}(u) | X \leqslant F_X^{-1}(u)\right]$$

For the lower tail, or:

$$\lambda^U = \lim_{u \longrightarrow 1^-} \Pr\left[Y > F_Y^{-1}(u) | X > F_X^{-1}(u)\right]$$

For the upper tail.

Of course the definition of tail dependence can be extended to arbitrary dimensions. In the same paper Embrechts, Mc Neil and Strautman prove that the multivariate normal distribution does not exhibit any tail dependence, except of the extreme case where $\varrho = 1$. This means that under plausible conditions, the multivariate normal distribution assigns zero probability to extreme co movements therefore is not able to describe financial returns. In a similar context, financial returns proved to be not only tail dependent but asymmetric as well. By using extreme value theory to model the tails of a multivariate distribution, Longin and Solnik (2001) showed that correlations tend to increase in bear markets and decrease in bull markets, and thus rejected the multivariate normality hypothesis. They proposed a measure of

this asymmetry named exceedance correlation defined as

$$Ex\_Corr\,(X,Y,\vartheta,\phi) = \begin{cases} Corr(X,Y|X \leq \vartheta, Y \leq \phi) & \text{for } \vartheta,\phi \leq 0 \\ Corr(X,Y|X \geq \vartheta, Y \geq \phi) & \text{for } \vartheta,\phi \geq 0 \end{cases}$$

and they proved that, asymptotically, exceedance correlation is zero for large positive returns and strictly positive for large negative returns. In a similar setting, Ang and Chen (2002) showed that correlations between stocks are much greater in downside moves than in upside moves. The existence of such stylized facts tends to reject the normality hypothesis. Furthermore the constant correlations hypothesis among financial returns failed to find empirical support, as in Engle and Sheppard (2001) who proposed a test that failed to find support of constant correlation in S&P500 portfolios. Constant correlation tests have also been proposed from Tse (2000) and Berra and Kim (2001). Again, these test reviled that correlations among assets tends to be time – varying.

To model this time varying dependence among financial returns, GARCH models were extended to the multivariate setting. In the multivariate GARCH family belong the vech model of Bollerslev, Engle and Wooldridge (1988), and its extension, the diagonal vech however the vast number of estimated parameters and the complexity of the conditions needed to ensure the positive – definiteness of the conditional covariance matrix make these models impractical, even for small dimensions. In order to eliminate the problem of ensuring positive definiteness Engle and Kroner (1995) proposed the BEKK model, which, in its general form, is also difficult to estimate,

since the number of estimated parameters is $O(k^4)$, where k is the dimension. In order to minimize the number of estimated parameters and therefore the computational burden, Alexander (2000) proposed the orthogonal GARCH model, which gave birth to a whole new class of models named factor GARCH models. The model proposed by Alexander is a linear combination of univariate GARCH models under a factor analysis framework. This model suffers from two limitations. It does not work well in cases where the series are not highly correlated and it is difficult to interpret the factors from a financial point of view. The multivariate GARCH model that gained most of the attention and is now considered as the benchmark model is the dynamic conditional correlation (DCC) model of Engle and Sheppard (2001). The main advantage of DCC is that allows a two step estimation of the parameters, since it decomposes correlation to variances and covariances. At the first step, working independence among margins is assumed and the conditional variances are estimated by univariate GARCH models. At the second step, the dynamic correlation is estimated, conditional on the results of the first step. Furthermore, DCC ensures that the estimated correlation matrices are positive definite under very mild conditions and the number of the estimated parameters of the second step can be reduced down to two, independent to the dimension of the series, therefore it can be estimated in problems of very large dimensions . Silvennoinen and Terasvirta (2008) or Bauwens et al (2006) give a complete overview of multivariate GARCH models.

The main criticism on multivariate GARCH models is that they estimated assuming multivariate normality, although Bollerslev and Wooldridge (1992) showed that the maximum likelihood estimations, assuming normality, are consistent, given that

both the conditional mean and conditional variance are specified correctly. Multivariate GARCH models that make non – normal distributional assumptions for the residuals can be found in Pelagatti (2004) or Fiorentini et al (2003) where the Student-T distribution is assumed or in Bauens and Laurent (2002) where the innovations follow skew-T distribution. Nevertheless the use of any standard distribution seldom fits the data well. If a returns process follows the Student –T distribution with v degrees of freedom, each margin follows the Student-T distribution with v degrees of freedom and the joint distribution is also T with the same degrees of freedom, an assumption that was found to be too restrictive in empirical applications.

The need for more flexible distributions that were able to better capture the characteristics of financial distributions led the researchers to the use of copulas. Strictly speaking a Copula is a d dimensional distribution function on $[0,1]^d$ with standard uniform margins. According to Sclar's theorem (1959), if a multivariate distribution has continuous margins there exists a unique copula that governs the dependence structure, implied by the distribution. For example the t-copula represents the dependence structure implied by the multivariate Student –T distribution. The main advantage of copulas is that the joint distribution can be factored to the dependence (the copula) and the margins. The marginal distributions are estimated separately from the dependence, which is governed by the copula. Patton (2002) extended the copula theory to cover the case of conditional multivariate distributions, allowing the construction of flexible joint conditional distributions. The combination of the conditional copulas and multivariate GARCH gave birth to a new class of models named copula GARCH (CGARCH) models, which gained considerable popularity

the past few years. For example Jondeau and Rockinger (2006) apply this methodology to investigate the dependence between daily stock index returns. Serban et al (2006) investigate the performance of a copula DCC model (a model with dependence implied by the T-Copula and correlation dynamics implied by the DCC) and they find it to be more efficient that multivariate GARCH models. In a similar setting Fantazzini used copula GARCH models to estimate the value at risk (VaR) of various index portfolios, Patton (2006) investigates the asymmetric dependence between exchange rates, while Ausin and Lopes (2006) used this setting to estimate time varying multivariate distributions in a Bayesian framework.

Copulas are now considered to be the "industry standard" tool to quantify dependence and are used in many financial applications. In the risk management framework, recent implementations of copula theory can be found at Wang et al (2010) where the VaR and CVaR of exchange rate portfolios are calculated for a basket of four international currencies, namely EUR, USD, JPY and HKD, or in Barges et al (2009)where the problem of capital allocation under tail value at risk (TVaR, Artzner et al 1999) is addressed. In a similar setting Perignon and Smith (2010) study whether the diversification effect is underestimated as a possible explanation to abnormal high levels of risk reported by international banks and its implications to reserve capital requirements of a banking institution Finally Huang et al (2009) propose a novel method to calculate the VaR of a portfolio based on copula GARCH models and showed that the t - copula based estimations outperform classical approach in terms of VaR violations in all confidence levels. Copulas are also used to quantify other sources of risk than market risk, like credit risk, as in Cousin and

Laurent (2008) or operational risk as in Chavez - Demoulin et al (2006).

In derivatives pricing Goorbergh et al (2005) is among the first who utilized copula theory to price options written on S&P500 and Nazdaq indices. They conclude that option prices implied by copula models differ significantly from those implied by the standard Gaussian assumption. Similar results are obtained in Zhang and Guegan (2008) where they conclude that prices obtained from a time varying t copula with GARCH marginal processes are the optimum way to calculate prices for derivatives, among their competing models. The pricing of derivatives issued on more than two assets is the problem addressed in Bedento et al (2010) where options written on a basket of five UK shares are studied. Their analysis reveals that results from copula models differ significantly from the standard gaussian approaches only in volatile but not in tranquil periods. Other authors, like Malo (2009) investigate the dependence between spot and future electricity prices using Markov switching multifractal models combined with copulas while Geman and Kharoubi (2009) study the assumption of negatively correlated equity and commodity returns with copula models and find that this assumption is mostly due to the Gaussian assumption previously used to quantify dependencies. Finally Chen et al (2008) uses copula to study the dependence structure between CDS's (Credit Default Swap's) and the kurtosis of equity returns and observed that lower rating classes display positive and asymmetric dependence structure unlike higher rating classes whose dependence structure seems to be almost symmetric.

Another research area where copulas are extensively used in financial contagion and the modern global financial crisis in general. The term financial contagion is used

to describe the interdependence between financial market and how a crisis originated in some market can spread out to affect other markets as well. One of the first studies of financial contagion is in Rodriguez (2007) who used copula models in a regime switching setup to study the changing dependence between financial market in periods of turmoil. In the same setting Aloiu et al (2010) and Kenourgios etal (2010) examine the effect of the US and UK markets to those of BRIC countries (Brasil, Russia India China) markets and found strong evidence of contagion.

