HAL

archives-ouvertes.fr

Dependence modelling of the joint extremes in a portfolio using Archimedean copulas : application to MSCI indices

Dominique Guegan, Sophie A. Ladoucette

▶ To cite this version:

HAL Id: halshs-00189214

https://halshs.archives-ouvertes.fr/halshs-00189214

Submitted on 20 Nov 2007

Dependence modelling of the joint extremes in a portfolio using Archimedean copulas : application to MSCI indices

Dominique GUEGAN, IDHE

Sophie A. LADOUCETTE, IDHE

2005.101

Dependence modelling of the joint extremes in a portfolio using Archimedean copulas: application to MSCI indices

Dominique Guégan   Sophie A. Ladoucette

Abstract: Using Archimedean copulas, we investigate the dependence structure existing between several series of financial assets log-returns that come from different markets. These series are considered as components of a portfolio and they are investigated on a long period including high shocks. To perform such a study, we model the tail of their joint distribution function using a dependence measure (Kendall's tau) and its relationship with the class of Archimedean copulas. Then, we define two different diagnostics to decide which copula best fits the tail of the empirical joint distribution. This approach permits us to understand the evolution of the interdependence of more than two markets in the tails, that is when extremal events corresponding to shocks induce some turmoil in the evolution of these markets.

1 Introduction

In the financial framework, the concept of tail modelling can be used to understand the dependence between series that come from several markets when extremal events or high frequency log-returns (which correspond to shocks) occur and induce some turmoil in the evolution of the markets. In this paper, we are interested to model the tail of the distribution of n components of a portfolio through Archimedean copulas, that is to model the dependence structure of the distribution of joint extremes.

Consider a general random vector $X = (X_1, \cdots, X_n)'$ which may represent n components of a portfolio measured at the same time. Assume that X has a n-dimensional joint distribution $F(x_1, \cdots, x_n) = \mathbb{P}[X_1 \leq x_1, \cdots, X_n \leq x_n]$. Assume further that for all $i \in \{1, \cdots, n\}$, the random variable X_i has a continuous marginal distribution F_i with $F_i(x) = \mathbb{P}[X_i \leq x]$. Under such assumptions, it has been shown by Sklar (1959) that the joint distribution F with marginals $F_1, \cdots, F_n$ can be written as:

$$F(x_1, \cdots, x_n) = C\big(F_1(x_1), \cdots, F_n(x_n)\big), \quad (x_1, \cdots, x_n) \in \mathbb{R}^n \qquad (1)$$

for an unique function C called the copula of F or X. Then, a copula C is a multivariate distribution with uniform marginals on $[0, 1]$, and it provides a natural link between F and $F_1, \cdots, F_n$. From (1), we see that the univariate marginals and the dependence structure can be separated, and it makes sense to interpret C as the dependence structure of the random vector X.

From a practical point of view, to model the tail of the joint distribution F of a multivariate random vector X with continuous marginal distributions, we need to choose the copula which best takes into account the dependence structure of X.

The data sets used in the empirical study of this paper consist of some Morgan Stanley Capital International (MSCI) daily closing prices for the American, French and Japanese markets. These data have been collected from Data Stream from January 1, 1985 to December 31, 2001, which provides a total of 4435 observations for each of the three markets. In the following, we consider the log-returns of these MSCI series that are denoted X_1 for the American market, X_2 for the French market and X_3 for the Japanese market. For these series, we aim at finding copulas that best model the tails of the empirical joint distributions for $X = (X_1, X_2)'$, $X = (X_1, X_3)'$ and $X = (X_2, X_3)'$ (bivariate case, $n = 2$), and for $X = (X_1, X_2, X_3)'$ (trivariate case, $n = 3$).

The particular copulas that we use in the sequel are chosen among the class of Archimedean copulas (e.g., Genest and MacKay, 1986a and 1986b). In-

deed, these particular copulas are useful for empirical studies since they are easily built through Kendall's tau on the one hand, and they allow to extend the modelling of bivariate series to n-variate series with $n \geq 3$ on the other hand. Kendall's tau is a dependence parameter which takes into account the existence of non-linear features inside data sets (e.g., Kendall and Stuart, 1979) and it is easily computable from real data.

In its spirit, this work is close for instance to the paper of Embrechts et al. (2001b) concerning the problem of Integrated Risk Management. Here, we adopt a statistical point of view and we show how Archimedean copulas can be used to adjust the best multivariate distribution to the n components of a given portfolio. There are numerous alternative applications of copulas techniques to Integrated Risk Management and our paper is a contribution among others. We can cite for instance the applications developed in Embrechts et al. (1997) on Danish fire data, the work of Rockinger and Jondeau (2001) in the 2-dimensional setting with Plackett's copula or the work of Blum et al. (2002) on Alternative Risk Transfer problem. Other references can be found in Embrechts (2000), Scaillet (2000) and Embrechts et al. (2001a).

