Copulas In Finance Ten Years Later

Cécile Kharoubi-Rakotomalala, ESCP Europe, France
Frantz Maurer, KEDGE Business School, France

ABSTRACT

Copula functions are mathematical tools that have been used in finance for approximately ten years. Their main selling point is to separate the dependence function (copula) from the marginal distributions. A little over a decade after the rise of copula modelling in finance, this article provides an initial assessment of their application in financial contexts. More specifically, the main purpose of this paper is to contribute to an ongoing debate in the field: the choice of copulas. Through an empirical study of two composite stock indices (S&P 500 and CAC 40) daily returns over the period 2002-2011, we show that this methodological challenge is still unsolved. With this in view, we suggest a method that enables to capture implicitly the empirical dependence structure without assuming any specific parametric form for it.

Keywords: Copulas; Dependence Structure; Empirical Copula; Kernel Estimator

1.   INTRODUCTION

$\mathcal{O}$ ver the last decade or so, the copula functions have literally inundated the financial sector. The use of copulas, very quickly adopted by both academics and practitioners, has been criticized sometimes (Mikosch, 2006) and was recently made responsible for being the source of the subprime crisis (Mac Kenzie and Spears, 2012). The aim of this article is to review the extent of the developments achieved, barely ten years on from the introduction of copulas in finance[1]. It will help the non-specialist reader to understand what a copula is and comprehend its different fields of application, whilst also putting into perspective the comparative advantages of this tool. The interest of this article is twofold. Firstly, it highlights a fundamental aspect that has not been studied sufficiently in the literature: the choice of the copula. Also, and in view of the limits of the selection methods available, it suggests a procedure that tackles the following question: how to know which copula to use?

From a technical point of view, copulas allow the dependence structure of random variables to be modelled without assumptions on the parametric form of the marginal distributions. To keep things simple, copulas can be seen as an extension of correlation in a non-Gaussian universe. Within the dominant paradigm underlying classical financial theories, the behaviour of securities returns is modelled by a normal distribution. This hypothesis of multivariate normality postulates that (*i*) the variance is a measure of risk and (*ii*) the correlation between two asset returns is a dependence measure, thus determining the portfolio diversification and the hedging strategies performance. The multivariate normal distribution is interesting because all marginals are Gaussian, as assumed in a great number of well-established financial models, and the link between two random variables can be fully described only with the knowledge of the marginals and an additional parameter (rho), the Pearson's linear correlation coefficient.

However, since 1965 we know, thanks to the influential contribution of Eugene Fama, that securities returns are not Gaussian, but are instead characterized by asymmetry (*skewness* is different from zero) and a probability of extreme events more important than in the Gaussian case (*kurtosis* is greater than 3). Therefore, the multivariate normal distribution cannot be used to model the joint evolution of securities returns. Of course, the statistical literature provides a huge number of other multivariate distributions. However, all these distributions are

---

[1]The notion of copula was introduced by the mathematician Abe Sklar in 1959. In the field of finance, the research group led by Paul Embrechts (ETH Zurich) is a pioneer. In 1999, he published in *Risk Magazine* (together with Alexander McNeil and Daniel Straumann) the first article dedicated to the use of copula functions in risk management. Even today, it remains the most frequently quoted article.

of a limited use because each of them is a generalization or an immediate extension from its univariate counterpart. As a result, it is not straightforward to extend them beyond the bivariate framework. To put it differently, the marginals of the multivariate distribution belong to the same family as these univariate distributions. Therefore, it is not possible to have a distribution with an Inverse Gaussian margin, a Beta margin, a uniform margin, etc. Copulas allow this problem to be solved. They are a very powerful tool to build multidimensional distributions with *given* (parametric) or *observed* (empirical) marginals. As such, copulas are a very powerful tool in financial modelling when it comes to take the lack of normality in (log) returns and the dependence between extreme values of various assets into account (Genest et *al.*, 2009).

This article is divided into two main sections. The first one recalls some elementary facts about copulas, explains why copulas reveal to be a very powerful tool in financial modelling, and lists some prominent applications in the field of finance. The second section focuses on the difficulties relating to the choice of copulas and suggests a solution to implicitly capture the empirical dependence structure without assuming any specific parametric form for it.

2.   COPULAS AND MAJOR AREAS OF APPLICATION IN FINANCE

2.1   What is a Copula?

A bivariate copula **C** is a probability distribution with uniform marginals. Let U_1 and U_2 be two uniform random variables and **U** the random vector (U_1, U_2). We have **C** $(u_1, u_2) = \Pr \{U_1 \leq u_1, U_2 \leq u_2\}$. The Sklar's (1959) theorem specifies the link defined by the copula **C** (determined from the joint distribution **F**) between the univariate marginal cumulative distribution functions $\mathbf{F}_1$ and $\mathbf{F}_2$ and the bivariate distribution **F**. This theorem allows copula functions to be built from the bivariate distributions.