2.2 Data and methodology

Let $R_t = (r_m, r_s)$ be a bivariate random vector of log - returns of a stock (r_s) traded in a market, and the returns of the market (r_m). The measure of market risk that the stock berries is known as *beta* of the the stock and can be referred to as a measure of the sensitivity of the asset's returns to market returns. Beta is estimated with the *market model*

$$r_s - r_f = a + b\,(r_m - r_f) + \epsilon \tag{2.1}$$

where r_f is the risk free rate. The expectations on the market model derive the model known as CAPM (Jack Treynor (1961, 1962), William Sharpe (1964), John Lintner (1965a,b) and Jan Mossin (1966))

$$E(r_s) = r_f + b \cdot E(r_m - r_f)$$

CAPM is one of the most extensively used models in finance. It holds that

$$b = \frac{cov(r_m, r_s)}{var(r_m)} = corr(r_m, r_s) \cdot \sqrt{\frac{var(r_s)}{var(r_{m)}}} \tag{2.2}$$

therefore, in order to estimate betas of stock the correlation between the stock and the market and the variance of the stock need to be estimated in the highest possible accuracy. Stock betas can be obtained by estimating the regression implied in (2.1) however this approach has two serious limitations

- It is based on the normality assumption.
- It provides static (not time varying) beta estimates

In section (2.1) we saw that both of these assumptions are too restrictive in real life situations. Instead we propose a different model based on copula functions, named t copula GARCH model (tCGARCH). My model assumptions are:

- The returns of both the stock and the market exhibit significant departures from Normality. More precisely, we assume that each series follows Hansen's Skew t distribution (Hansen, 1994), which can accommodate asymmetry and kurtosis in returns with time varying variances modelled with a GARCH type process.
- The dependence structure between each stock and the index is time varying and it is modeled with a t - copula with static degrees of freedom and time varying correlations that follow a specification similar to the DCC model of Engle.

Our assumptions are similar to those in Jondeau and Rockinger (2006) with only difference that time varying correlations are model with the TVC model in Jondeau and Rockinger[1] To estimate the betas of some major stock of the financial sector during the modern crisis period, three models where estimated: A simple market model implied by (2.1) a DCC GARCH model (that is a gaussian copula with gaussian margins) and a tCGARCH model described earlier. The first model provides static estimates for betas unlike the other two. The DCC model is used to estimate time varying betas under the wrong assumption of normality. It is worth noting that the tCGARCH model nest the DCC model, therefore both models can be compared with standard likelihood tests or with an information criterion like BIC or AIC. The aims of this empirical investigation are:

- To compare betas from the current period (2006 - 2010) to a less volatile previous period (2002 - 2006) to see if the level or variablity of betas has increased.

- To compare various models that are (or can be) used to calculate time varying betas, in terms of their statistical accuracy

- To quantify the severity of model misspecification in risk management applications, by comparing the one step ahead VaR forecasts of the alternative models.

[1]I also estimated the model implied by Jondeau and Rockinger whoever log likelihood as well as Akaike values are smaller in TVC than DCC in all cases that is why the DCC was my model of choice.

2.2.1 Data and descriptive statistics

My data set consists of the log returns six major financial institutions namely AXP, BAC, C, GS JPM and MS[2] , listed in the standard and poor 500 index and the log return of the S&P 500 index, from 03/01/2006 (the first working day in 2006) till 25/08/2010, a period that nests the global financial crisis. The descriptive statistics of the series are presented in table 2.1

	AXP	BAC	C	GS	JPM	MS	S&P500
mean	-0.0002	-0.0011	-0.0022	0.0001	-0.0001	-0.0007	-0.0002
median	0.0002	0.0000	-0.0012	-0.0001	-0.0004	0.0000	0.0008
IQR	0.0238	0.0223	0.0267	0.0265	0.0232	0.0276	0.0119
MAD	0.0121	0.0113	0.0135	0.0135	0.0117	0.0139	0.0060
st. dev	0.0330	0.0476	0.0517	0.0324	0.0359	0.0466	0.0162
Skewness	0.1236	-0.1873	-0.3117	0.3400	0.3575	1.5857	-0.2151
kurtosis	9.3831	16.3194	22.7704	13.6024	12.4702	39.5787	11.1869
JB test	0	0	0	0	0	0	0

Table 2.1: Descriptive statistics for the seven log returns series. MAD stands for mean absolute deviation and IQR is the interquantile range. In JB test row, the p - values of the normality hypothesis are presented.

From table 2.1 it is evident that the use of Hansen's skew t distribution for our log returns is justified since all series exhibit significant asymmetry and excess kurtosis. Further the hypothesis H_0: *The series is normally distributed* is rejected by the

[2]These are the tickers of the corresponding stocks. AXP is American express, BAC is the Bank of America, C is Citycorp, GS is Goldman Sachs, JPM is J.P. Morgan and MS is Morgan Stanley

Jarque Berra test in all confidence levels.

2.3 Empirical results

In order to estimate the parameters of our six bivariate copulas we resort to the two - step estimation method of Joe and Xu (1996) presented thoroughly in subsection 1.4.2. Each of the seven univariate returns series $r_{i,t}$ is modeled with a GARCH type process as follows:

$$r_{i,t} = c_{0,i} + c_{1,i} r_{i,t-1} + \varepsilon_{i,t}$$

$$\begin{aligned} \varepsilon_{i,t} &= \epsilon_{i,t}\sqrt{h_{i,t}} \\ h_{i,t} &= \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i h_{i,t-1} + \gamma_i \varepsilon_{i,t-1}^2 \cdot I\left(\varepsilon_{i,t-1} < 0\right) \\ \epsilon_{i,t} &\sim iid\ ST(d_i, \lambda_i) \end{aligned} \tag{2.3}$$

This is the GJR model of Glosten et al (1993) specifically designed to capture asymmetric effects in conditional volatility process $h_{i,t}$ and $ST(d_i, \lambda_i)$ is the univariate skew t distribution of Hansen, with density:

$$f_{ST}(z; d, \lambda) = \begin{cases} bc\left(1 + \frac{1}{d-2}\left(\frac{bz+a}{1-\lambda}\right)^2\right)^{-(d+1)/2} & \text{if } z < -a/b \\ bc\left(1 + \frac{1}{d-2}\left(\frac{bz+a}{1+\lambda}\right)^2\right)^{-(d+1)/2} & \text{if } z \geq -a/b \end{cases} \tag{2.4}$$

where d and λ are the degree of freedom and asymmetry parameter, respectively, and

$$a \equiv 4\lambda c \frac{d-2}{d-1}, \quad b \equiv 1 + 3\lambda^2 - a^2, \quad c \equiv \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)\sqrt{\pi\left(d-2\right)}}$$

When $\lambda = 0$ Hansen's Skew t distribution collapses to the central t - distribution with d degrees of freedom. Further if both $\lambda = 0$ and $d \longrightarrow \infty$ the distribution collapses the standard normal distribution, therefore this density nests the other two densities that dominate the applications in GARCH processes. Having estimated the marginal models independently, the standardized residuals are transformed to uniform with the probability integral transformation

$$\epsilon_{i,t} \sim iid\ ST(d_i, \lambda_i) \ \text{ then } \ u_{i,t} = F_{st}\left(\epsilon_{i,t}; d_i, \lambda_i\right) \sim U(0,1)$$

where F_{st} denotes the skew t cumulative distribution function. Having transformed the standardized residuals to uniform, we create six bivariate data sets with uniform margins, of the form $\{u_{i,t}, u_{m,t}\}_{t=1}^{T}$ with $i = \{AXP, BAC, C, GS, JPM, MS\}$. To each of the six data sets a $t-$ copula with time varying correlations ϱ_t and static

degrees of freedom v is fitted

$$\begin{aligned}
\boldsymbol{c}_t\left(u_{i,t},u_{m,t};\varrho_t,v|\mathcal{I}_{t-1}\right) &= \mathcal{C}\left(\varrho_t\right)\cdot\Omega\left(v\right)\cdot\eta\left(\varrho_t,v\right)\\
\mathcal{C}\left(\varrho_t\right) &= \frac{1}{\sqrt{1-\varrho_t^2}}\\
\Omega\left(v\right) &= \frac{\Gamma\left(\frac{v+2}{2}\right)\Gamma\left(\frac{v}{2}\right)}{\Gamma\left(\frac{v+1}{2}\right)^2}\\
\eta\left(\varrho_t,v\right) &= \frac{\prod_{i=1}^{2}\left(1+\frac{z_{i,t}^2}{v}\right)^{\frac{v+1}{2}}}{1+\frac{z_{1,t}^2+z_{2,t}^2-2\varrho_t z_{1,t}z_{2,t}}{\left(1-\varrho_t^2\right)v}}
\end{aligned}$$

with $z_{1,t}=t_v^{-1}\left(u_{i,t}|\mathcal{F}_{t-1}\right)$, $z_{2,t}=t_v^{-1}\left(u_{m,t}|\mathcal{F}_{t-1}\right)$ and $\boldsymbol{z}_t=\left(z_{1,t},z_{2,t}\right)'$ where t_v^{-1} denotes the inverse cdf of the univariate central t - distribution with v degrees of freedom and $\mathcal{F}_{t-1}$is the $\sigma-$field generated by all available information at time $t-1$.The dynamics for the time varying correlations follow the DCC type process of Engle (2002)

$$\begin{aligned}
R_t &= \begin{bmatrix}1 & \varrho_t\\ \varrho_t & 1\end{bmatrix}\\
R_t &= diagQ_t^{-1/2}\cdot Q_t\cdot diagQ_t^{-1/2}\\
Q_t &= \bar{Q}(1-a-b)+a\boldsymbol{z}_{t-1}\cdot\boldsymbol{z}_{t-1}'+\beta Q_{t-1}
\end{aligned}$$