The paper is organized as follows. In Section 2, we present the class of Archimedean copulas which are of interest, and we specify the link existing between Kendall's tau and these particular copulas. Section 3 is devoted to the empirical study. We begin by providing a statistical presentation of the data sets. Using the Peak Over Threshold method, we estimate the tails of the marginal univariate distributions F_i of X_i, $i \in \{1, 2, 3\}$, with Generalized Pareto Distributions. Then, using Archimedean copulas, we model the tails of the bivariate joint distributions (Paragraph 3.2) and, under some assumptions, of the trivariate joint distribution (Paragraph 3.3) of the various series of MSCI log-returns. Finally, we propose two diagnostics (one using a numerical criterion and one graphical with a QQ-plot method) which permit to retain the best copula within those proposed. We apply these diagnostics to our data sets. In Section 4, we formulate some conclusions.

2   Measuring dependence through Archimedean copulas

For the purpose of this paper, we concentrate on an important class of copulas called Archimedean copulas (e.g., Genest and MacKay, 1986a and 1986b). This class of copulas is worth studying. In particular, they allow for a great variety of different dependent structures and, in contrast to the family of elliptical copulas, they have closed form expressions and they are not derived from multivariate distributions using Sklar's Theorem (e.g., Cambanis et al.,

1981 and Hult and Lindskog, 2001). Nevertheless, a disadvantage of this last feature is that extensions of bivariate Archimedean copulas to multivariate ones need some technical assumptions on their parameters. This point will be considered in Paragraph 3.3.

A (bivariate) Archimedean copula C_α has the property to be generated by a convex function φ_α that is continuous and strictly decreasing from $[0,1]$ to $[0,\infty[$ with $\varphi(0) = \infty$ and $\varphi(1) = 0$, and which depends on a dependence parameter α so that, it has the following form:

$$C_\alpha(u,v) = \varphi_\alpha^{-1}(\varphi_\alpha(u) + \varphi_\alpha(v)), \quad (u,v) \in [0,1]^2.$$

An important characteristic of Archimedean copulas, which will be used in the sequel, is that there exists a formula linking the dependence parameter α of the generator φ_α and a measure of dependence called Kendall's tau. We refer to Kendall and Stuart (1979) for details about this coefficient.

The general definition of Kendall's tau τ for two random variables X_1 and X_2 is the probability of concordance minus the probability of discordance:

$$\tau(X_1, X_2) = \mathbb{P}[(X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) > 0] - \mathbb{P}[(X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) < 0]$$

where $(\tilde{X}_1, \tilde{X}_2)'$ is an independent copy of the vector $(X_1, X_2)'$. For a general copula C, Kendall's tau can be expressed as a double integral of C (e.g., Nelsen, 1999). However, for an Archimedean copula C_α, Genest and MacKay (1986) have shown that it depends on the generator φ_α and its derivative in the simple following form:

$$\tau = 1 + 4 \int_0^1 \frac{\varphi_\alpha(t)}{\varphi_\alpha'(t)} dt. \tag{2}$$

For our empirical study, the class of Archimedean copulas is of fundamental interest. Indeed, from real data we estimate Kendall's tau which permits to compute the dependence parameter of various generators φ_α through (2) and then to construct the corresponding copulas C_α.

The four Archimedean copulas that we use in this paper are the following:

- Gumbel copula, see Gumbel (1958):

$$G_\alpha^G(u,v) = \exp\Big(-\big(|\log u|^\alpha + |\log v|^\alpha\big)^{1/\alpha}\Big), \ \alpha \in [1,\infty[.$$

- Cook and Johnson copula, see Cook and Johnson (1981):

$$C_\alpha^{CJ}(u,v) = \Big(u^{-\alpha} + v^{-\alpha} - 1\Big)^{-1/\alpha}, \ \alpha \in]0,\infty[.$$

4

- Ali-Mikhail-Haq copula, see Ali et al. (1978):

$$C_\alpha^{AMH}(u,v) = \frac{uv}{1 - \alpha(1-u)(1-v)}, \ \alpha \in [-1,1].$$

We remark that when $\alpha = 1$, we get as particular case the Cook and Johnson copula.

- Frank copula, see Frank (1979):

$$C_\alpha^F(u,v) = \log_\alpha \left(1 + \frac{(\alpha^u - 1)(\alpha^v - 1)}{\alpha - 1} \right), \ \alpha \in \ ]0, \infty[\backslash\{1\}.$$

We refer to Avouyi-Dovi et al. (2002) for the analytical expression of the generator φ_α and for the exact relationship between α and τ, for each of these four copulas.

3 Empirical study

Now, we propose an empirical study which aims at modelling the tail of the joint distribution of n components of a portfolio through Archimedean copulas. For this study, we use the series of MSCI daily log-returns that come from the American, French and Japanese markets.

3.1 Statistical presentation of the data sets

We briefly present the basic statistics of the log-return series of the MSCI daily indices before investigating the distribution of their joint extremes. The log-returns are of interest rather that the prices in order to achieve stationarity. We recall that the log-return series are denoted X_1 for the American market, X_2 for the French market and X_3 for the Japanese market.

