# Calibrating and Simulating Copula Functions in Financial Applications

*Annalisa Di Clemente\* and Claudio Romano*

*Department of Social and Economic Sciences, Sapienza University of Rome, Rome, Italy*

Copula functions can be utilized in financial applications to determine the dependence structure of the financial asset returns in the portfolio. Empirical evidence has proved the inadequacy of the multi-normal distribution, traditionally adopted to model the financial asset returns distribution. Copula functions can be employed in a flexible way for building efficient algorithms and to simulate a more adequate distribution of the financial assets. This paper aims to describe some simple statistical procedures currently employed to calibrate the copula functions to the financial market data. Furthermore, we present some useful methods for choosing which copula function better fits the real financial data. Also, some algorithms to simulate random variates from certain types of copula functions are illustrated. Finally, for illustration purposes, the previous procedures described are applied to two Italian equities. In particular, we show how to generate efficient Monte Carlo scenarios of equity log-returns in the bivariate case using different types of copula functions.

**Keywords: Copula functions, dependence structure, multivariate distribution function, extreme joint co-movements, financial applications, stock market**

## INTRODUCTION

The study of the copula functions is very relevant because of their implementation in the field of financial portfolio risk management. The copula functions are used in financial applications since 2000, following the seminal researches of [1, 2]. The crucial matter is the real distribution of financial data. Empirical evidence has widely proved that the multinormal distribution is inadequate to model portfolio's financial asset returns distribution at least from two points of view:

(1) The empirical marginal distributions are skewed and fat-tailed.
(2) The normal distribution does not consider the possibility of extreme joint co-movements for financial asset returns.

In other words, the real dependence structure of the financial assets is different from the Gaussian one and especially under situations of market stress [3]. For this reason, the copula functions can be a useful and simple tool for implementing efficient algorithms and to simulate the financial asset returns distribution more realistically.

The copulas allow us to model the dependence structure independently from the marginal distributions. In this way, we may construct a multivariate distribution with different margins and the dependence structure given from a particular type of copula function. Therefore, a crucial step in this context is the choice and the calibration of the most adequate copula function from the real financial data.

In this paper, a group of useful methods for calibrating, selecting, and simulating copula functions is presented. We aim to collect and to describe in a very simple manner the principal contributions in this field provided by the most famous and accredited international literature (for example [4–18]). In our opinion, the study of copula functions is very important because these quantitative tools are valuable for developing advanced financial portfolio models. Potential applications are in the field of market, credit, and operational risks. In particular, the copula functions can model the dependence structure between risk factors (for example, equity returns, interest rate returns, and foreign exchange returns) in a more reliable way.

In particular, the credit assets clearly show a non-normal return distribution and the phenomena of asymmetry, leptokurtosis, and tail dependence. To take into account these empirical characteristics of credit asset return distributions, with particular regard to multiple default events, we may model the default dependence structure by using different types of copulas. Technically, the times until default of each obligor in the loan portfolio may be simulated following a copula-based approach first illustrated in Li [1].

The copula functions have been also applied in the field of portfolio credit asset allocation where the optimal portfolio composition [19] may change by utilizing various types of the copula. The crucial aspect is the choice of the kind of copula better describing the default dependence structure of credit assets in a portfolio (see, for example [20]).

Further applications of copula functions in the financial risk management have been also developed by [21–25]. Nowadays, the copula functions find a precious application also in the estimate of the systemic risk generated by the Systemically Important Financial Institutions (see, for example [26–28]).

The implications in the field of the macro-prudential regulation on the global financial system are evident [29, 30].

Finally, in this paper, the described statistical methods for calibrating, selecting, and simulating copula functions are implemented to an empirical financial data set concerning the log-returns of two Italian equities: Olivetti and TIM. When it is possible, we show as the copula approach performs better than the bivariate normal distribution in modeling the real financial data.