where $\bar{Q}$ is usually taken to be the sample covariance of the transformed standardized residuals $\boldsymbol{z}_t$.Therefore in each bivariate model there at 19 parameters in total that need to be estimated, 8 parameters for each of the two margins and 3 copula parameters. The asymptotic standard errors of the parameters vector ϑ are calculated with

the *Godambe Information Matrix* known as the sandwich estimator: $\hat{\Sigma} = \hat{M}^{-1}\hat{V}\hat{M}^{-1}$ where $\hat{M}$ is the sum of the log likelihood Hessians and $\hat{V}$ is the outer product of sums of the log likelihood gradients, based on the estimate $\hat{\vartheta}$ of ϑ.The estimated marginal parameters along with their corresponding standard errors are presented in table 2.3. All calculations were conducted with MATLAB 7.7 (R2008b) by employing the *fminunc* solver, for nonlinear unconstrained optimization. The reason that fminunc was chosen instead of fmincon (than can accept linear or / and nonlinear constraints) is that fminunc provides more accurate approximations of the log likelihood Hessian than fmincon, which results to more accurate standard errors. However since the problem is constrained, I reparametrized the constrained parameters, to make the unconstarined. The reparametrizations are illustrated in table 2.2

Marginal Parameter	Parameter space	Reparametrization
α	$(0,1)$	$\alpha = \frac{1}{1+\exp(-x)}$
β	(0,1)	$\beta = \frac{1}{1+\exp(-x)}$
d	(2,+∞)	$d = 2 + \exp(x)$
λ	(-1,1)	$\lambda = \frac{\exp(x)-1}{\exp(x)+1}$
Copula Parameter		
a	$(0,1)$	$a = \frac{1}{1+\exp(-x)}$
b	(0,1)	$b = \frac{1}{1+\exp(-x)}$
v	(2,+∞)	$v = 2 + \exp(x)$

Table 2.2: Reparametrizations of the model parameters. x is the corresponding unconstrained parameter. Since $\alpha + \beta < 1$ (and $a + b < 1$), whenever this constrained was violated α (a) was taken $\alpha = 0.9999 - \beta$ $(a = 0.9999 - b)$

	AXP	BAC	C	GS	JPM	MS	S&P500
c_0	0.0002	-0.0002	-0.0003	0.0008	0.0002	$<-10^{-6}$	0.0002
	(0.0005)	(0.0004)	(0.0004)	(0.0006)	(0.0005)	(0.0006)	(0.0003)
c_1	*-0.0994*	-0.0041	0.0268	-0.0554	*-0.0967*	0.0034	*-0.0811*
	(0.0287)	(0.0297)	(0.0295)	(0.0294)	(0.0294)	(0.0287)	(0.0282)
$\omega \times 10^6$	2.337	*2.566*	*3.146*	*11.293*	*3.673*	*6.919*	*1.195*
	(1.358)	(1.242)	(1.450)	(3.802)	(1.634)	(2.806)	(0.456)
α	*0.0527*	*0.1009*	*0.0990*	0.0184	*0.0450*	0.0319	$\sim 10^{-8}$
	(0.0227)	(0.0308)	(0.0288)	(0.0167)	(0.0203)	(0.0220)	($\sim 10^{-5}$)
β	*0.9060*	*0.8616*	*0.8494*	*0.8987*	*0.8822*	*0.9004*	*0.9121*
	(0.0158)	(0.0195)	(0.0190)	(0.0191)	(0.0191)	(0.0168)	(0.0124)
γ	*0.1041*	*0.1273*	*0.1641*	*0.1419*	*0.1721*	*0.1442*	*0.1641*
	(0.0360)	(0.0440)	(0.0481)	(0.0366)	(0.0416)	(0.0401)	(0.0273)
d	*5.0824*	*4.3520*	*4.4996*	*5.2970*	*5.2663*	*4.5370*	*6.3250*
	(0.8940)	(0.6554)	(0.7048)	(0.8037)	(0.8806)	(0.7136)	(1.3498)
λ	0.0102	-0.0361	0.0155	-0.0027	0.0229	-0.0351	*-0.1644*
	(0.0382)	(0.0382)	(0.0367)	(0.0411)	(0.0387)	(0.0375)	(0.0376)
LogL	2794	2797.1	2639.8	2743	2812.1	2594.5	3600.1

Table 2.3: Parameter estimates for the marginal models and robust standard errors. Parameters in italics are statistically significant in the 5% level.

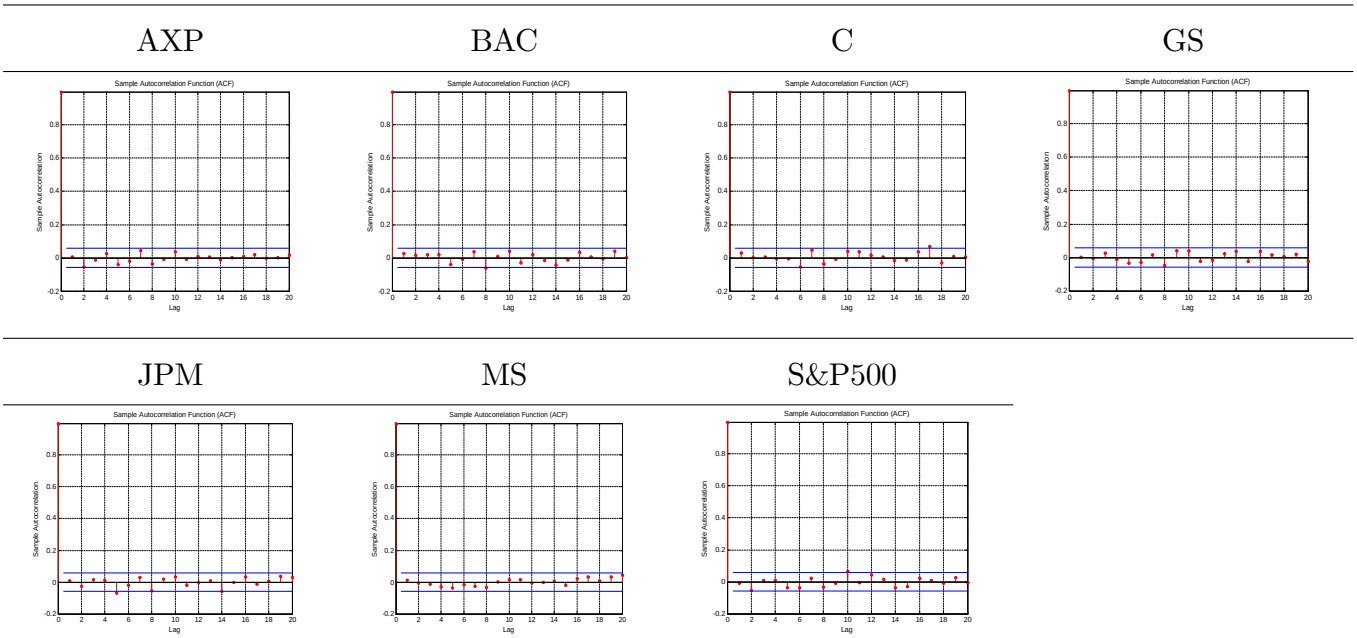


Figure 2.1: Autocorrelation function graphs for the standardized residuals of the seven series.

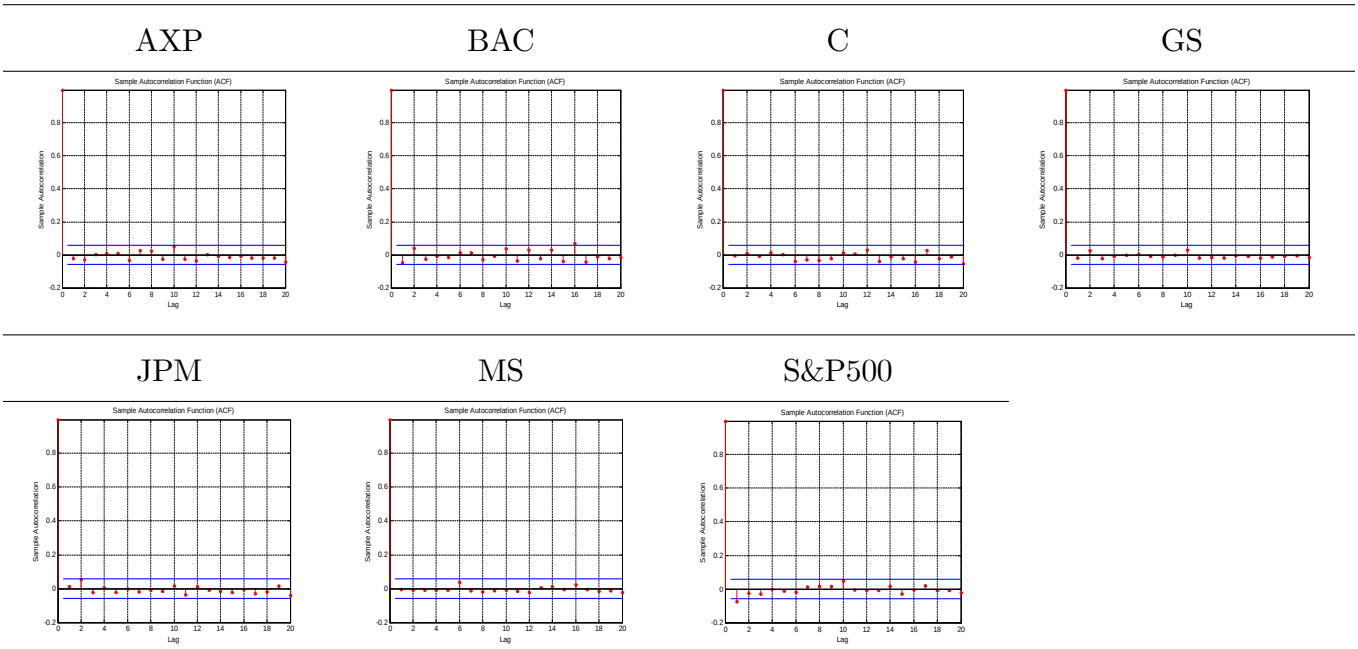


Figure 2.2: Autocorrelation function graphs for the squared standardized residuals of the seven series.