Figure 1 displays the trajectories and the empirical distributions of the three series on the sample period from 01/01/1985 to 31/12/2001 with a total of $N = 4434$ points.

Series	Mean	Standard deviation	Skewness	Kurtosis
X_1	4.34 10^{-4}	1.04 10^{-2}	-2.67	59.93
X_2	5.27 10^{-4}	1.23 10^{-2}	-0.37	7.03
X_3	1.94 10^{-4}	1.47 10^{-2}	-0.10	12.69

Table 1: Statistics for the series X_1, X_2 and X_3 on the period 01/01/1985-31/12/2001.

In Table 1, we summarize the computations of the first four empirical moments of the series. We remark that the empirical skewness are far from
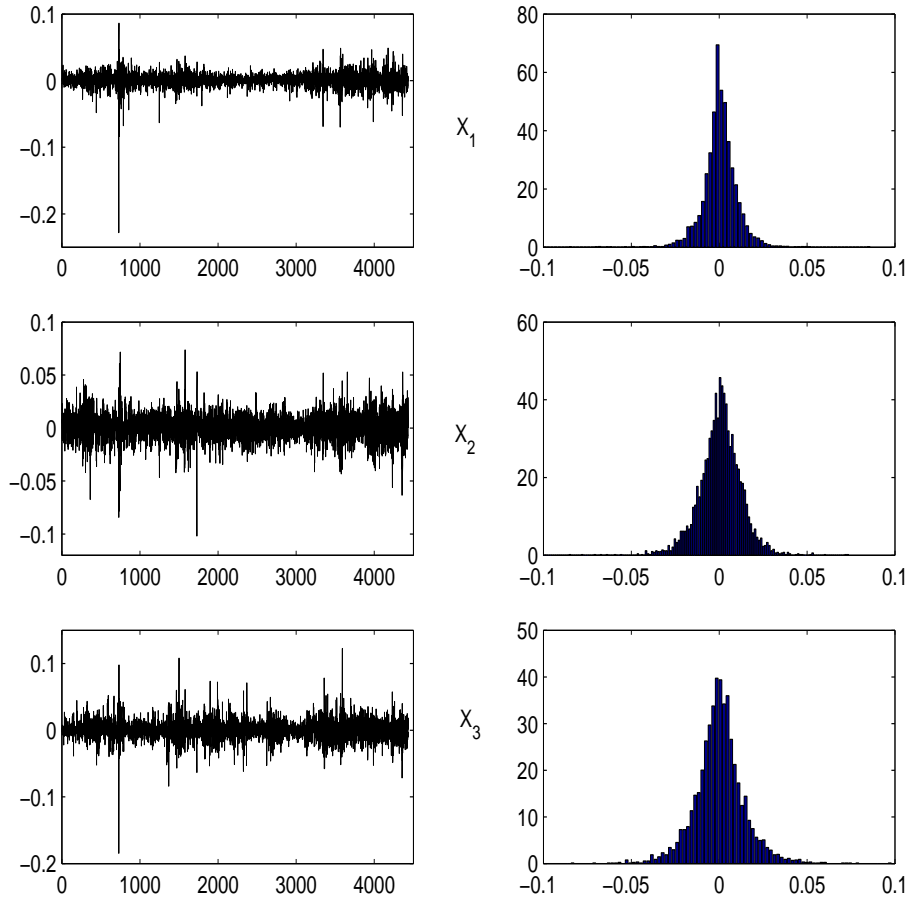
5

Figure 1: Trajectories and histograms of the series X_1, X_2, X_3 on the period 01/01/1985-31/12/2001.

zero and that the series exhibit excess kurtosis relative to the Gaussian distribution. We confirm that the series are non Gaussian since they follow logLaplace distributions that have been adjusted using a Kolmogorov-Smirnov test with 95% level, see Avouyi-Dovi et al. (2002).

3.2 Bivariate case

In order to adjust the best copula on the tail of the empirical joint distribution of $(X_i, X_j)'$, $i \neq j \in \{1, 2, 3\}$, we proceed in five steps. First of all, using the Peak Over Threshold method, we estimate the tail distribution of each series (e.g. Embrechts et al., 1997). Then, we compute empirical Kendall's tau in the tails for each pair of series. Using these values, we deduce the empirical dependence parameter α for the four Archimedean copulas C_α^G, C_α^{CJ},

C_α^{AMH} and C_α^F. Thus, applying Sklar's Theorem, we model the bivariate distribution of each pair with these different copulas. Finally, we build two diagnostics to decide the copula that best models the tail of each bivariate series.