Sklar's Theorem. Let **F** be a two-dimensional distribution function whose marginals are $\mathbf{F}_1$ and $\mathbf{F}_2$. Then, **F** admits a copula representation:

$$\mathbf{F}(x_1, x_2) = \mathbf{C}\,(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)) \tag{1}$$

The copula **C** is unique if marginals are continuous. This theorem is very important, since we can associate a copula with each bivariate distribution (which may be unique). We therefore have a *canonical* representation of the distribution. On the one hand, the marginals $\mathbf{F}_1$ and $\mathbf{F}_2$, that is to say the one-dimensional directions. On the other hand, a copula, that links these marginals. As such, the copula defines the dependence between the one-dimensional directions. If the bivariate distribution is absolutely continuous, then it admits a density and we have:

$$f(x_1, x_2) = c\,(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)) \times f_1(x_1) \times f_2(x_2) \tag{2}$$

with $c(u_1, u_2)$ the density of the copula **C**.

In order to better understand eq. (2), let us consider a portfolio including two risk factors: IBM (x_1) and Google (x_2) stocks. In that case:

- $f(x_1, x_2)$ represents the joint density of the two risk factors for the portfolio, i.e. the simultaneous behaviour of the two IBM and Google stocks.

- $\left.\begin{array}{c} f_1(x_1) \\ f_2(x_2) \end{array}\right\}$ represents the marginal densities, i.e. the individual behaviour of each risk factor.

- $c\left(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)\right)$ represents the dependence between the two risk factors, i.e. how one varies in relation to each other.

Although many copula families are available, analytical tractability is a striking feature in the search for an appropriate form for the dependence structure. Parametric copulas (i.e., functions that depend on one or several parameters) are the most commonly used in biostatistics, actuarial science, or even finance. The Gaussian, Frank, Clayton (also called Cook-Johnson) and Gumbel copulas belong to this family. For example, they are of significant interest for risk management because they allow to build parametric or semi parametric models. Table 1 presents the aforementioned copulas, which were also chosen for their complementary characteristics in terms of tail dependence.

Table 1. Four Parametric Copulas (with $\tilde{u} = -\ln u$)

Copula[a]	Bivariate copula C(u_1, u_2)
Gaussian ($-1 \leq \rho \leq 1$)	$\Phi_\rho(\Phi^{-1}(u_1), \Phi^{-1}(u_2))$, where: ρ is the correlation coefficient, Φ_ρ is a bivariate normal distribution, Φ^{-1} is the inverse of the univariate standard normal distribution.
Frank ($\theta \neq 0$)	$-\dfrac{1}{\theta} \ln \left(1 + \dfrac{\left(e^{-\theta u_1} - 1\right)\left(e^{-\theta u_2} - 1\right)}{e^{-\theta} - 1} \right)$
Gumbel ($\theta \geq 1$)	$\exp\left(-\left(\tilde{u}_1^{\,\theta} + \tilde{u}_2^{\,\theta}\right)^{1/\theta}\right)$
Clayton ($\theta > 0$)	$\left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta}$

[a] Franck, Gumbel and Clayton copulas belong to the *Archimedean* copulas family (Genest and MacKay, 1986). The Gaussian copula is an *elliptical* copula. This table summarizes the four parametric copulas used in this study.

When the bivariate distributions are represented by a copula function, the tail dependence of the distribution is an asymptotic property of the copula. Reference can be made to Joe (1997) to define the tail dependence of the distributions and then we can study the asymptotic properties of the different copulas presented in Table 1. In Table 2, we consider these asymptotic properties. We talk about *upper tail* (*resp. lower tail*) dependence meaning that the positive (*resp. negative*) extreme returns are correlated. If $\lambda_U = 0$ (*resp. $\lambda_L = 0$*) then the C copula does not show any *upper tail* (*resp. lower tail*) tail-dependence.

Table 2. Tail Dependence

	Upper Tail Dependence	*Lower Tail* Dependence	Conditions
	λ_U	λ_L	
Gaussian	0	0	Asymptotic independence if $-1 < \rho < 1$
Frank	0	0	Asymptotic independence
Gumbel	$2 - 2^\theta$	0	Asymmetric upper dependence if $\theta < 1$
Clayton	0	$2^{-1/\theta}$	Asymmetric lower dependence if $\theta > 0$