The paper is structured as follows. In *Definition of the Copula Function* Section, a brief definition of copula function is given, describing the main families of copulas utilized in financial practical applications[1]. In *Parameter Estimation of a Given Copula* Section, some quantitative approaches to estimate the parameters of a determined copula function from real data are presented. The procedures for selecting the type of copula function which better fits empirical data are shown in *Selecting the Right Copula* Section. In *Simulation Algorithms* Section, the algorithms to simulate random variates from some types of copula are illustrated. An application to a time series of the log-returns for two Italian equities is performed in *Application to Two Italian Equities* Section. Finally, in *Concluding Remarks* section, we draw some concluding remarks.

# DEFINITION OF THE COPULA FUNCTION

An n-dimensional copula[2] is a multivariate distribution function (d.f.) C, with uniform distributed margins in [0, 1] (U(0, 1)) and the following properties:

1. C: $[0, 1]^n \rightarrow [0, 1]$;
2. C is grounded and n-increasing;
3. C has margins $C_i$ which satisfy $C_i(u) = C(1, \ldots, 1, u, 1, \ldots, 1) = u$ for all $u \in [0, 1]$.

It is obvious, from the above definition, that if $F_1, \ldots, F_n$ are univariate distribution functions, $C(F_1(x_1), \ldots, F_n(x_n))$ is a multivariate d.f. with margins $F_1, \ldots, F_n$, because $U_i = F_i(X_i)$, $i = 1, \ldots, n$, is a uniform random variable. Copula functions are a useful tool to construct and simulate multivariate distributions.

The following theorem is known as Sklar's Theorem [31, 32]. It is the most important theorem about copula functions because it is used in many practical applications.

**Theorem**[3]: *Let F be an n-dimensional d.f. with continuous margins $F_1, \ldots, F_n$. Then it has the following unique copula representation*:

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)). \quad (1)$$

From Sklar's Theorem we see that, for continuous multivariate distribution functions, the univariate margins and the multivariate dependence structure can be separated. The dependence structure can be represented by a proper copula function. Moreover, the following corollary is attained from **(1)**.

**Corollary**: *Let F be an n-dimensional d.f. with continous margins $F_1, \ldots, F_n$ and copula C (satisfying **(1)**). Then, for any $\mathbf{u} = (u_1, \ldots, u_n)$ in $[0, 1]^n$*:

$$C(u_1, \ldots, u_n) = F(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)), \quad (2)$$

*where $F_i^{-1}$ is the generalized inverse of $F_i$.*

A trivial example is the copula of independent random variables (the product copula). It takes the following form:

$$C_{ind}(u_1, \ldots, u_n) = u_1 \cdot \ldots \cdot u_n.$$

Another example is the Farlie–Gumbel–Morgenstern (FGM) copula, which in the bivariate case is defined by:

$$C(u_1, u_2) = u_1 u_2 [1 + \alpha(1 - u_1)(1 - u_2)], \quad -1 \leq \alpha \leq 1.$$

## Elliptical Copulas

The class of elliptical distributions provides useful examples of multivariate distributions because they share many of the tractable properties of the multivariate normal distribution. Furthermore, they allow modeling multivariate extreme events and forms of non-normal dependencies. Elliptical copulas are simply the copulas of elliptical distributions. Simulation from

---

[1]i.e.: the class of the elliptical copulas and the class of the Archimedean copulas.

[2]The original definition is given by Sklar (1959) [31–33].
[3]For the proof, see Sklar (1996) [32].

elliptical distributions is easy to perform. Therefore, as a consequence of Sklar's Theorem[4], the simulation of elliptical copulas is also easy.

## Normal Copula

The Gaussian (or normal) copula is the copula of the multivariate normal distribution. The random vector $\mathbf{X} = (X_1, \ldots, X_n)$ is multivariate normal iff:

(1) The univariate margins $F_1, \ldots, F_n$ are Gaussians;
(2) The dependence structure among the margins is described by a unique copula function C (the normal copula) such that[5]:

$$C_R^{Ga}(u_1, \ldots, u_n) = \Phi_R\left(\varphi^{-1}(u_1), \ldots, \varphi^{-1}(u_n)\right), \qquad (3)$$

where $\Phi_R$ is the standard multivariate normal d.f. with linear correlation matrix R and $\varphi^{-1}$ is the inverse of the standard univariate Gaussian d.f.