If the specifications of the marginal models are correct then the standardized residuals and squared standardized residuals should not exhibit any residual autocorrelations. The sample autocorrelations of the residuals and squared residuals are depicted in figures 2.1 and 2.2 respectively, out of which is evident that the AR(1) - GJR(1,1) model is adequate to solve both problems for all seven series.

2.3.1 Copula results and time varying parameters

Having estimated the marginal parameters, the standardized residuals were transformed to uniform and the tCGARCH model was fitted to the six data sets. The copula parameters are depicted in table 2.5.

	AXP	BAC	C	GS	JPM	MS
v	*7.4350*	*7.0483*	*7.0594*	*5.8901*	*5.4090*	*2.1271*
	(1.1678)	(1.6175)	(1.6139)	(1.447)	(0.9273)	(0.1917)
a	*0.0644*	*0.0364*	*0.0347*	*0.032*	*0.0465*	*0.009*
	(0.0240)	(0.0112)	(0.0111)	(0.0122)	(0.212)	(0.0037)
b	*0.7847*	*0.9383*	*0.9443*	*0.9285*	*0.7836*	*0.8408*
	(0.0897)	(0.0238)	(0.0211)	(0.0293)	(0.1132)	(0.387)
LogL	554.5667	457.7818	412.8283	452.1737	566.2456	560.8979

Table 2.5: Copula parameters estimates and their corresponding robust standard errors. Parameters in italics are statistically significant in 5% level

From the results in table 2.5 we observe that the choice of the t copula is clearly justified from the small value of the degree of freedom parameter. The degree of

freedom parameter affects the tail dependence coefficient, the probability of o joint extreme event to occur. The probability of a joint negative extreme event is called lower tail dependence (λ^L) while the opposite is called upper tail dependence (λ^U). The Gaussian copula has tail dependence zero, while the tail dependence of the t copula is symmetric $\left(\lambda = \lambda^U = \lambda^L\right)$ and are given by:

$$\lambda = 2T_{v+1}\left(-\sqrt{v+1}\sqrt{\frac{1-\varrho_t}{1+\varrho_t}}\right)$$

Since correlations are time varying so are the tail dependence coefficients whose graphs are depicted in figure 2.3

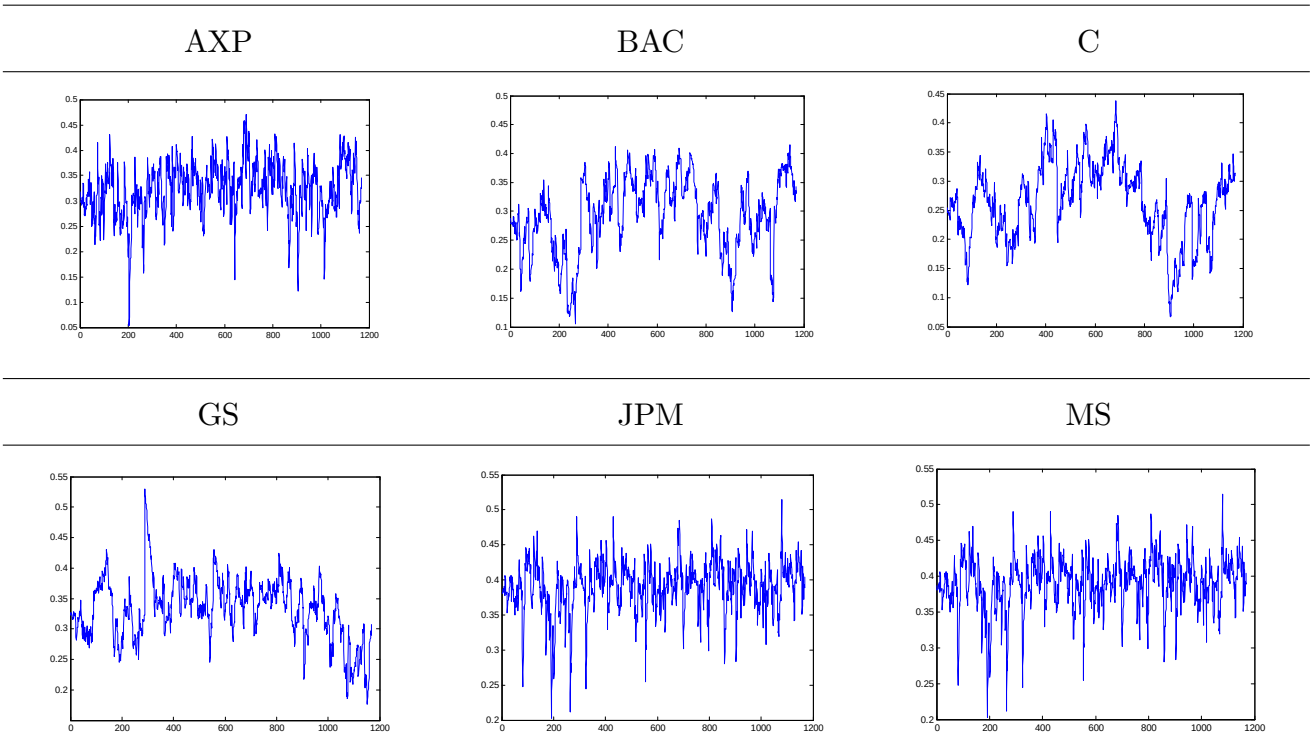


Figure 2.3: Time plots of the tail dependence coefficient for the six estimated t - copulas

Since the tail dependence of each of the six t - copulas is constantly away from zeros we conclude that the use of the Gaussian copula would have clearly underestimated the probability of a joint extreme event a fact that would evidently have significant impact in risk management decisions.

2.3.2 Model comparison and time varying betas

The main purpose of this thesis is to statistically evaluate the alternative models and quantify their implications in terms of risk management decisions. The two models considered are the tCGARCH model and a Gaussian DCC model. Since both models are nested their comparison can be conducted with standard likelihood ratio tests

$$LR = -2\ln\left(\frac{L_{NDCC}}{L_{tCGARCH}}\right) = -2\left(LL_{NDCC} - LL_{tCGARCH}\right) \sim \chi^2(3)$$

where L_X and LL_X denote the likelihood and log likelihood of model X, respectively. Table 2.6 presents the total log likelihood of the two competing models, the corresponding likelihood ratio tests.

	AXP	BAC	C	GS	JPM	MS
$LL_{tCGARCH}$	6948.6	6855.0	6652.7	6795.3	6978.4	6755.4
LL_{NDCC}	6857.3	6772.8	6552.5	6686.8	6863.4	6534.2
LR	182.6***	164.4***	200.4***	217***	230***	442.4***

Table 2.6: Likelihood ratio test for the six bivariate data sets. The critical value of the test in 1% significance is 11.3449.

The likelihood ratio test reject the hypothesis of model equivalence in every confi-

dence lever therefore we conclude that the proposed tCGARCH model is statistically superior than the NDCC model. This is something we expected since the degree of freedom parameter in both the copula and margins was very small in all cases which is a clear indication of departure from normality. But what about betas? In figure 2.4 the time varying betas implied from the tCGARCH model, along with the static beta estimates from the simple market model are plotted.