Tail dependence of a bivariate copula measures the probability of simultaneous extreme events. Table 2 shows that the Gaussian copula has no tail-dependence (except in the case of a perfect positive correlation, i.e. $\rho = 1$). Thus, even with a correlation coefficient equal to 90%, extreme returns are asymptotically independent. In other words, the Gaussian copula does not allow for the correlation of extreme values. The Gumbel copula has an *upper tail* dependence but no *lower tail* dependence (upper asymmetric dependence). It is the opposite for the Clayton copula (lower asymmetric dependence). Figure 1 graphs the densities of the four copulas considered in this article. One can visualize the asymptotic characteristics of the four distinct dependence structures relating to Tables 1 and 2. The four copulas parameters of Figure 1 are calibrated using the method of moments, which will be presented later.
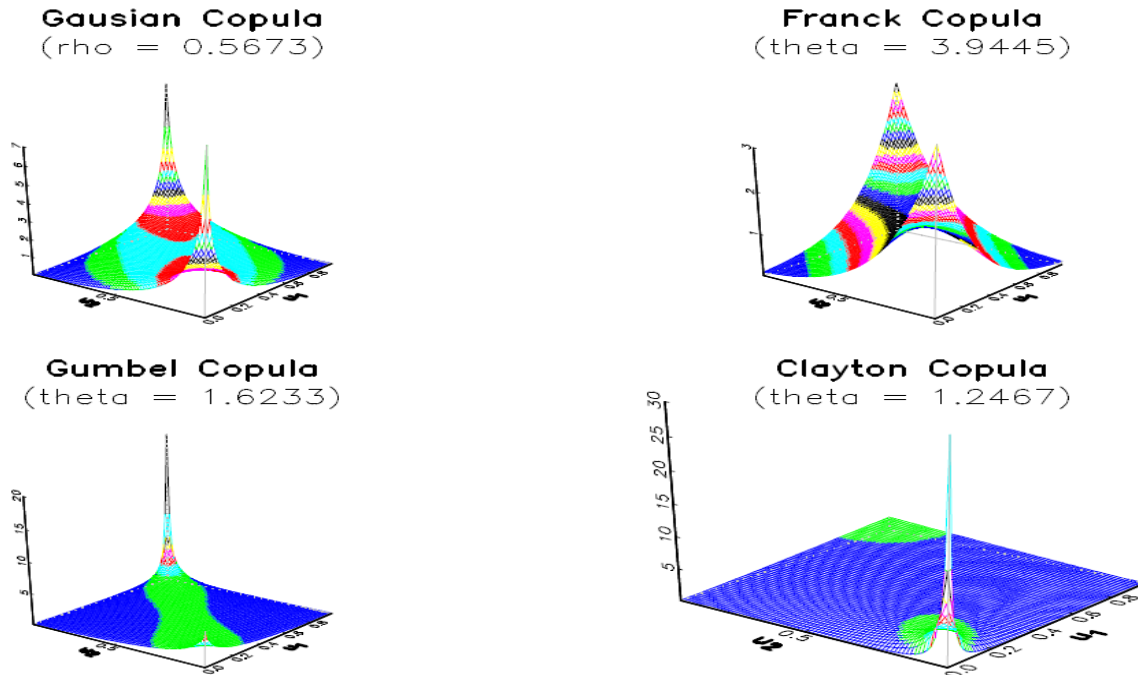
Figure 1. **Four Bivariate Copula Density Functions**
Note: To illustrate the diversity of shapes obtained with the parametric copulas presented in Table 1 and used in this study, we show here the bivariate copula density functions obtained with the parameters estimated (through the method of moments) on the joint distribution S&P 500–CAC 40 daily returns. The estimated parameters values into brackets are displayed in Table 3.

2.2   Correlation and Copulas: What is the Difference?

The correlation coefficient is a canonical measure of dependence within the multivariate Gaussian framework, and more generally, for the elliptical distributions. However, empirical research in finance has shown that empirical distributions of stock returns do not belong to this family of distributions. Thus, using the linear correlation coefficient in a non-Gaussian universe is theoretically inappropriate. As noted by Embrecht *et al.* (1999), "one does not have to search far in the literature of financial risk management to find misunderstanding and confusion about correlation" wrongly used as a dependence measure outside of its legitimate framework. This often comes from confusion between the concepts of correlation and dependence. When the variables are independent, the linear correlation coefficient is zero, but the reciprocal is wrong. The correlation may thus be zero whilst the random variables are not independent. The inverse relationship is only true within the multivariate Gaussian framework. This is no longer the case when the marginal distributions are Gaussian and the joint distribution is not Gaussian[2].

The linear correlation presents other pitfalls. For example, it is not invariant under non-linear increasing transformations. That is undoubtedly its most salient shortcoming. From a financial point of view, this means that the correlation is not a coherent measure of dependence. Let us consider two vectors of stock returns (X and Y) and two vectors of the logarithmic returns of X and Y. Here, the transformation applied is the logarithm function. We can compute the linear correlation coefficient of the two vectors of prices. Of course, we can do the same with the two vectors of logarithmic returns. Both coefficients obtained are not equal, although the informational content is identical for the two pairs of vectors. Unlike the linear correlation, the dependence structure conveyed by a copula function C is preserved under non-linear increasing transformations[3]. Let $T_1, T_2, \ldots, T_n$ denote a set of increasing functions, then the vectors $(X_1, X_2, \ldots, X_n)$ and $(T_1(X_1), T_2(X_2), \ldots, T_n(X_n))$ have the same structure of dependence

---

[2] For example, let $X \sim N(0, 1)$ be a random variable following a standard Gaussian distribution. If $Y = X^2$, then the covariance between X and Y equals zero, whilst X and Y are perfectly dependent.
[3] The complete proof is established in Nelsen (2006).

1558   *2013 The Clute Institute*

C. Stated differently, the degree of dependence is the same when considering price vectors and vectors of logarithmic returns.

Additional difficulty, the correlation coefficient is defined in relation to volatility. This means that it depends directly on the level of volatility. It does not allow the dependence among random variables to be studied regardless of volatility. Yet, many empirical studies have shown that volatility fluctuated stochastically. Unlike the linear correlation coefficient, dependence functions are not determined by the volatility of the random variables considered. This has two notable consequences. On the one hand, with copula functions, it becomes possible to study the variations of co-movements without heteroskedastic bias. On the other hand, as we mentioned it previously, multivariate distributions allow isolating marginal behaviours from dependence itself, suggesting a two stage statistical procedure: estimate the marginal distributions and the copula function separately from each other. As a result, even in the case of heavy-tailed univariate distributions, the dependence measure obtained through the copula parameter can still be defined. However, this does not apply to the correlation coefficient whose univariate densities must have finite variance in order to be defined. To summarize, the great interest of copula functions is to overcome the aforementioned difficulties: they generate all multivariate distributions with flexible marginals.