If $n = 2$, **Eq. 3** can be written as:

$$C_R^{Ga}(u, v) = \int_{-\infty}^{\varphi^{-1}(u)} \int_{-\infty}^{\varphi^{-1}(v)} \frac{1}{2\pi\left(1 - R_{12}^2\right)^{1/2}} \exp\left\{-\frac{s^2 - 2R_{12}st + t^2}{2\left(1 - R_{12}^2\right)}\right\} ds dt,$$

where $R_{12}$ is simply the linear correlation coefficient between the two random variables.

## t-Student Copula

The copula of the multivariate t-Student distribution is the t-Student copula. Let $\mathbf{X}$ be a vector with an n-variate t-Student distribution with $\nu$ degrees of freedom, mean vector $\boldsymbol{\mu}$ (for $\nu > 1$), and covariance matrix $\frac{\nu}{\nu-2}\Sigma$ (for $\nu > 2$)[6]. It can be represented in the following way:

$$\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + \frac{\sqrt{\nu}}{\sqrt{S}}\mathbf{Z}, \qquad (4)$$

where $\boldsymbol{\mu} \in R^n S \sim \chi_\nu^2$ and the random vector $\mathbf{Z} \sim N_n(\mathbf{0}, \Sigma)$ are independent.

The copula of vector $\mathbf{X}$ is the t-Student copula with $\upsilon$ degrees of freedom. It can be analytically represented in the following way:

$$C_{\nu,R}^t(\mathbf{u}) = t_{\nu,R}^n\left(t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_n)\right), \qquad (5)$$

where $R_{ij} = \Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}$ for $i, j \in \{1, \ldots, n\}$ and where $t_{\nu,R}^n$ denotes the multivariate d.f. of the random vector $\sqrt{\nu}\mathbf{Y}/\sqrt{S}$, where the random variable $S \sim \chi_\nu^2$ and the random vector $\mathbf{Y}$[7] are independent. $t_\nu$ denotes the margins[8] of $t_{\nu,R}^n$.

For $n = 2$, the t-Student copula has the following analytic form:

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

where $R_{12}$ is the linear correlation coefficient of the bivariate t-Student distribution with $\nu$ degrees of freedom, if $\nu > 2$.

# Archimedean Copulas

An Archimedean copula can be written in the following form:

$$C(u_1, \ldots, u_n) = \psi^{-1}\left[\psi(u_1) + \cdots + \psi(u_n)\right] \qquad (6)$$

for all $0 \leq u_1, \ldots, u_n \leq 1$ and where $\psi$ is a function often called the generator, satisfying:

i. $\psi(1) = 0$;
ii. for all $t \in (0, 1)$, $\psi'(t) < 0$, i.e. $\psi$ is decreasing;
iii. for all $t \in (0, 1)$, $\psi''(t) \geq 0$, i.e. $\psi$ is convex.

Examples of bivariate Archimedean copulas are the following:

Product Copula

$$\psi(t) = -\ln t; \quad C(u_1, u_2) = u_1 \cdot u_2.$$

Clayton Copula[9]

$$\psi(t) = t^{-\alpha} - 1, \quad \alpha > 0; \quad C(u_1, u_2) = \left(u_1^{-\alpha} + u_2^{-\alpha} - 1\right)^{-1/\alpha}.$$

Gumbel Copula[10]

$$\psi(t) = (-\ln t)^\alpha, \ \alpha \geq 1; \quad C(u_1, u_2)$$
$$= \exp\left\{-\left[(-\ln u_1)^\alpha + (-\ln u_2)^\alpha\right]^{1/\alpha}\right\}.$$

Frank copula[11]

$$\psi(t) = -\ln\frac{e^{-\alpha t} - 1}{e^{-\alpha} - 1}, \ \alpha \in \mathbf{R}; \quad C(u_1, u_2)$$
$$= -\frac{1}{\alpha}\ln\left(1 + \frac{(e^{-\alpha u_1} - 1)(e^{-\alpha u_2} - 1)}{e^{-\alpha} - 1}\right).$$

Extensions to the multivariate case are the following:

Cook–Johnson Copula[12]

---

[4]See **Eqs. 1 and 2**.