First step. To model the tails of the univariate distributions $\hat{F}_1$, $\hat{F}_2$ and $\hat{F}_3$ of the three series X_1, X_2 and X_3, we use the Peak Over Threshold (POT) method whose we briefly recall the principle. If a random variable X follows a distribution function F, we define the associated distribution of excesses over a high threshold u as:

$$F_u(y) = P[X - u \le y | X > u] = \frac{F(y+u) - F(u)}{1 - F(u)} \qquad (3)$$

for $0 \le y < x_+ - u$, where $x_+ \le \infty$ is the upper endpoint of F. For a large class of distributions F (including all the common continuous distributions), the excess function F_u converges to a Generalized Pareto Distribution (GPD), denoted $G_{\xi,\beta}$, as the threshold u is raised. By the way, we can assume that the GPD models can approximate the unknown excess distribution F_u, i.e. for a certain threshold u and for some ξ and β (to be estimated), we have:

$$F_u(y) = G_{\xi,\beta}(y). \qquad (4)$$

By setting $x = u + y$ and combining expressions (3) and (4), we get:

$$F(x) = (1 - F(u))G_{\xi,\beta}(x - u) + F(u), \ x > u \qquad (5)$$

which permits us to get an approximation of the tail of the distribution F.

From an empirical point of view, if we deal with a time series with unknown underlying distribution F, we build an estimate for $F(u)$ using $1 - N_u/N$, where N_u is the number of data exceeding the fixed threshold u, and if we estimate the parameters ξ and β of the GPD, we get the following estimator for the tail distribution:

$$\hat{F}(x) = 1 - \frac{N_u}{N}\Big(1 + \hat{\xi}\frac{x - u}{\hat{\beta}}\Big)^{-1/\hat{\xi}} \qquad (6)$$

which is only valid for $x > u$.

Now, we consider the log-return series X_i, $i \in \{1, 2, 3\}$, from the three markets. For each series, we choose a threshold u that corresponds to the 95th sample percentile: on the one hand, the threshold has to be chosen sufficiently high so that the approximation (4) can be applied, and on the other hand it has to be considered sufficiently low to have sufficient data for the estimation procedure. This means that we define the tails of the empirical

distributions of the three series in considering the upper 5% of the total number of observations (given the $N = 4434$ data, this implies $N_u = 222$ threshold exceedances). Then, we fit the GPD to the N_u excesses using Maximum Likelihood Estimation (MLE) of the parameters ξ and β and we compute the confidence intervals at the 95% level for the parameters' estimates using a bootstrap procedure. The results have been summarized in Table 2.

Series	Threshold and GPD estimates
X_1	$u_1 = 1.4182$, $\hat{\beta}_1 = 0.6129$ [0.5891,0.6714], $\hat{\xi}_1 = 0.1033$ [0.0454,0.1306]
X_2	$u_2 = 1.5071$, $\hat{\beta}_2 = 0.5688$ [0.5396,0.6143], $\hat{\xi}_2 = 0.0818$ [0.0205,0.1185]
X_3	$u_3 = 1.6037$, $\hat{\beta}_3 = 0.5996$ [0.5564,0.6442], $\hat{\xi}_3 = 0.1704$ [0.0844,0.2064]

Table 2: Estimates (with confidence intervals in brackets) for the parameters of the GPD adjusted on the tails of the series X_1, X_2, X_3 and values of the threshold that corresponds to the 95th percentile.

For each of the three markets, the parameter ξ being significantly non null and positive, the distribution F_i of the whole series X_i belongs to the domain of attraction of the Fréchet distribution.

Now, using the tail estimator (6) with the estimated values of $\hat{\xi}$ and $\hat{\beta}$ given in Table 2, we can compute the tail of the empirical marginal distribution $\hat{F}_i$ of X_i for $x_i > u_i$, $i \in \{1, 2, 3\}$.

Second step. We compute the empirical values $\hat{\tau}$ of Kendall's tau between the different pairs of series (X_1,X_2), (X_1,X_3) and (X_2,X_3) considered in their tails. The tails are defined by the points on which we have adjusted the GPD in the first step presented before. We report the values in Table 3.

Pair	$\hat{\tau}$
(X_1,X_2)	0.0688
(X_1,X_3)	0.0136
(X_2,X_3)	0.0239

Table 3: Empirical Kendall's tau for each pair of series considered in the tails defined by the 95th sample percentile.

The values $\hat{\tau}$ show that the bivariate series are almost uncorrelated for high-

probability events whereas they assume high values jointly.

Third step. We compute the parameters α of the Archimedean copulas that we fit to the tail of the joint distribution of the pairs (X_1, X_2), (X_1, X_3) and (X_2, X_3).

Using the formula (2) and the values of $\hat{\tau}$ given in Table 3, we easily compute the dependence parameter $\hat{\alpha}$ of the different Archimedean copulas as follows: $\hat{\alpha} = \frac{1}{1-\hat{\tau}}$ for the Gumbel copula and $\hat{\alpha} = \frac{2\hat{\tau}}{1-\hat{\tau}}$ for the Cook and Johnson copula. For the Ali-Mikhail-Haq copula and for the Frank copula, we need numerical resolutions since such easy expression between α and τ is not available. The values of $\hat{\alpha}$ that we get for the different pairs are reported in Table 4.