2.3   Finance-Related Applications of Copulas

The applications of copula in finance are mainly in risk management and option pricing, even though other applications have been proposed, such as portfolio management and pricing derivatives. Whenever the Gaussian multivariate distribution is used "by default", copulas allow modelling to be improved. In the field of risk management, Embrecht et *al.* (1999) showed that the worst-case scenarios of Value-at-Risk (VaR) were not those obtained for the maximum value of the linear correlation coefficient, i.e. when the random variables are comonotonic. Cherubini and Luciano (2001) use Archimedean copulas to investigate the tail behavior of distributions and VaR. However, they do not include in their study the Gaussian copula and are unable to select an optimal copula. Rosenberg and Schuermann (2006) use the copula modelling to aggregate market risk, credit risk and operational risk. Copulas have also been applied to the default risk modelling. In this area, the paper on default correlation by Li (2000) is worth mentioning. He introduces a random variable called "time-until-default" to denote the survival time of each financial instrument, and define the default correlation between two sources of credit risks as the correlation coefficient between their survival times.

Similar to risk management, the pricing of options on several underlying assets has been enriched thanks to copulas. For example, Rosenberg (2003) develops a non- parametric method for evaluating options on two underlying assets. The non-parametric estimation of the risk-neutral multivariate density is grounded on option prices traded on the market. The non-parametric dependence function is then estimated by using historical returns. However, he does not deal with the problem of the risk-neutral copula estimation. Cherubini and Luciano (2002) duplicate the study of Rosenberg and use conventional parametric copula functions, but the problem of the risk-neutral copula is totally ignored. Finally, Coutant et *al.* (2001) use copulas to define risk-neutral multivariate distributions, and in particular, the risk-neutral copula, enabling them to derive formulas for evaluating options with several underlying assets.

Malevergne and Sornette (2003) use a copula representation to redefine the portfolio selection theory. However, they select a Gaussian copula – on which the conventional approach of portfolio theory is grounded – to develop their theoretical framework. Finally, some researchers use copulas to study how the dependence structure of financial assets changed over time (e.g. Longin and Solnik, 2001; Poon et *al.*, 2004; Patton, 2004). In this aim, they model the co-movements of financial markets with copula functions.

Very few authors were interested in the choice of the "optimal" copula, i.e. the copula that will converge to the real dependence structure underlying the data. Notable exceptions are Durrleman et al. (2000), Ané and Kharoubi (2003), Kole et *al.* (2007). As a matter of fact, the vast majority of published articles in peer-review journals either uses the Gaussian copula *de facto*, or provides selection methods that do not apply to the Gaussian copula. This can lead to biased results because the copula function includes all information about dependence. Therefore, the choice of the copula that is going to fit the data is very important.

3.   IN SEARCH OF THE "OPTIMAL" COPULA

Several graphical tools are available to assess the goodness-of-fit of a copula to the dependence structure of two random variables. Unfortunately, they cannot always be easily interpreted. In this paper, we examine the joint distribution of the S&P 500–CAC 40 pair of daily returns over the period January 1, 2002, through December 31, 2011[4]. Data are extracted from the Datastream (Thomson-Reuters) database and correspond to daily closing prices. We will consider the four parametric copulas displayed in Table 1. In order to select which copula is the best one – the so-called "optimal" copula – we have to calibrate each one of them. For the purpose of visualizing the material we will be working with, Figure 2 graphs the density of the empirical joint distribution of the S&P 500–CAC 40 stock index daily returns. It is reconstructed through the bivariate Gaussian kernel method.
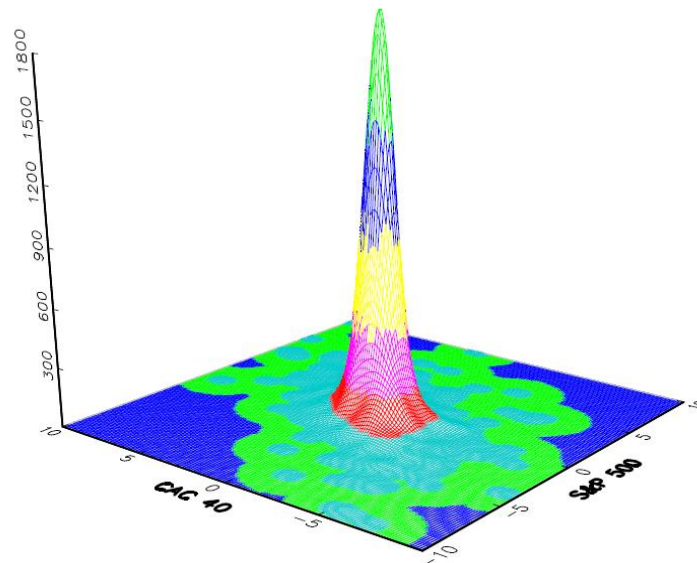