[5]As one can easily deduce from (**Eq. 2**).

[6]If $\nu \leq 2$, then the covariance matrix is not defined.

[7]$\mathbf{Y}$ has an n-dimensional normal distribution with mean vector $\mathbf{0}$ and covariance matrix R.

[8]All the margins are equally distributed.

[9]Clayton (1978) [33].
[10]Gumbel (1960) [34], Hougaard (1986) [35].
[11]Frank (1979) [36].
[12]It is a multivariate extension of the Clayton copula.

$$C(u_1, \ldots, u_n) = \left( \sum_{j=1}^{n} u_j^{-\alpha} - n + 1 \right)^{-1/\alpha.}$$

Gumbel–Hougaard Copula

$$C(u_1, \ldots, u_n) =$$

$$\exp\left\{ -\left[ (-\ln u_1)^\alpha + (-\ln u_2)^\alpha + \cdots + (-\ln u_n)^\alpha \right]^{\frac{1}{\alpha}} \right\}$$

Frank Copula

$$C(u_1, \ldots, u_n) = -\frac{1}{\alpha} \ln\left( 1 + \frac{(e^{-\alpha u_1} - 1) \cdot (e^{-\alpha u_2} - 1) \cdot \cdots \cdot (e^{-\alpha u_n} - 1)}{(e^{-\alpha} - 1)^{n-1}} \right).$$

# PARAMETER ESTIMATION OF A GIVEN COPULA

## The Maximum Likelihood (ML) Method

Let f be the density of the joint distribution F:

$$f(x_1, \ldots, x_n) = c(F_1(x_1), \ldots, F_n(x_n)) \prod_{i=1}^{n} f_i(x_i)$$

where $f_i$ is the univariate density of the marginal distribution $F_i$ and c is the density of the copula given by the following expression:

$$c(u_1, \ldots, u_n) = \frac{\partial C(u_1, \ldots, u_n)}{\partial u_1 \ldots \partial u_n}.$$

We suppose a set of T empirical data of n financial asset log-returns, $\chi = \{(x_1^t, \ldots, x_n^t)\}_{t=1}^{T}$. Let $\vartheta = (\vartheta_1, \ldots, \vartheta_n, \alpha)$ be the parameter vector to estimate, where $\vartheta_i$, $i = 1, \ldots, n$ is the vector of parameters of the marginal distribution $F_i$ and $\alpha$ is the vector of the copula parameters. The log-likelihood function is the following:

$$l(\vartheta) = \sum_{t=1}^{T} \ln c(F_1(x_1^t; \vartheta_1), \ldots, F_n(x_n^t; \vartheta_n); \alpha) + \sum_{t=1}^{T} \sum_{i=1}^{n} \ln f_i(x_i^t; \vartheta_i).$$
$$(7)$$

The ML estimator $\widehat{\vartheta}$ of the parameter vector $\vartheta$ is the one that maximizes **Eq. 7**, i.e.:

$$\widehat{\vartheta} = \mathrm{argmax}\, l(\vartheta).$$

## The Method of Inference Functions for Margins (IFM)

According to the IFM method[13], the parameters of the marginal distributions are estimated separately from the parameters of the copula. In other words, the estimation process is divided into the following two steps:

i. Estimating the parameters $\vartheta_i$, $i = 1, \ldots, n$ of the marginal distributions $F_i$ using the ML method:

$$\widehat{\vartheta}_i = \mathrm{argmax}\, l^i(\vartheta_i) = \mathrm{argmax} \sum_{t=1}^{T} \ln f_i(x_i^t; \vartheta_i)$$

where $l^i$ is the log-likelihood function of the marginal distribution $F_i$;

(ii) Estimating the copula parameters $\alpha$, given the estimations performed in step (i):

$$\widehat{\alpha} = \mathrm{argmax}\, l^c(\alpha) = \mathrm{argmax} \sum_{t=1}^{T} \ln c\left( F_1(x_1^t; \widehat{\vartheta}_1), \ldots, F_n(x_n^t; \widehat{\vartheta}_n); \alpha \right)$$

where $l^c$ is the log-likelihood function of the copula.