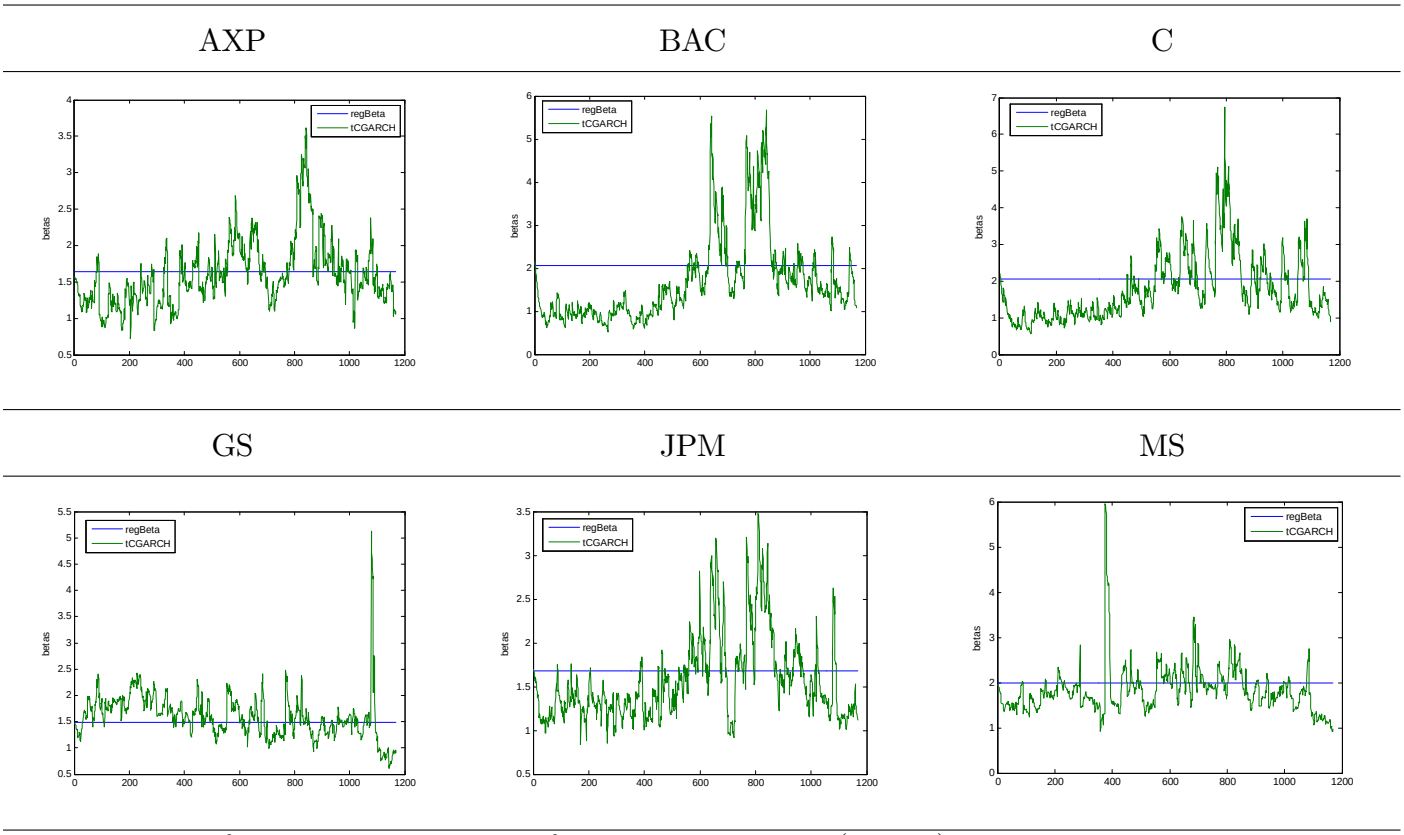


Figure 2.4: Plot of the static beta estimates from the market model (regBeta) against the time varying estimates from the tCGARCH model

From figure 2.4 we observe that stock betas vary significantly in our sample period, therefore the use of the static beta estimates may have serious limitations

in risk management decisions. To compare the betas from the two dynamic models (tCGARCH and NDCC) I plotted their differences. These six plots are presented in figure 2.5. In this case we observe that inspite the clear statistical superiority of the tCGARCh model, the beta estimates from the two dynamic models are very close. I believe that this fact is due to characteristics of correlation. Linear correlation is a measure of dependnece in a Gaussian world, therefore the characteristics (like tail dependence) that constitute the tCGARCH model statistically superior than the NDCC model are not incorporated in the correlation estimates. For more information on the pitfalls of using linear correlation as a measure of dependence, the interested reader is referred to Embrechts et al (2001).

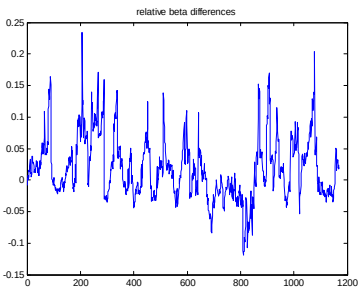
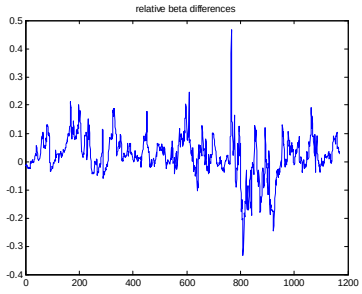
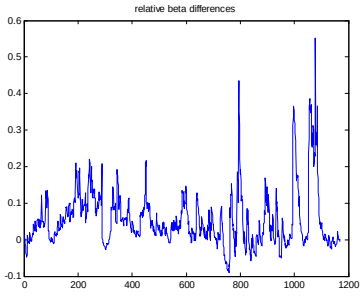
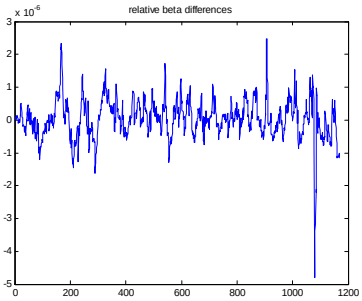
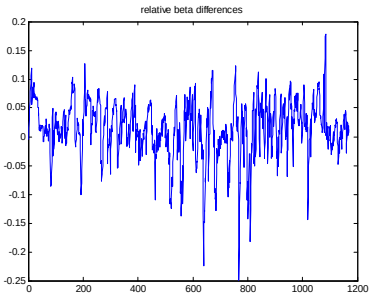
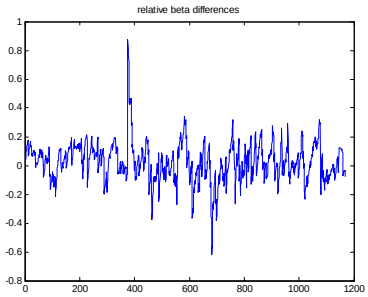
AXP	BAC	C
		
GS	**JPM**	**MS**
		

Figure 2.5: Differences of betas $beta_{tCGARCH} - beta_{NDCC}$ for the six stocks

To quantify the effect of the misspecified model to risk management decisions we calculated one step ahead forecasts of the Value at Risk, for each data set, assuming equally weighted portfolios, for the last 169 days of the portfolio. The one step ahead forecast, for day $t+1$, is calculated as follows:

- Use the copula parameters at time $t+1$ to simulate N correlated uniform values $\{u_1, u_2\}$ that follow the copula of choice.

- Transform these values from $U(0,1)$ to the marginal distributions of choice, using the inverse probability integral transformation. That is if F_i is the marginal distribution, and ϕ_i is the vector of marginal parameters for the $i-$ margin, $i = 1, 2$ then

$$\epsilon_i = F_i^{-1}(u_i; \phi_i)$$

  Each of the N values in ϵ is a path (a realization) of the simulated standardized residuals for time $t+1$

- Transform ϵ to data by

$$y_{t+1}^i = \epsilon_i \sqrt{h_{t+1}^i} + \mu_{t+1}^i$$

- Create the portfolio

$$p_{t+1} = \sum_{i=1}^{2} w^i y_{t+1}^i$$

  where the vector w contains values that sum to unity and calculate the one step ahead $a - VaR$ as the $a-$ empirical quantile, of the values in p_{t+1}

Figure 2.6 contains these VaR forecast for the six data sets assuming the tC-GARCH and the NDCC models.