Figure 2. Empirical Joint Distribution of the S&P 500–CAC 40 Daily Returns

3.1   Estimation of a Parametric Copula

Several methods of statistical inference are available for tackling this problem: the method of the maximum likelihood, the two-step IFM (*inference from margins*) method (Shih and Louis, 1995; Joe and Xu, 1996), the *omnibus* procedure (Genest *et al.*, 1995; Shih and Louis, 1995). One can also estimate the parameter of a copula such that it fits the concordance measures *Kendall*'s tau (denoted as τ) and *Spearman*'s rho (denoted here as ζ), two well-known quantities based on ranks which are considered as some kind of dependence measures[5]. This is the so-called method of moments. Only this method is considered in the sequel. This methodological choice is justified by the fact, highlighted by Genest and Favre (2007), that the dependence structure captured by a copula has nothing to do with the individual behaviour of the variables. These authors argue that the ranks of the observations are the best summary of the joint behaviour of random pairs. It would make sense, therefore, to rely any inference about the parameter of a copula on rank-based procedures.

---

[4] The total number of observations is equal to 2,608.

[5]The estimator of Kendall's tau is $\hat{\tau} = (c - d)/(c + d)$ where c and d are respectively the number of concordant and discordant disjoint pairs. Roughly speaking, Spearman's rho corresponds to the linear correlation of the rank statistics.

In some sense, Kendall's tau and Spearman's rho are a generalization of the linear correlation in the case of a non-Gaussian bivariate distribution. Both can be defined in terms of a copula, with the following relationships (Nelsen, 2006, p. 164-167):

$$\tau = 1 - 4 \iint_{[0,1]^2} \partial_{u_1} \mathbf{C}(u_1, u_2) \, \partial_{u_2} \mathbf{C}(u_1, u_2) \, du_1 \, du_2 \quad \text{(Kendall's tau)} \tag{3}$$

$$\varsigma = 12 \iint_{[0,1]^2} \mathbf{C}(u_1, u_2) du_1 \, du_2 - 3 \quad \text{(Spearman's rho)} \tag{4}$$

For one-parameter bivariate copulas, such as those of Table 1, analytical solutions linking Kendall's tau and the copula parameter are available. For the Gaussian copula, for instance, the relationship between the parameter to be estimated (ρ) and the Kendall's τ is given by: $2\pi^{-1}\arcsin(\rho)$. Then we have the following estimate $\hat{\rho}$ of the Gaussian copula parameter: $\hat{\rho} = \sin\left(\dfrac{\pi}{2}\hat{\tau}\right)$.

The relationship with Spearman's ζ is given by: $6\pi^{-1}\arcsin(\rho/2)$. Thus we have: $\hat{\rho} = 2\sin\left(\dfrac{\pi}{6}\hat{\varsigma}\right)$. With the daily returns of the S&P 500–CAC 40 pair, we obtain $\hat{\tau} = 38.39\%$ and $\hat{\varsigma} = 52.51\%$, so it comes that $\hat{\rho}(\hat{\tau}) = 56.73\%$ and $\hat{\rho}(\hat{\varsigma}) = 54.31\%$ for the Gaussian copula, as indicated in Table 3. The estimated parameters for each copula function are displayed in Table 3.

Table 3. Estimated Parameters of the Copula Functions

	Copula	$\hat{\theta}_{\mathbf{C}}(\hat{\tau})$	$\hat{\theta}_{\mathbf{C}}(\hat{\varsigma})$
C_1	Gaussian	0.5673	0.5431
C_2	Frank	3.9445	3.6814
C_3	Gumbel	1.6233	✓
C_4	Clayton	1.2467	✓

Note: All parameters are estimated through the method of moment as explained in section 3.1. The estimated parameters for each copula considered in this paper and the S&P 500–CAC 40 pair of daily returns are summarized in this table.

3.2   Which Copula Provides the Best Fit to the Observations?

Among the graphical diagnostics available to assess the goodness-of-fit of a copula to the real underlying dependence structure, some apply only to the family of Archimedean copulas. As a result, they exclude the Gaussian copula. This is the case of the methods developed by Genest and Rivest (1993) and Frees and Valdez (1998). That is the reason why these procedures will not be presented here. The Gaussian copula underlies the main theoretical financial models, and that is why this copula serves as a benchmark dependence structure (e.g. Junker and May, 2005).

Durrleman et *al.* (2000) suggest a selection criterion allowing a specific class of copula alternatives – including the Gaussian copula – to be compared. To test the adequacy of a specific dependence structure, they compute the distance between each class of copula alternatives and the empirical copula introduced by Paul Deheuvels in 1979. This distance is defined on the discrete L^p norm as follows:

$$\bar{d}_2\left(\hat{\mathbf{C}}_{(T)}, \mathbf{C}_k\right) = \left\| \hat{\mathbf{C}}_{(T)}, \mathbf{C}_k \right\| L^2$$

$$= \left(\sum_{t_1=1}^{T} \cdots \sum_{t_n=1}^{T} \cdots \sum_{t_N=1}^{T} \left[\hat{\mathsf{c}}_{(T)}\left(\frac{t_1}{T},\ldots,\frac{t_n}{T},\ldots,\frac{t_N}{T} \right) - \mathsf{c}_k\left(\frac{t_1}{T},\ldots,\frac{t_n}{T},\ldots,\frac{t_N}{T} \right) \right]^2 \right)^{\frac{1}{2}} \tag{5}$$

where T denotes the order of the empirical copula. The optimal copula is defined as the copula which minimizes the distance $\bar{d}_2(\hat{\mathsf{c}}_{(T)}, \mathsf{c}_k)$. We are interested in identifying which one of the four copulas considered in Table 1 (Gaussian, Frank, Gumbel and Clayton copulas) fits best the empirical dependence structure of the (S&P 500, CAC 40) daily returns data. We have reported in Table 4 the distance measures for these copulas.