## The Canonical Maximum Likelihood (CML) Method

The CML method differs from the IFL method because no assumptions are made about the parametric form of the marginal distributions. The estimation process is performed in two steps:

i. Transforming the dataset $(x_1^t, \ldots, x_n^t)$, $t = 1, \ldots, T$, into uniform variates $(\widehat{u}_1^t, \ldots, \widehat{u}_n^t)$, using the empirical distributions[14];

ii. Estimating the copula parameters as follows:

$$\widehat{\alpha} = \mathrm{argmax} \sum_{t=1}^{T} \ln c(\widehat{u}_1^t, \ldots, \widehat{u}_n^t; \alpha).$$

For example, we can estimate the parameter **R** of the Gaussian copula (**Eq. 3**) with the CML or the IFM method in the following way[15]:

$$\widehat{R}_{IFM/CML} = \frac{1}{T} \sum_{t=1}^{T} \varsigma_t^{\mathrm{T}} \varsigma_t$$

where $\varsigma_t = (\Phi^{-1}(u_1^t), \ldots, \Phi^{-1}(u_n^t))$. In this notation $u_i^t = \widehat{u}_i^t$ when we are using the CML method and $u_i^t = F_i(x_i^t; \widehat{\vartheta}_i)$ when we are using the IFM method, $i = 1, \ldots, n$.

The following recursive procedure[16] is used to estimate the parameter **R** of the $t_v$-Student copula (**Eq. 5**):

i. Let $\widehat{R}_1$ be the IFM/CML estimator of the **R** parameter for the Gaussian copula;

ii. $\widehat{R}_{m+1} = \frac{1}{T} \left( \frac{v+n}{v} \right) \sum_{t=1}^{T} \frac{\varsigma_t^{\mathrm{T}} \varsigma_t}{1 + \frac{1}{v} \varsigma_t \widehat{R}_m^{-1} \varsigma_t^{\mathrm{T}}}, \quad m = 1, 2, \ldots,$

where $\varsigma_t = (t_v^{-1}(u_1^t), \ldots, t_v^{-1}(u_n^t))$;

---

[13]Joe and Xu (1996) [13].

[14]In other words, the variates $(\widehat{u}_1^t, \ldots, \widehat{u}_n^t)$ are generated from the empirical copula.
[15]See [7].
[16]See [4].

iii. Step (ii) is repeated until $\widehat{R}_{\mathbf{m}+1} = \widehat{R}_{\mathbf{m}}$. So, the IFM/CML estimator of the parameter $\mathbf{R}$ for the $t_\nu$-Student copula is $\widehat{R}_{\mathbf{IFM/CML}} = \widehat{R}_\infty$.

Mashal and Zeevi [37] suggest using the following algorithm to estimate the parameters $\nu$ and $\mathbf{R}$ of the $t_\nu$-Student copula:

i. Transforming the dataset $(x_1^t, \ldots, x_n^t)$, t = 1, …, T, into uniform variates $(\widehat{u}_1^t, \ldots, \widehat{u}_n^t)$, using the empirical marginal distributions.
ii. Estimate $\widehat{R}$ using the Kendall's $\tau$ non parametric estimator: $\widehat{R}_{ij} = \sin(\frac{\pi}{2}\widehat{\tau}_{ij})$ , i,j = 1,…,n.
iii. Perform a numerical search for $\widehat{\nu}$, i.e.,

$$\widehat{\nu} = \underset{\nu \in (2,\infty]}{\mathrm{argmax}} \left[ \sum_{t=1}^{T} \log\left( c(u_1^t, \ldots, u_n^t; \nu, \widehat{R}) \right) \right], \text{ where } c(u_1, \ldots, u_n;$$