Fourth step. Using the tail of the empirical distribution $\hat{F}_i$ of X_i for $x_i > u_i$, $i \in \{1, 2, 3\}$, that we have computed in the first step and applying Sklar's Theorem, we get the following relationship for each pair (X_i, X_j):

$$\hat{F}(x_i, x_j) = C_{\hat{\alpha}}\big(\hat{F}_i(x_i), \hat{F}_j(x_j)\big), \quad x_i > u_i, x_j > u_j \tag{7}$$

where $C_{\hat{\alpha}}$ denotes one of the four Archimedean copulas with parameter $\hat{\alpha}$ computed in the third step.

Fifth step. We propose two different diagnostics: a numerical method and a graphical method. These diagnostics permit us to decide, among the range of the Archimedean copulas C_{α}^G, C_{α}^{CJ}, C_{α}^{AMH} and C_{α}^F, the copula that best models the tail of the empirical joint distribution of the pairs (X_1, X_2), (X_1, X_3) and (X_2, X_3).

The first method needs a numerical criterion that we denote $D_C^{(2)}$ and that corresponds to:

$$D_C^{(2)} = \sum_{x_1, x_2} \left| C_{\hat{\alpha}}\big(\hat{F}_1(x_1), \hat{F}_2(x_2)\big) - \hat{F}(x_1, x_2) \right|^2.$$

Then, the copula $C_{\hat{\alpha}}$ for which we get the lowest $D_C^{(2)}$ will be chosen as the best copula. For the various copulas and for each pair of series, the $D_C^{(2)}$ values that we have computed are given in Table 4.

According to the criterion $D_C^{(2)}$, we get the following results. For the pair (X_1, X_2), we get the best model using the Cook and Johnson copula. For the pair (X_1, X_3), the fit based on the Ali-Mikhail-Haq copula has the lowest $D_C^{(2)}$ value, so should be chosen. For the pair (X_2, X_3), we get the best model using the Ali-Mikhail-Haq copula.

Pair		C_α^G	C_α^{CJ}	C_α^{AMH}	C_α^F
(X_1,X_2)	$\hat{\alpha}$	1.0739	0.1478	0.2863	0.4982
	$D_C^{(2)}$	0.3034	0.0743	0.0794	0.0838
(X_1,X_3)	$\hat{\alpha}$	1.0138	0.0276	0.0603	0.6862
	$D_C^{(2)}$	0.0551	0.0530	0.0526	0.0567
(X_2,X_3)	$\hat{\alpha}$	1.0245	0.0490	0.1047	0.6571
	$D_C^{(2)}$	0.2289	0.0455	0.0425	0.0533

Table 4: Values of $\hat{\alpha}$ and $D_C^{(2)}$ for the different pairs with respect to the various Archimedean copulas.

The second method is graphical. From the definition of a copula C, we know that if U and V are two uniform random variables on $[0,1]$ then the random variables:

$$C(V|U) = \frac{\partial C}{\partial U}(U,V)$$

and

$$C(U|V) = \frac{\partial C}{\partial V}(U,V)$$

are also uniformly distributed on $[0,1]$. Using the quantile transformation of X_i, we know that the distribution of $\hat{F}_i(X_i)$ is uniform on $[0,1]$, $i \in \{1,2,3\}$. For a fixed pair (X_i, X_j), the copula $C_{\hat{\alpha}}$ for which the distributions of $C_{\hat{\alpha}}(\hat{F}_j(X_j)|\hat{F}_i(X_i))$ and $C_{\hat{\alpha}}(\hat{F}_i(X_i)|\hat{F}_j(X_j))$ are uniformly distributed on $[0,1]$ would be the copula that best models the tail of the empirical joint distribution. This checking can be carried out by the way of the classical QQ-plot method.

As an example, we present our results for the pair (X_1, X_2). For each copula, we represent in Figure 2 the QQ-plot of the empirical distribution $C_{\hat{\alpha}}(\hat{F}_2(X_2)|\hat{F}_1(X_1))$ against the uniform distribution on $[0,1]$, and the more straight the line is, the best the fit of the tail of the empirical joint distribution of $(X_1, X_2)'$ by the copula $C_{\hat{\alpha}}$ is. The QQ-plots for $C_{\hat{\alpha}}(\hat{F}_1(X_1)|\hat{F}_2(X_2))$ are not presented since they are close to the QQ-plots for $C_{\hat{\alpha}}(\hat{F}_2(X_2)|\hat{F}_1(X_1))$. The Cook and Johnson copula provides the straightest line and then best models the dependence structure of $(X_1, X_2)'$, as we have concluded using the numerical criterion $D_C^{(2)}$. The same study has been derived for $(X_1, X_3)'$ and $(X_2, X_3)'$, and similar results have been obtained using either the graphical method or the numerical method.