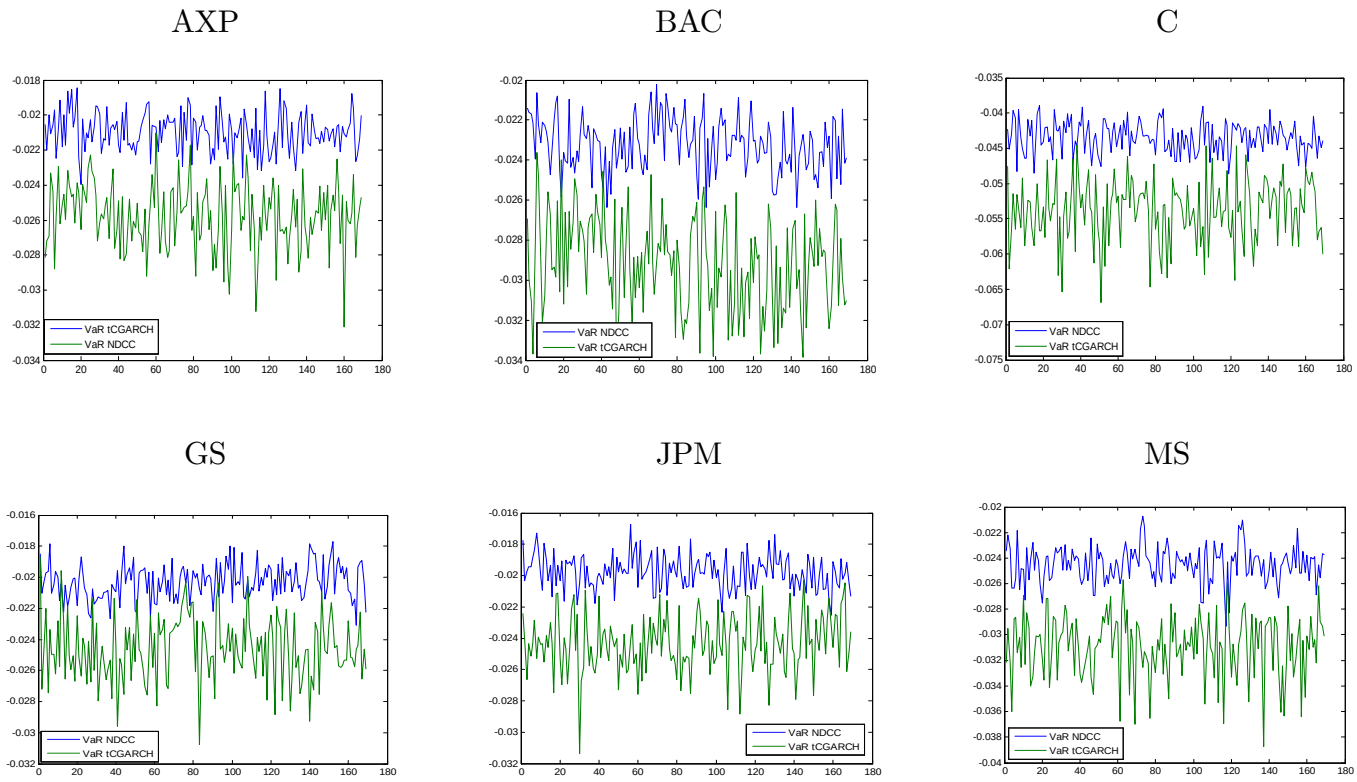


Figure 2.6: One day ahead forecast of 95% VaR's assuming equally weighted portfolios

From figure 2.6 it is clear that the NDCC model underestimates the VaR in all cases, therefore we conclude that although the two models provide similar beta estimates, the statistically superior tCGARCH model clearly outperforms the NDCC model in terms of VaR forecasts.

2.4 Time evolution of betas

One of the main questions arising is if betas have increased during the modern global financial crisis period. The beginning of this period has not been fully clarified yet but most authors believe that the beginning of the period starts in the end of 2006 or in the early 2007. To observe the effect of the new crisis to the stock betas I calculated time varying betas with the tCGARCH model for four years prior to 2006, from 03/01/2002 to 30/12/2005, total 1006 observations. The graphs of the six stocks betas are depicted in figure 2.7

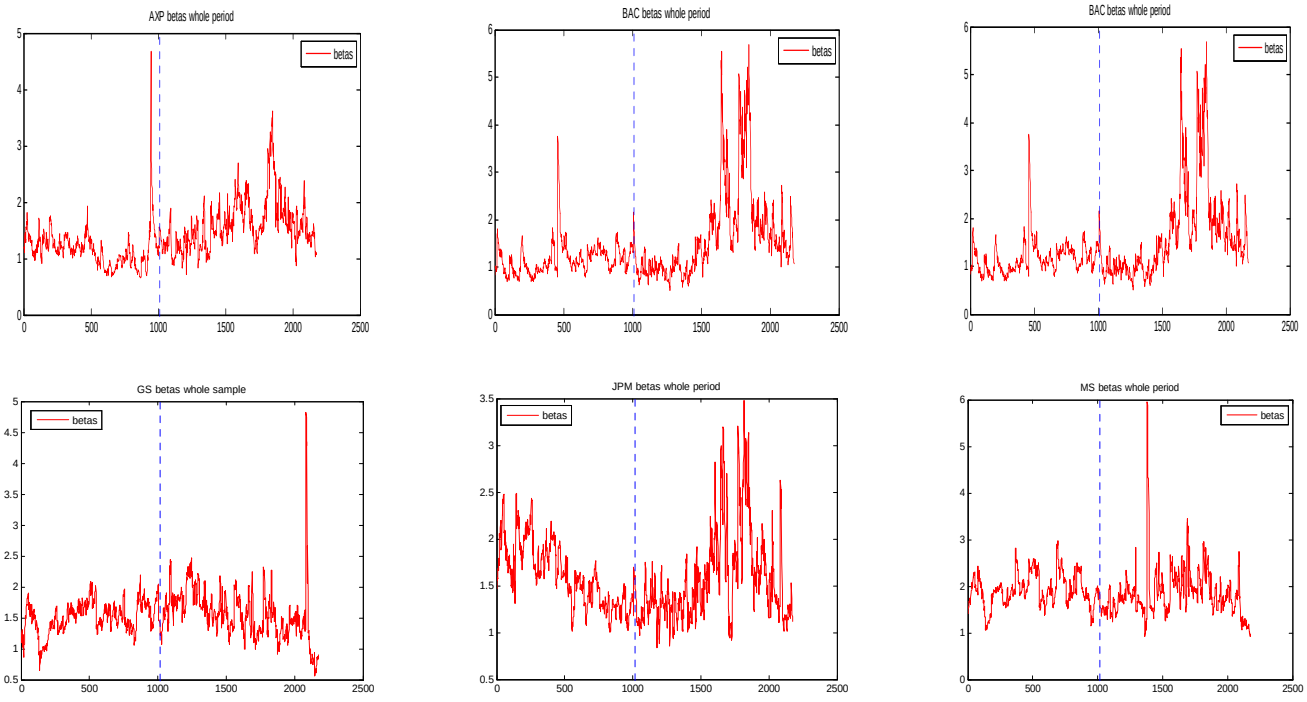


Figure 2.7: Time varying betas for the six stocks, for the whole sample period (03/01/2002 - 25/08/2010). The dashed blue line corresponds to 03/01/2006.

From figure 2.7 it is evident that both the average level of betas and variability of

betas have increased since the beginning of 2006. To quantify this result I calculated the mean, MAD and standard deviations of the betas for the two sub - periods, prior to 03/01/06 and after 03/01/06. Table 2.7 illustrates these descriptive statistics

	AXP	BAC	C	GS	JPM	MS
mean	1.598	1.688	1.774	1.562	1.571	1.831
	1.159	1.163	1.374	1.512	1.613	1.917
MAD	0.254	0.473	0.466	0.228	0.245	0.248
	0.174	0.185	0.166	0.169	0.258	0.194
st. dev	0.455	0.982	0.882	0.420	0.477	0.509
	0.319	0.341	0.266	0.252	0.332	0.343

Table 2.7: Descriptive statistics of betas for the two sub - samples. The upper value is for sub - sample 2 (from 03/01/2006 to 25/08/2010).

From table 2.7 we observe that there was a significant increase in the mean level of betas for half of the stocks (AXP, BAC and C) in the second period, while the mean level of betas for the other three has remained unchanged. However the variability of betas has significantly increased in alla cases, by 42% in the case of AXP to 330% in the case of C. Therefore I argue that betas tend to behave like stock returns; it is not clear that mean returns change significantly during crisis periods but their volatility is significantly increased.

2.5 Other copula models

Apart from the t and the Gaussian copulas, there are other copulas that are used in financial applications. In shake of completeness I also estimated the following models

- A Clayton copula with time varying association parameter
- A Symmetrized Joe Clayton (SJC) copula with both static and time varying copula parameters.
- A static t - copula

The first two copulas belong to the family of Archimedean copulas and they are used in finance because, unlike the t and the Gaussian copula, they exhibit asymmetric dependence. The Clayton copula is tail dependent only in the lower tail (which has the most extreme implications in risk management) but can only incorporate positive dependence, therefore it should be used only when all series are positively correlated. On the other hand the SJC copula exhibits both lower and upper tail dependence and thus it is the most flexible copula used in this study. Table 2.7 contains the log likelihood values of the competing copula models[3]The choice for the best fitting copula is being made with an information criterion like the Akaike (AIC) or the Bayesian (BIC) information criterion. The AIC criterion does not depend on the number of parameters unlike BIC which imposes a penalty for the number of parameters. From table 2.8 the superiority of the t - copula is clear. It is

[3]In all cases the margins were modeled with an AR(1) - GJR(1,1) SkewT model, as in section 2.3

the superior copula in all sets for both criteria. The tCGARCH model is found to be the best model in three cases according to the BIC and in five cases according to the AIC, while the static t - copula is the best fitting copula, when the tCGARCH isn't.