$$\nu, \mathbf{R}) = \frac{\Gamma\left((\nu+n)/2\left[\Gamma(\nu/2)\right]^{n-1}\left(1 + \mathbf{y}'\mathbf{R}^{-1}\mathbf{y}\right)^{-(\nu+n)/2}\right.}{|\mathbf{R}|^{1/2}[\Gamma((\nu+1)/2)]^n \prod_{i=1}^{n} (1+y_i^2/\nu)^{-(\nu+1)/2}} \text{ and}$$

$$\mathbf{y} = (y_1, \ldots, y_n) = (t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_n)).$$

## Parameter Estimation and Dependence Measures

This method works only with one-parameter bivariate copulas. The main dependence measures[17] can be written as a function of the copula [15]. In some cases analytical solutions are available and the copula parameter can simply be written as a function of the dependence measure. Otherwise, a numerical procedure is necessary.

For instance, for the Gaussian copula we obtain:

$$R_{12} = 2\sin\left(\frac{\pi}{6}\rho_S\right) \text{ and } R_{12} = \sin\left(\frac{\pi}{2}\tau\right)$$

For the Clayton copula:

$$\alpha = \frac{2\tau}{1-\tau}.$$

For the Gumbel copula:

$$\alpha = (1-\tau)^{-1}.$$

For the Morgenstern copula:

$$\alpha = 3\rho_S \text{ and } \alpha = \frac{9}{2}\tau.$$

## Non-parametric Estimation

So far, the parameters of a given type of copula are been estimated. Now the empirical copula (or the Deheuvels copula [38]) is constructed from the sample data. This is any copulas of the empirical multivariate distribution.

Let $\{x_1^{(t)}, \ldots, x_n^{(t)}\}$ be the order statistics and $\{r_1^t, \ldots, r_n^t\}$ be the rank statistics, t = 1,…, T of the dataset. We have: $x_i^{(r_i^t)} = x_i^t$, i = 1,…, n.

---

[17]i.e. the rank correlation coefficients: the Spearman's rho, $\rho_S$, and the Kendall's tau, $\tau$.

Any Function

$$\widehat{C}\left(\frac{t_1}{T}, \ldots, \frac{t_n}{T}\right) = \frac{1}{T} \sum_{t=1}^{T} \prod_{i=1}^{n} 1_{\left[r_i^t \le t_i\right]} \quad (8)$$

defined on the lattice $\ell = \left\{ \left(\frac{t_1}{T}, \ldots, \frac{t_n}{T}\right) : 1 \le i \le n; \ t_i = 0, \ldots, T \right\}$ is an empirical copula.

The empirical copula density [15] has the following expression:

$$\widehat{c}\left(\frac{t_1}{T}, \ldots, \frac{t_n}{T}\right) = \sum_{i_1=1}^{2} \cdots \sum_{i_n=1}^{2} (-1)^{i_1+\cdots+i_n} \widehat{C}\left(\frac{t_1 - i_1 + 1}{T}, \ldots, \frac{t_n - i_n + 1}{T}\right).$$

# SELECTING THE RIGHT COPULA

In *Parameter Estimation of a Given Copula* Section, some methods to calibrate the parameters of a given analytical representation of copula function are illustrated. Now the issue is selecting the type of copula function which fits better the empirical data.

## Selecting an Archimedean Copula

The method described in this section (see [13]) can select the Archimedean copula which fits better real data. An Archimedean copula has the analytical representation given by **Eq. 6**. So, to select the copula, it is sufficient to identify the generator, $\psi$.

In the bivariate case ($n$ = 2), Genest and Rivest [13] defined a univariate function, K, which is related to the generator of the Archimedean copula through the following expression:

$$K_\psi(z) = z - \frac{\psi(z)}{\psi'(z)}. \quad (9)$$

A non parametric estimation of (**9**) is the following:

$$\widehat{K}(z) = \frac{1}{T} \sum_{t=1}^{T} 1_{[\vartheta_t \le z]} \quad (10)$$

where $\vartheta_i = \frac{1}{T-1} \sum_{t=1}^{T} 1_{[x_1^t < x_1^i, \ x_2^t < x_2^i]}$, i = 1,…,T.