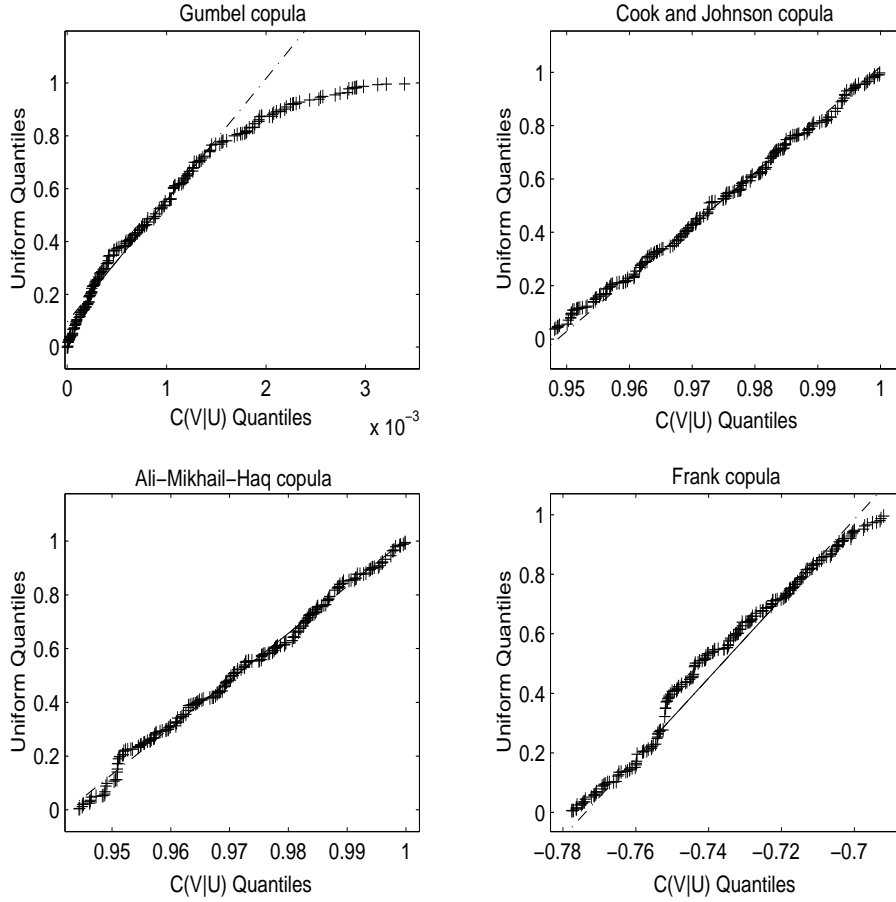On the basis of our analysis, we summarize the models that we would choose

Figure 2: QQ-plots of the empirical partial derivatives against the uniform distribution for $(X_1, X_2)'$ with respect to the various Archimedean copulas.

for the tails of the empirical joint distributions of the various pairs of series:

$$(X_1, X_2): \quad \hat{F}(x_1, x_2) = C_{0.1478}^{CJ}\big(\hat{F}_1(x_1), \hat{F}_2(x_2)\big), \quad x_1 > u_1, x_2 > u_2$$

$$(X_1, X_3): \quad \hat{F}(x_1, x_3) = C_{0.0603}^{AMH}\big(\hat{F}_1(x_1), \hat{F}_3(x_3)\big), \quad x_1 > u_1, x_3 > u_3$$

$$(X_2, X_3): \quad \hat{F}(x_2, x_3) = C_{0.1047}^{AMH}\big(\hat{F}_2(x_2), \hat{F}_3(x_3)\big), \quad x_2 > u_2, x_3 > u_3.$$

Now, we formulate an important remark. The concept of upper tail dependence relates to the amount of dependence in the upper-right-quadrant tail of the bivariate distribution. For an overview of this concept, we refer to Joe (1997). According to the $D_C^{(2)}$ criterion or the QQ-plot method, we have shown that the Gumbel copula performs worst for all the pairs (X_1, X_2), (X_1, X_3) and (X_2, X_3). This means that these series are not dependent in

the upper-right-quadrant tail of their bivariate distribution since the Gumbel copula has upper tail dependence. In contrast, the Cook and Johnson, Ali-Mikhail-Haq and Frank copulas have no upper tail dependence, and we recall that our results were good enough with these copulas. Thus, the choice of the copula to reconstruct the tail of the joint distribution is fundamental, and we can introduce a misspecification if we do not take care about this concept.

3.3 Extension to the multivariate case

A bivariate family of Archimedean copula can be extended in a natural way to a n-variate family of Archimedean copula, $n \geq 3$, under some constraints (e.g., Joe, 1997). First of all, to get this extension, we need that all bivariate marginal copulas of the multivariate copula belong to the given bivariate family. Secondly, we need that all multivariate marginal copulas of order 3 to $n - 1$ have the same multivariate form. In the following, we consider the trivariate case. We specify the constraints that permit such an extension and that deal with the dependence parameters.