Table 4. Distance Measures for the Parametric Copula Functions

	Copula	$\bar{d}_2(\hat{\mathsf{c}}_{(T)}, \mathsf{c}_k)/T$
C_1	Gaussian	7.01×10^{-3}
C_2	Frank	9.87×10^{-3}
C_3	Gumbel	8.34×10^{-3}
C_4	Clayton	12.66×10^{-3}

Note: We put in bold characters the best or "optimal" copula. In order to select among the different copulas, we used the distance between each parametric copula and the Deheuvels (1979) or empirical copula, as explained in eq. (5). This table gives a goodness-of-fit measure for the estimated copula functions. T denotes the empirical copula order.

Table 4 indicates that the Gaussian copula is the "optimal" copula, i.e., the closest to the empirical copula. This result seems to be inconsistent with a few studies published in this area. However, it should be mentioned that the markets or periods of time observed were different. Ané and Kharoubi (2003) examine the returns of five composite stock indices over the period January 2, 1987, through December 31, 2000. The selection procedure they applied on fifteen pairs of bivariate returns, identifies the Clayton copula as the best choice in all cases. Another empirical study − also based on the Deheuvels or empirical copula − clearly supports the Student copula, and rejects both the Gaussian and Gumbel copulas (Kole *et al.*, 2007). Durrleman *et al.* (2000) examine the dependence between the spot prices of Aluminum (AL) and the 15 months forward prices (AL-15). They conclude that the Gaussian copula does not fit the empirical copula properly, and that the Frank copula is the optimal copula to capture the underlying dependence of the (AL, AL-15) returns vector. However, Malevergne and Sornette (2003) underlines that the Gaussian copula hypothesis cannot be rejected in a number of financial series, such as composite stock indices and exchange rates.

In the next section, we argue that grounding the choice of copulas on the distance to the empirical copula can be misleading. Yet, this choice is essential, especially in the field of risk management, where it determines dependence between one-dimensional risk factors, such as interest rates, commodity prices or FX rates.

3.3 Reconstructing the Empirical Dependence Structure through the Inverted Kernel Method

The selection procedure of the "optimal" copula based on eq. (5) requires an estimated parameter from each alternative copula. The method that we suggest here adopts a different approach. It consists of deriving the empirical density of a bivariate distribution by inverting the kernel method. The optimal copula is then identified by comparing the empirical density (estimated non-parametrically from the kernel method) with the theoretical density in three dimensions, or in the form of level curves. The goal is to capture implicitly the empirical dependence structure without assuming any specific parametric form on it. In the univariate case, the nonparametric estimator of the density function obtained with the kernel method is defined by (Silverman, 1986, p. 15):

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h} \right) \tag{6}$$

1562   *2013 The Clute Institute*

where h is the window width and K the kernel. In particular, the univariate Gaussian kernel is defined by:

$$K(t) = \frac{1}{\sqrt{2\pi}} \exp^{-(1/2)t^2}.$$
(7)

In the multivariate case, this estimator becomes (Silverman, p. 76):

$$\hat{f}(\mathbf{x}) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left\{\frac{1}{h}(\mathbf{x} - \mathbf{X}_i)\right\}.$$
(8)

The kernel function K(**x**) is now a function, defined for d-dimensional **x**, satisfying: $\int_{R^d} K(\mathbf{x})d\mathbf{x} = 1$. The multivariate Gaussian kernel is then defined by:

$$K(\mathbf{x}) = (2\pi)^{-d/2} \exp\left(-\frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{x}\right).$$
(9)

As we have seen, where the Sklar's (1959) theorem is used, the density f of a two-dimensional distribution can be written according to the density c of the associated copula **C** and the marginal densities f_1 and f_2, such as:

$$f(x_1, x_2) = c(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)) \times f_1(x_1) \times f_2(x_2)$$
(10)

with $c(u_1, u_2)$ the density of copula **C**. From the relationship linking the bivariate density $f(x_1, x_2)$ and the univariate densities $f_1(x_1)$ and $f_2(x_2)$, one can calculate the expression for the density of the copula:

$$c(u_1, u_2) = \frac{f\left(\mathbf{F}_1^{-1}(u_1), \mathbf{F}_2^{-1}(u_2)\right)}{f_1\left(\mathbf{F}_1^{-1}(u_1)\right) \times f_2\left(\mathbf{F}_2^{-1}(u_2)\right)}.$$
(11)