	AXP	BAC	C	GS	JPM	MS
SJC tv 6 parameters	543.296	449.795	403,706	451.810	554.872	558.758
Clayton tv 3 parameters	449.675	381.201	342.541	379.99	474.9	481.42
t static 1 parameter	**551.735**	443.222	395.803	444.5	**563.766**	**562.03**[1]
SJC static 2 parameters	533.566	437.067	388.443	443.111	542.445	527.918
tCGARCH 3 parameters	554.567[1]	**457.782**[1]	**412.828**[1]	**452.174**[1]	566.246[1]	560.898

Table 2.8: Loglikelihood values of the five alternative copula models. Values in boldface letters denote the best fitting copula, according to Bayesian information criterion while values with the superscript 1 are the best fitting copula according to Akaike information criterion

2.6 Conclusion

The multivariate distribution of a portfolio plays a significant role in finance since a number of theories are based upon the statistical characteristics of the returns series.

In the univariate setting there are many distributions that seem to fit financial data adequately however this is not the case when the problem in hand is multivariate. In fact none of the existing multivariate distributions can accommodate all characteristics of financial data. The remedy to this problem is to separate the multivariate distribution to the marginal distributions and the dependence structure which is described by a copula. This decomposition allows researcher to use different parametric models for the dependence and the margins that significantly adds flexibility to the model. For example a multivariate t distribution enforces both the margins and the dependence to have the same degree of freedom parameter which is rather unrealistic. Instead one could use a t - copula with t - margins to allow all margins and the copula to have their own degree of freedom parameter. This study aimed to measure stock betas of some major financial institutions during the contemporary period of global financial crisis by applying the results of copula theory, investigate the implications of the proposed model in risk management applications and compare the betas of the current regime to previous, less volatile periods.. The proposed methodology, unlike traditional models, does not suffer from two serious limitations: It does not assume normality and it can provide time varying beta calculations, unlike the standard methodology used to calculate betas (the market model). I established the statistical superiority of the proposed modeling scheme in comparison with other sophisticated models used in finance and I concluded that, although the proposed model provides similar results to other models that are based on the wrong assumption of normality, it should be the model of choice when risk management decisions are involved, since traditional models clearly underestimate the risk of the

portfolio, in comparison with the tCGARCH model. Further I provided empirical evidence that betas behave like stock returns in terms of behavior during periods of crisis. Even though the mean level of betas does not change during turmoil periods, the variability of betas (like stock volatility) is significantly increased.

References

1. Aas, K., Haff, H., (2006), "The generalized hyperbolic Skew Student's t distribution." *Journal of financial econometrics*, 4 (2), 275 – 309..

2. Alexander, C., (2000), "Orthogonal GARCH." *Mastering risk volume 2.*

3. Aloui R., Mohamed Safouane Ben Aissa, Duc Khuong Nguyen (2010), "Global Financial crisis, extreme interdependences, and contagion effects: The role of economic structure?" *Journal of Banking & Finance* (forthcoming)

4. Andersen, E.W. (2003), "Composite likelihood and two stage estimation in family studies". *Biostatistics* 5, 15 – 30.

5. Ausin, M.C., Lopez, H., (2009), "Time – varying joint distribution through copulas." *Computational statistics and data analysis*, article in press.

6. Barges M., Helene Cossette, Etienne Marceau, (2009), "TVaR-based allocation with copulas.", *Insurance:Mathematics and Economics*

7. Bartram S. M., Stephen J. Taylor, Yaw-Huei Wang, (2006), "The Euro and European financial market dependence.", J*ournal of Banking & Finance*

8. Bawens, L., Laurent, S., (2002), "A new class of multivariate skew densities, with application to GARCH models." *Journal of business and economic statistics.*

9. Bawens, L., Laurent, S., Rombouts, J., (2006), "Multivariate GARCH models: A survey." *Journal of applied econometrics*, 21, 79 – 109

10. Bedendo M., Francesca Campolongo, Elisabeth Joossens, Francesco Saia (2009), "Pricing multiasset equity options:How relevant is the dependence function?" *Journal of Banking & Finance.* 34 788-801

11. Bera, A.K., Kim, S., (2002), "Testing constancy of correlationand other specifications of the BGARCH model." *Working paper, CORE.*

12. Besag, J.E. (1974), "Spatial interaction and the statistical analysis of lattice systems." *Journal of Royal statistical society, Series B (Methodological)*, 34, 192 – 236.

13. Bessler, D. A., Yang, J. (2003). "The structure of interdependence in international stock markets." *Journal of international money and finance* 22, 261 - 287

14. Black, F., Scholes, M., (1973), "The Pricing of Options and Corporate Liabilities". *Journal of Political Economy* 81 (3), 637–654.

15. Bollerslev, T., Engle, R., Wooldridge, J.M., (1988) "A capital asset pricing model with time – varying covariances", *The journal of political economy*, 96, 116 – 131.

16. Bollerslev, T., (1986), "Generalized autoregressive conditional heteroscedasticity." *Journal of Econometrics* 31 (3), 307 - 327

17. Bollerslev, T., Wooldridge, J.M., (1992), "Quasi – maximum likelihood and inference in dynamic models with time – varying covariances." *Econometric reviews* 11 (2), 143 – 172.

18. Berndt,E.K. Hall, B.H. and Hall, R.E. (1974) "Estimation and Inference in Nonlinear Structural Models." *Annals of Economic and Social Measurement*, Vol 3, (4)

19. Caporin, M., Mc Aleer, M., (2006), "Dynamic asymmetric GARCH." *Journal of financial econometrics,* 4 (3), 385 – 412.

20. Chan, Y., Li, H., (2008) "Tail dependence for multivariate t - copulas and its monotonicity." *Insurance: Mathematics and Economics*, Vol42 (2), 763-770

21. Chavez-Demoulin V., P. Embrects, J. Neslehova, (2005), "Quantitative models for operational risk:Extremes, dependence and aggregation.", *Journal of Banking & Finance* 30 2635-2658

22. Cherubini, U., Luciano, E., Vecchiato, W., (2004), "Copula methods in finance." *John Wiley and sons, London.*

23. Chen X., Fan, Y (2006) "Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification" *Journal of Econometrics*, 135, 1-2, 125-154

24. Chen, X, Wu, W.B., Yi, Y., (2009) "Efficient estimation of copula-based semiparametric Markov models." *Annals of statistics*, vol 37 (6B), 4214-4253.

25. Chen Y.H., Anthony H. Tu, Kehluh Wang (2006), "Dependence structure between the credit default swap return and the kurtosis of the equity return distribution: Evidence from Japan.", *Journal of International financial markets, institutions & money* 18 259-271

26. Chollete, L., Heinen, A., Valdesogo, A., (2009), "Modeling international financial returns with a multivariate regime switching copula." *Journal of financial econometrics*, 7 (4), 437 - 480.

27. Cousin A., Jean - Paul Laurent (2008), "Comparison results for exchangeable credit risk portfolios.", *Insurance: Mathematics and Economics.*42 1118-1127

28. Creal, D., Koopman, S.J. and Lucas, A., (2010) "A general framework for observation driven time varying parameter models." *Discussion paper, Tinbergen university.*

29. Davison, A,C., Smith, R.L., (1990), "Models for exceedances over high thresholds." *Journal of Royal statistical society, Series B (Methodological)*, 52 (3), 393 – 442.

30. Demarta, S., Mc Neil, A., (2005), "The t – copula and related copulas." *International statistical review*, 73 (1), 111 – 129.

31. Diamantopoulos, K., Vrontos, D., (2010) "A student - t full factor multivariate GARCH model." Computational economics 35 (1), 63 - 83

32. Dias, A. (2004) "Copula inference for finance and insurance." *PhD thesis, ETH Zurich*

33. Dias, A., Embrechts, P., (2004). "Dynamic copula models for multivariate high frequency data in finance." *ETH working paper.*

34. Dowd, K. (2001), "Measuring Market Risk", *John Wiley and sons.*

35. Durrelman, V., Nikeghbali, A., Roncalli, T., (2000), "A simple transformation of copulas." *Technical report, Credit Lyonnais.*

36. Engle, R.F., (2002), "Dynamic Conditional Correlation: A simple class of multivariate generalized autoregressive conditional heteroscedasticity models." J*ournal of business and economic statistics,* 20, 339 – 350.

37. Engle R.F., Kroner, K. F., (1995), "Multivariate simultaneous generalized ARCH.", *Econometric theory*, 11, 122 – 150.

38. Engle, R.F., Sheppard, K., (2001), "Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH." *NBER working paper 8554*

39. Engle, R.F., Shephard, N., Sheppard, K., (2008), "Fitting vast time – varying covariance models." *Draft paper, SSRN.*

40. Embrechts, P., Mc Neil, A., Straumann, D., (2001). "Correlation and dependence in risk management: Properties and Pitfalls." *Risk management: value at risk and beyond. Cambridge university press.*

41. Fantazzini, D., (2008), "Dynamic Copula modeling for Value at Risk." *Frontiers in finance and economics*, 5 (2), 72 – 108.

42. Fermanian, J.D. and Scailet, O. (2004) "Some statistical pitfalls in copula modeling for financial applications." "*FAME research paper series*, rp 108.

43. Fermanian, J.D. and Wegkamp, M.J. (2004) "Time dependent copulas." *Working paper, centre de researche en economie et en statistique, Paris France*

44. Fernandez, V., (2008) "Copula-based measures of dependence structure in assets returns." *Physica A: Statistical Mechanics and its Applications*, Vol 387 (14), 3615-3628

45. Fiorentini G., Sentana E., CalzolariG., (2003), "Maximum likelihood estimation and inference in multivariate conditionally heteroscedastic dynamic regression models with Student t innovations." *Journal of business and economic statistics* 21, 532 – 546.