Consider $n = 3$ and assume that the (i, j) bivariate marginal $(i \neq j \in \{1, 2, 3\})$ has dependence parameter $\alpha_{i,j}$. If $\alpha_1 \leq \alpha_2$ with $\alpha_1 = \alpha_{1,3} = \alpha_{2,3}$ and $\alpha_2 = \alpha_{1,2}$, then trivariate Archimedean copulas have the following form:

$$C_{\alpha_1,\alpha_2}(u_1, u_2, u_3) = \varphi_{\alpha_1}^{-1}\Big(\varphi_{\alpha_1} \ o \ \varphi_{\alpha_2}^{-1}\big(\varphi_{\alpha_2}(u_1) + \varphi_{\alpha_2}(u_2)\big) + \varphi_{\alpha_1}(u_3)\Big) \quad (8)$$

with $(u_1, u_2, u_3) \in [0, 1]^3$.

In the sequel, for two random variables X_1 and X_2, we denote by $\alpha(X_1, X_2)$ the dependence parameter deduced from Kendall's tau, denoted by $\tau(X_1, X_2)$, by means of the formula (2). For a random vector $X = (X_1, X_2, X_3)^{'}$ with joint distribution F and continuous marginal distributions F_1, F_2 and F_3, the expression (1), using (8), becomes for all $(x_1, x_2, x_3) \in \mathbb{R}^3$:

$$\begin{aligned} F(x_1, x_2, x_3) &= C_{\alpha_1,\alpha_2}\big(F_1(x_1), F_2(x_2), F_3(x_3)\big) \\ &= C_{\alpha_1}\Big(C_{\alpha_2}\big(F_1(x_1), F_2(x_2)\big), F_3(x_3)\Big) \quad (9) \end{aligned}$$

if $\alpha_1 \leq \alpha_2$ with $\alpha_1 = \alpha(X_1, X_3) = \alpha(X_2, X_3)$ and $\alpha_2 = \alpha(X_1, X_2)$.

Now, we apply this extension to the three series of MSCI daily log-returns X_1, X_2 and X_3 in order to model the tail of their empirical joint distribution $\hat{F}$ by means of trivariate Archimedean copulas.

As in the previous paragraph, we estimate the tail of the empirical distribution $\hat{F}_i$ of X_i using the threshold u_i that corresponds to the 95th sample

percentile, $i \in \{1, 2, 3\}$. Thus, we use the estimated parameters $\hat{\xi}$ and $\hat{\beta}$ given in Table 2 to get the tail estimator $\hat{F}_i$ of X_i for $x_i > u_i$, $i \in \{1, 2, 3\}$.

To be allowed to use (9), the dependence parameter of one pair of series defined in the tail has to be greater or equal than the two others which have to be equal. According to Table 4, for the Gumbel, Cook and Johnson and Ali-Mikhail-Haq copulas, the largest dependence parameter α is computed for the pair (X_1, X_2) and the parameters are not equal for the two other pairs (X_1, X_3) and (X_2, X_3). For the Frank copula, the largest parameter is computed for the pair (X_2, X_3). In spite of these results, we decide nevertheless to continue our empirical study to show how the expression (9) works.

For each of the four Archimedean copulas, we decide to set $\hat{\alpha}_1 = \hat{\alpha}(X_1, X_3)$ and $\hat{\alpha}_2 = \hat{\alpha}(X_1, X_2)$. We also do the same choice for the Frank copula in a view of being able to compare the results. Then, using these parameters, we build the trivariate Archimedean copula (9) for each bivariate family of Archimedean copula. To choose the trivariate copula $C_{\hat{\alpha}_1, \hat{\alpha}_2}$ that best models the tail of the empirical joint distribution $\hat{F}$ of the series X_1, X_2 and X_3, we derive an extension of the numerical criterion $D_C^{(2)}$ in the following way:

$$D_C^{(3)} = \sum_{x_1, x_2, x_3} \left| C_{\hat{\alpha}_1}\left(C_{\hat{\alpha}_2}\left(\hat{F}_1(x_1), \hat{F}_2(x_2)\right), \hat{F}_3(x_3)\right) - \hat{F}(x_1, x_2, x_3) \right|^2.$$

Then, the copula $C_{\hat{\alpha}_1, \hat{\alpha}_2}$ for which we get the lowest $D_C^{(3)}$ will be chosen as the best copula. For the various copulas, the quantities $D_C^{(3)}$ that we have computed are reported in Table 5.

	$C_{\alpha_1, \alpha_2}^{G}$	$C_{\alpha_1, \alpha_2}^{CJ}$	$C_{\alpha_1, \alpha_2}^{AMH}$	$C_{\alpha_1, \alpha_2}^{F}$
$D_C^{(3)}$	26.5131	2.1215	2.1410	2.2540

Table 5: Values of $D_C^{(3)}$ with respect to the various trivariate Archimedean copulas for $(X_1, X_2, X_3)'$.