Thanks to the inverted kernel method, one can obtain a nonparametric estimation of the dependence structure. It proceeds in two steps: (*i*) estimating the numerator of equation (11) through the bivariate Gaussian kernel, and (*ii*) estimating each component of the product at the denominator of equation (11) through the univariate Gaussian kernel. Figure 3 graphs the resulting empirical dependence of the S&P 500–CAC 40 daily returns.
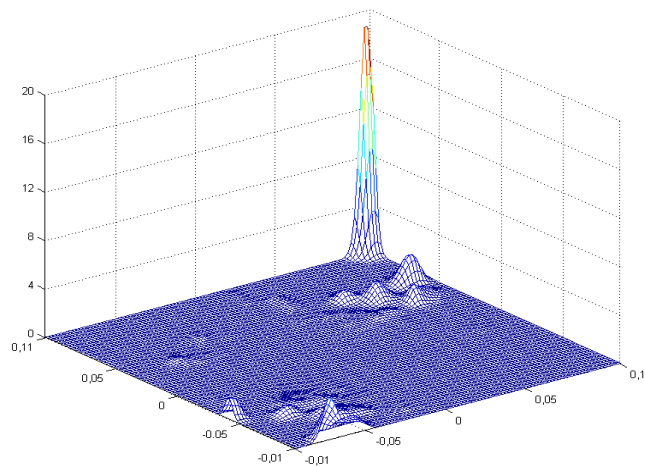

Figure 3. **Empirical Dependence Structure of the S&P 500–CAC 40 Daily Returns**
Gaussian Kernel ($h = 1.9933$)

In order to identify the "optimal" copula for modeling the dependence structure of the pair S&P 500–CAC 40, we then compare its empirical density (Figure 3) – reconstructed through the inverted Gaussian kernel method – with the theoretical density (Figure 1) of each alternative parametric copula considered in this paper. This provides a powerful testing procedure against the specific class of copula alternatives of Table 1. Through a visual check, there is no difficulty to identify the density of the alternative theoretical copula (Figure 1) whose shape is closest to that of the empirical dependence structure (Figure 3). It appears that the Gumbel copula is the "best" or "optimal" copula for modeling the empirical density of the S&P 500–CAC 40 daily returns. It may be worth mentioning that the criterion selection based on the distance to the Deheuvels (1979) or empirical copula, as defined by equation (5), led to a different result and argued for the Gaussian copula. Yet, as highlighted earlier, these two copulas are very different in terms of their tail dependence. In all fairness, it should be noted, however, that the nonparametric estimation of a density function through the kernel method is very sensitive to the window width h used (see Silverman, p. 15-18, for very convincing graphical proofs). The effect of varying the bandwidth h may be significant. This constant acts as a smoothing parameter. If h is chosen too small then spurious fine structure becomes visible, while if h is too large then main features of the actual density structure may be obscured.

4.   CONCLUSION

Identifying the so-called "optimal" copula that provides the best fit to the data at hand still poses serious difficulties. The results obtained for a pair of composite stock indices daily returns, namely the S&P 500 and CAC 40 – observed over a ten-year period from 2002 to 2011 – show that there is still no systematic rigorous method for the choice of copulas. However, the study by Berg (2009) demonstrates that strategies for tackling this problem should be based (sometimes not directly) on the Deheuvels (1979) or empirical copula.

Praised for their great flexibility when it comes to implement multivariate analysis, copula functions have been applied to an eclectic mix of subjects in the field of empirical finance. When doing copulas selection, it is recommended to examine various diagnostic tests such as goodness-of-fit plots. This may give valuable information on the fit of a copula. In this paper, we provide an intuitive and informative method to identify the so-called "optimal" copula.

Finally, a word of warning: The choice of a copula should be related to the problem at hand and not on mathematical convenience. As emphasized by Thomas Mikosch, in a paper published in 2006 in the journal *Extremes*, a multivariate non-Gaussian distribution with Gaussian marginal distributions is usually considered pathological and does not have a great practical interest. However, such a distribution corresponds to a copula. In the footsteps of Mikosch, one wonders why this copula would not be "pathological" as well.

AUTHOR INFORMATION

Cécile Kharoubi-Rakotomalala is Professor in the finance department at ESCP Europe Paris Campus. After graduating from Paris Dauphine and Sorbone Universities, she obtained a Doctorate from dauphine and ESSEC Business School in 2003. She joined ESCP Europe in September 2004 and created a financial market specialization. Her research focuses on modeling and managing financial risks, and the hedge funds industry. She is an author of many articles on these topics. Contact Information: ESCP Europe, 79 avenue de la République, 75011 Paris (France). Telephone: +33 (0)1 49 23 21 57. E-mail: ckharoubi@escpeurope.eu

Frantz Maurer is an Associate Professor at KEDGE Business School and an Affiliate Professor in the finance department at ESCP Europe Paris Campus. He holds a Doctorate in management Science and a post-doctoral degree from the University of Bordeaux. Frantz is the author of several publications directed to academics and practitioners on risk management. His current research interests focus primarily on risk measures for corporates and Enterprise

Risk Management for financial institutions. Contact Information: KEDGE Business School, 680 cours de la Libération, 33405 Talence cedex (France). Telephone: +33 (0) 556 848 577. E-mail: fmaurer@escpeurope.eu (Corresponding author)