We choose a parametric representation for the generator, $\psi$. Then, the parameter, $\alpha$ of the selected Archimedean copula is estimated using, for instance, the following estimation of the Kendall's $\tau$:

$$\tau = \binom{T}{2}^{-1} \sum_{i<j} sign\left[\left(x_1^i - x_1^j\right) \cdot \left(x_2^i - x_2^j\right)\right].$$

The parameter $\alpha$ may also be estimated using the IFM or the CML method. Using $\alpha$, a parametric estimation of (**9**) is easily obtained.

All the steps described above are repeated for different choices of the generator $\psi$. To select the Archimedean copula which fits better the dataset, Frees and Valdez [10] propose to use a QQ-plot between **Eqs. 9** and **10**.

The optimal copula may also be selected by minimizing the distance based on the L$^2$ norm between **Eqs. 9** and **10**:[18]

$$d_2\left(\widehat{K}, K\right) = \int_0^1 \left[K(z) - \widehat{K}(z)\right]^2 dz.$$

The method described in this section may also be used to graphically estimate the parameter $\alpha$ of a given Archimedean copula.

## Selecting the Right Copula Using the Empirical Copula

Let $\{C_k\}_{1 \le k \le K}$ be the set of the available copulas. We choose the copula $C_k$ which minimize the following distance, based on the discrete L$^n$ norm, between the same $C_k$ and the empirical copula as defined in **(8)**:

$$\overline{d}_n\left(\widehat{C}, C_k\right) = \left(\sum_{t_1=1}^{T} \cdots \sum_{t_n=1}^{T} \left[\widehat{C}\left(\frac{t_1}{T}, \ldots, \frac{t_n}{T}\right) - C_k\left(\frac{t_1}{T}, \ldots, \frac{t_n}{T}\right)\right]^2\right)^{1/2.} \quad (11)$$

The distance (**Eq. 11**) may also be used to estimate the vector of parameters $\vartheta \in \Theta$ of a given copula $C(\mathbf{u}; \vartheta)$ in the following way:

$$\widehat{\vartheta} = \arg\min_{\vartheta \in \Theta} \left(\sum_{\mathbf{u} \in \ell} \left[\widehat{C}(\mathbf{u}) - C(\mathbf{u}; \vartheta)\right]^2\right)^{1/2.}$$

## SIMULATION ALGORITHMS

In this section, we show a collection of algorithms to simulate random variates $(u_1, \ldots, u_n)$ from certain types of copula C. For the definition of the copula, these random variates $u_i$ are a determination of correlated uniform(0,1) distributed random variables. So, to simulate random variates $(x_1, \ldots, x_n)$ from a multivariate distribution F with given margins $F_i$, $i = 1, \ldots, n$, and copula C, we have to invert each $u_i$ using the marginal distributions: $x_i = F_i^{-1}(u_i)$, $i = 1, \ldots, n$.

## Simulation From the Gaussian Copula

To generate random variates from the Gaussian copula (**Eq. 3**), we can use the following procedure. If the matrix R is positive definite, then there is some $n \times n$ matrix A such as $R = AA^T$. It is also assumed that the random variables $Z_1, \ldots, Z_n$ are independent standard normal. Then, the random vector $\boldsymbol{\mu} + A\mathbf{Z}$ (where $\mathbf{Z} = (Z_1, \ldots, Z_n)^T$ and the vector $\boldsymbol{\mu} \in R^n$) is multivariate normal distributed with mean vector $\boldsymbol{\mu}$ and covariance matrix R.