According to the criterion $D_C^{(3)}$, we get the best model using the trivariate Cook and Johnson copula. On the basis of our analysis, we explicit the expression (9) for the tail of the empirical joint distribution of the three series X_1, X_2 and X_3:

$$\hat{F}(x_1, x_2, x_3) = C_{\hat{\alpha}_1}^{CJ}\left(C_{\hat{\alpha}_2}^{CJ}\left(\hat{F}_1(x_1), \hat{F}_2(x_2)\right), \hat{F}_3(x_3)\right)$$

for $x_1 > u_1$, $x_2 > u_2$, $x_3 > u_3$, with $\hat{\alpha}_1 = \hat{\alpha}(X_1, X_3)$ and $\hat{\alpha}_2 = \hat{\alpha}(X_1, X_2)$.

4 Conclusion

In this paper, we have performed models to fit the tail of the joint distribution of n components in a portfolio. We have successively considered the cases $n = 2$ and $n = 3$. These models have been built using Archimedean copulas and thanks to their easy relationship with Kendall's tau.

Archimedean copulas are worth studying since they allow for a great variety of different dependent structures. Unlike the family of elliptical copulas, they are not derived from multivariate distributions using Sklar's theorem. A positive consequence of this feature is that we do not have to take into account the marginal distributions. This point of view is not considered here, but will be investigated in a companion paper. However, a disadvantage of this feature is that extensions of bivariate Archimedean copulas to multivariate ones suffer from lack of free dependence parameters choice. In particular, some of these parameters are forced to be equal (see Paragraph 3.3 for the trivariate case). In spite of these technical conditions, it is important to remark that the generalization to the n-variate case, with $n \geq 3$, is possible without great difficulties. However, the generalization is very restrictive using elliptical copulas because in that case we need to use elliptical marginals.

Also, we have proposed two methods to decide which copula best models the tail of the joint distribution of n components in a portfolio. The first method is based on a numerical criterion and the second method is graphical and use QQ-plots.

Thanks to the notion of copula, our approach permits to obtain some information about the dependence between more than two markets, and this is of primary importance because most measures of dependence do not permit to get similar results.

References

Ali, M.M., N.N. Mikhail and M.S. Haq, 1978, A class of bivariate distributions including the bivariate logistics, Journal of Multivariate Analysis, 8, 405-412.

Avouyi-Dovi, S., D. Guégan and S. Ladoucette, 2002, What is the best approach to measure the interdependence between different markets?, Preprint 02-03, GRID, ENS Cachan.

Blum, P., A. Dias and P. Embrechts, 2002, The ART of dependence modelling: the latest advances in correlation analysis, in: M. Lane, Alternative

Risk Strategies, Risk Books, 339-356, London.

Cambanis, S., S. Huang and G. Simons, 1981, On the theory of elliptically contoured distributions, Journal of Multivariate Analysis, 11, 368-385.

Cook, R.D. and M.E. Johnson, 1981, A family of distributions for modelling nonelliptically symmetric multivariate data, Journal of the Royal Statistical Society, B, 43, 210-218.

Embrechts, P., 2000, The Bell curve is wrong: so what?, in: P. Embrechts, Extremes and Integrated Risk Management, Risk Waters Group, 25-28.

Embrechts, P., A. Hoeing and A. Juri, 2001a, Using copulae to bound the Value-at-Risk for functions of dependent risks, Preprint ETH, Zürich.

Embrechts, P., F. Lindskog and A. McNeil, 2001b, Modelling Dependence with Copulas and Applications to Risk Management, Preprint ETH, Zürich.

Embrechts, P., C. Klüppelberg and T. Mikosch, 1997, Modelling Extremal Events, Springer-Verlag, Berlin.

Frank, M.J., 1979, On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$, Aequationes Mathematicae, 19, 194-226.

Genest, C. and R.J. MacKay, 1986a, The joy of copulas: Bivariate distributions with uniform marginals, The American Statistician, 40, 280-283.

Genest, C. and R.J. MacKay, 1986b, Copules Archimédiennes et familles de lois bidimensionnelles dont les marges sont données, The Canadian Journal of Statistics, 14, 145-159.

Gumbel, E.J., 1958, Distributions à plusieurs variables dont les marges sont données, Comptes Rendus de l'Académie des Sciences Paris, 246, 2717-2719.

Hult, H. and Lindskog F., 2001, Multivariate extremes, aggregation and dependence in elliptical distributions, Preprint ETH, Zürich.

Joe, H., 1997, Multivariate models and dependence concepts, Monograph on Statistics and Applied Probability, Chapman and Hall, London.

Nelsen, R., 1999, An introduction to copulas, Springer, New York.

Rockinger, M. and E. Jondeau, 2001, Conditional dependency of financial series: an application of copulas, NER 82, Banque de France, Paris.

Scaillet, O., 2000, Nonparametric estimation of copulas for time series, Working Paper, Département des Sciences Economiques, Université Catholique de Louvain, Belgium.

Sklar, A., 1959, Fonctions de répartition à n dimensions et leurs marges, Publications de l'Institut de Statistique de l'Université de Paris, 8, 229-231.