46. Glosten, R.T., 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 (5), 1779 – 1801

47. Geman H., Cecile Kharoubi(2008), "WTI crude oil Futures in portfolio diversification:The time -to-maturity effect.", *Journal of Banking and Finance*

48. Genest, C., Mackay, J., (1986). " The joy of copulas: bivariate distributions with uniform marginals". *American statisticial*, vol 40, number 4.

49. Genest, C., Ghoudi, K., Rivest, L., (1995), "A semi – parametric estimation procedure of dependence parameters in multivariate families of distributions." *Biometrica* 82, 543 – 552.

50. Genest, C., Verret, F., (2005) "Locally most powerful rank tests of independence for copula models." *Journal of nonparametric statistics*, vol 17 (5), 521 – 539

51. Genest, C., Quessy, J.F. and Remillard, B. (2006) "Goodness-of-fit Procedures for Copula Models Based on the Probability Integral Transformation." *Scandinavian Journal of Statistics*, 33 (2), 337-366

52. Genest, C. and Favre, A. (2007). "Everything you always wanted to know about copula modeling but were affraid to ask." *Journal of hydrologic engineering,* 12 (4), 347 - 368

53. Genest, C., Gendron, M., Bourdeau - Brien, M., (2009). "The Advent of Copulas in Finance." The *European journal of finance.*

54. Giacomini, E., Hardle, W., Ignatieva,E., and Spokoiny, V., (2006) "Inhomogeneous dependency modeling with time varying copula." *SFB 649 discussion paper*

55. Hamilton, J.D. (1994) "Time series analysis." *Princeton university press, Princeton*

56. Hansen, B.E. (1994),"Autoregressive conditional density estimation." *International Economic Review* 35 (3), 705 – 730

57. Huang, J.J., Lee, K.J., Liang, H., and Lin, W.F., (2009) "Estimating value at risk of portfolio by conditional copula-GARCH method." *Insurance: Mathematics and economics*, 45 (3), 315 - 324

58. Hult, H., Lindskog, F., (2002), 'Multivariate extremes, Aggregation and Dependence in Elliptical Distributions.' *Advances in Applied Probability*, 34 (3), 587 – 608

59. Hurlimann V., (2004). "Fitting bivariate cumulative returns with copulas." *Computational statistics and data analysis* 45, 355 - 372 .

60. Joe, H., (1996,1997), "Multivariate models and dependence concepts." *Chapman and Hall, London.*

61. Johnson, N.L., (1949), "Systems of frequency curves generated by method of translation." *Biometrica* 36, 149 – 176.

62. Jondeau, E. and Rockinger, M. (2006),"The Copula – GARCH modelof conditional dependencies: An international stock market application.", *Journal of International money and finance*, 25, 827 – 853.

63. Jorion, P. (1997), "Value at risk: The new benchmark for managing financial risk", *MC Grow – Hill*

64. Kenourgios D., Ar. Samitas, N.Paltalidis (2010), "Financial c rises and stock market contagion in a multivariate time-varying asymmetric framework.". *Journal of International Financial Markets, Institutions & Money*

65. Kim, G., Silvapulle, M.J., Silvapulle, P. (2007). "Comparison of semiparametric and parametric methods for estimating copulas." *Computational statistics and data analysis* 51, 2386 - 2850.

66. Kole, E., Koedijk, K., Verbeek, M., (2007), "Selecting copulas for risk management." *Journal of banking and finance*, 31 (8), 2405 – 2423.

67. Kumar M.S., Tatsuyoshi Okimoto (2010), "Dynamics of International integration of government securities' markets.", *Journal of Banking & Finance* forthcoming

68. Liu, Y. and Luger R., (2009) "Efficient estimation of copula GARCH models." *Computational statistics and data analysis* 53, 2284 - 2297

69. Longin, F., Solnik, B., (1995), "Is the correlation in equity markets constant: 1960 – 1990?" *Journal of International money and finance* 14 (1), 3 – 26.

70. Malo P., (2009), "Modeling electricity spot and futures price dependence:A multifrequency approach.", *Physica A* 388 4763-4779

71. Manner, H., Reznikova, O., (2009) "Time varying copulas: A survey" *IAP statistical network*: http://www.stat.ucl.ac.edu/IAP

72. Markowitz, H.M. (1952), "Portfolio Selection". *The Journal of Finance* 7 (1), 77–91.

73. Mashal, R., Zeevi,A. (2002). "Beyond correlation: Extreme co - movements between financial assets." *Working paper, Columbia graduate school of business.*

74. Mendez, B., Souza, R. (2004). "Measuring financial risks with copulas." *International review of financial analysis* 13, 27 - 45 .

75. Merton, R. C. (1973). "Theory of Rational Option Pricing". *Bell Journal of Economics and Management Science* 4 (1), 141–183.

76. Morgan guaranty Trust company, Global research (1995), '*Riskmetrics technical document.*'

77. Nelson, D.B. (1991), "Conditional heteroscedasticity in asset returns: A new approach," *Econometrica*, 59, 347 – 70

78. Nielsen, R.B., (1999, 2006), "An introduction to copulas." *Springer, first and second edition.*

79. Ning C., Tony S. Wirjanto (2009), "Extreme return-volume dependence in East -Asian stock markets:A copula approach.", *Finance Research Letters* 6 202-209

80. Ning C.,(2009), "Dependence structure between the equity market and the foreign exchange market- A copula approach.", *Journal of International Money and Finance* 29 743-759

81. Panchenko. V., (2005) "Goodness of fit tests for copulas." *Physica A: Statistical Mechanics and its Applications*, 355 (1), 176 - 182

82. Patton, A., J. (2007). "Copula-Based Models for Financial Time Series" in T.G. Andersen, R.A. Davis, J.-P. Kreiss and T. Mikosch (eds.) *Handbook of Financial Time Series*, Springer Verlag.

83. Patton, A.J., (2006). "Modelling Asymmetric Exchange Rate Dependence." *International economic review*, 47, 527 – 556.

84. Patton, A.J., (2006), "Estimation of copula models for time series of possibly different lengths." *Journal of applied econometrics*, 21 (2), 147 – 173.

85. Patton, A.J. ,(2002),"Modeling Time-Varying Exchange Rate Dependence Using the Conditional Copula", *Working Paper 01-09, Department of Economics, University of California, San Diego.*

86. Pelagatti, M.M., Rondena, S., (2004), "Dynamic Conditional Correlation for elliptical distributions." *Discussion paper, universita di Milano – Bicocca.*

87. Peletier, D., (2006) "Regime switching for dynamic correlations." *Journal of econometrics* 131, 445 - 473

88. Perignon C., Daniel R. Smith (2009), "Diversification and Value-at-Risk.", *Journal of Banking & Finance* 34 55-66

89. Rodriguez, J.C., (2007). "Measuring financial contagion: A copula approach." *Journal of empirical finance* 14, 401 - 423 .

90. Serban, M., Brockwell, A., Lehoczky, J., Srivastava, S., (2006), "Modeling the dynamic dependence structure in multivariate financial time series." *Journal of time series analysis*, 28, (5), 763-782

91. Silvennoinen A., Terasvirta, T., (2008), "Multivariate GARCH models." *Handbook of Financial Time Series. New York: Springer*

92. Sklar, A., (1959) "Functions de repartition a n dimensions et leurs marges." *Publications de I'Institut Statistique de l'Universite de Paris* 8, 229 – 231

93. Terasvirta, T., (2006), "An introduction to univariate GARCH models." *SSE/EFI working papers in economics and finance,* No 646.

94. Tse, Y.K., (2000) "A test for constant correlations in a multivariate GARCH model." *Journal of econometrics* 98 (1), 107 – 127.

95. Tse, Y.K., Tsui, K.C., (2002), "A multivariate generalized autoregressive conditional heteroscedasticity model with time – varying correlations." *Journal of business and economic statistics*, 20, 351 – 362.

96. Tsay, R.S. (2002), "Analysis of financial time series." *John Wiley and sons.*

97. Van den Goorbergh, R.W.J., Genest, C., Werker, B.J.M. (2005), "Bivariate option pricing using dynamic copula models." *Insurance: Mathematics and economics*, 37, 101 – 114.

98. Van rer Vaart (1998). "Asymptotic statistics." *New York: Cambridge university press.*

99. Wang, W., Wells, M., (2000) "Estimation of Kendall's tau under censoring." *Statistica sinica*, 10,1199-1215

100. Wang Z. R., Xiao-Hong Chen, Yan-Bo Jin, Yan-ju Zhou (2010), "Estimating risk of foreign exchange portfolio: Using VaR and CVaR based on GARCH-EVT-Copula model.", *Physica A* 389 4918-4928

101. Zhang, J., Guegan D. (2009) "Pricing bivariate option under GARCH processes with time-varying copula." *Insurance: Mathematics and Economics*, Vol 42 (3), 1095-1103