REFERENCES

1. Ané, T., & Kharoubi, C. (2003). Dependence Structure and Risk Measure. *Journal of Business*, 76 (3), 411-438.
2. Berg, D. (2009). Copula Goodness-of-fit Testing: an Overview and Power Comparison. *The European Journal of Finance*, 15 (7-8), 675-701.
3. Cherubini, U., & Luciano, E. (2001). Value-at-Risk Trade-off and Capital Allocation with Copulas. *Economic Notes*, 30, 235–256.
4. Cherubini, U., & Luciano, E. (2002). Bivariate Option Pricing with Copulas. *Applied Mathematical Finance*, 9 (2), 69-85.
5. Coutant, S., Durrleman, V., Rapuch, G., & Roncalli, T. (2001). Copulas, Multivariate Risk - Neutral Distributions and Implied Dependence Functions. Working paper, Groupe de Recherche Opérationnelle, Crédit Lyonnais.
6. Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés − Un test non paramétrique d'indépendance. *Académie Royale de Belgique − Bulletin de la Classe des Sciences − 5e Série*, 65, 274-292.
7. Durrleman, V., Nikeghbali, A., & Roncalli, T. (2000). Which Copula Is the Right One?. Working paper, Groupe de Recherche Opérationnelle, Crédit Lyonnais.
8. Embrechts, P., McNeil, A., & Straumann, D. (1999). Correlation: Pitfalls and Alternatives. *Risk Magazine*, 12 (5), 69-71.
9. Fama, E. (1965). The Behavior of Stock Market Prices », *Journal of Business*, 34–105.
10. Frees, E.W., & Waldez, E.A. (1998). Understanding Relationships Using Copulas. *North American Actuarial journal*, 2 (1), 1-25.
11. Genest, C., & Favre, A.-C. (2006). Everything You Always Wanted to Know About Copula Modelling but Were Afraid to Ask. *Journal of Hydrologic Engineering*, 12, 347-368.
12. Genest, C., Gendron, M., & Boudeau-Brien, M. (2009). The Advent of Copulas in Finance. *The European Journal of Finance*, 15, 609-618.
13. Genest, C., & MacKay, J. (1986). Copules Archimédiennes et familles de lois bidimensionnelles dont les marges sont données. *Canadian Journal of Statistics*, 14, 145-159.
14. Genest C., & Rivest, L.-P. (1993). Statistical Inference Procedure for Bivariate Archimedean Copulas. *Journal of the American Statistical Association*, 88, 1034-1043.
15. Genest, C., Ghoudi, K., & Rivest, L.-P. (1995). A Semiparametric Estimation Procedure for Dependence Parameters in Multivariate Families of Distributions. *Biometrika*, 82, 543-552.
16. Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Monographs on Statistics and Applied Probability, 73, London, Chapman and Hall.
17. Joe, H. & Xu, J.J. (1996). The Estimation Method of Inference Functions for Margins for Multivariate Models. Technical report, 166, Department of Statistics, University of British Columbia.
18. Junker, M. & May, A. (2005). Measurement of Aggregate Risk with Copulas. *Econometrics Journal*, 8, 428-454.
19. Kole, E., Koedijk, K., & Verbeek, M. (2007). Selecting Copulas for Risk Management. *Journal of Banking & Finance*, 31, 2405-2423.
20. Longin, F., & Solnik, B. (2001). Extreme Correlation of International Equity Markets. *Journal of Finance*, 649–676.
21 Li, D. (2000). On Default Correlation: a Copula Function Approach. *Journal of Fixed Income*, March (4), 43-54.
22. MacKenzie, D. & Spears, T. (2012). The Formula that Killed Wall Street? The Gaussian Copula and the Material Cultures of Modeling. Working paper, University of Edinburgh.
23. Malevergne, Y., & Sornette, D. (2003). Testing the Gaussian Copula Hypothesis for Financial Assets Dependences. *Quantitative Finance*, 3 (4), 231-250.
24. Mikosch, T. (2006). Copulas: Tales and Facts (with discussion). *Extremes*, 9 (1), 3-20.
25. Nelsen, R.B. (2006). *An Introduction to Copulas*. Springer Series in Statistics, 2nd ed., Springer.

26.   Patton, A. (2006), Modeling Asymmetric Exchange Rate Dependence. *International Economic Review*, 47 (2), 527-556.
27.   Poon, S., Rockinger, M., & Tawn, J. (2004). Extreme Values Dependency in International Stock Markets ». *Review of Financial Studies*, 17, 581-610.
28.   Rosenberg, J. (2003). Nonparametric Pricing of Multivariate Contingent Claims. *Journal of Derivatives,* Spring, 9-26.
29.   Rosenberg, J. & Schuermann, T. (2006). A General Approach to Integrated Risk Management with Skewed, Fat-tailed Risks. *Journal of Financial Economics*, 79, 569-614.
30.   Shih, J.H., & Louis, T.A. (1995). Inferences on the Association Parameter in Copula Models for Bivariate Survival Data. *Biometrics*, 51, 1384-1399.
31.   Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. Monographs on Statistics and Applied Probability, 26, Chapman & Hall.
32.   Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistiques de l'Université de Paris*, 8, 229-231.