The matrix A can be easily determined with the Cholesky decomposition of R. This decomposition is the unique lower-triangular matrix L such as $LL^T = R$. Hence, one can generate random variates from the n-dimensional Gaussian copula running the following algorithm:

Find the Cholesky decomposition A of the matrix R;
Simulate n independent standard normal random variates $\mathbf{z} = (z_1, \ldots, z_n)^T$;
Set $\mathbf{x} = A\mathbf{z}$;
Determine the components $u_i = \varphi(x_i)$, $i = 1, \ldots, n$;
The vector $(u_1, \ldots, u_n)^T$ is a random variate from the n-dimensional Gaussian copula, $C_R^{Ga}$.

## Simulation From the t$_\nu$-Student Copula

To simulate random variates from the t-Student copula (**5**), $C_{\nu,R}^t$, we can use the following algorithm, which is based on **Eq. 4**:

Find the Cholesky decomposition, A, of R;
Simulate n independent random variates $\mathbf{z} = (z_1, \ldots, z_n)^T$ from the standard normal distribution;
Simulate a random variate, s, from $\chi_\nu^2$ distribution, independent of $\mathbf{z}$;
Determine the vector $\mathbf{y} = A\mathbf{z}$;
Set $\mathbf{x} = \frac{\sqrt{\nu}}{\sqrt{s}}\mathbf{y}$;
Determine the components $u_i = t_\nu(x_i)$, $i = 1, \ldots, n$;
The resultant vector is: $(u_1, \ldots, u_n)^T \sim C_{\nu,R}^t$.

## Simulation From the Cook–Johnson Copula

This algorithm is a particular case of the one suggested by [39] for the generation of multivariate outcomes from a compound copula. To generate random variates from the Cook–Johnson copula with a parameter $\alpha$, we have to perform the steps below:

Generate n independent random variates, $y_1, \ldots, y_n$ from the exponential distribution[19] with parameter $\lambda = 1$;
Generate a random variate, z, from a $\text{Gamma}(1/\alpha, 1)$ distribution independent of $y_1, \ldots, y_n$;
Set $u_j = (1 + y_j/z)^{-1/\alpha}$, $j = 1, \ldots, n$;
The vector $\mathbf{u} = (u_1, \ldots, u_n)$ is generated from the Cook–Johnson copula.

The Cook–Johnson copula reproduces a positive dependence structure. A negative dependence structure may be obtained for some of the variables by setting $u_i^* = 1 - u_i$.

## Simulation From the Morgenstern Copula

The following algorithm [40] generates bivariate random variates from the Farlie–Gumbel–Morgenstern copula:

Generate independent uniform(0,1) random variates $v_1$ and $v_2$;
Set $u_1 = v_1$;
Calculate $A = \alpha(2u_1 - 1) - 1$ and $B = [1 - \alpha(2u_1 - 1)]^2 + 4\alpha v_2 (2u_1 - 1)$;
Set $u_2 = 2v_2/(\sqrt{B} - A)$;

---

[18]See [7].

[19]The exponential distribution has the following form: $F(x; \lambda) = 1 - e^{-\lambda x}$, $x > 0$.

The vector $(u_1, u_2)$ is generated from the Farlie–Gumbel–Morgenstern copula.

## A General Algorithm to Simulate a Copula

This method is based on the conditional distributions of a random vector $\mathbf{U} = (U_1, \ldots, U_n)$. In the bivariate case, we have:

$$\Pr\{U_2 \leq u_2 \backslash U_1 = u_1\} = C_{2\backslash 1}(u_1, u_2)$$

where $C_{2\backslash 1}(u_1, u_2) = \lim\limits_{\Delta u \to 0^+} \dfrac{c(u_1 + \Delta u, u_2) - C(u_1, u_2)}{\Delta u} = \dfrac{\partial C(u_1, u_2)}{\partial u_1}$.

The Algorithm[20] is the Following:

Generate two independent uniform(0,1) random variates $v_1$ and $v_2$;
Set $u_1 = v_1$;
Let $C(u_2; u_1) = C_{2\backslash 1}(u_1, u_2)$. Set $u_2 = C^{-1}(v_2; u_1)$;
The vector $(u_1, u_2)$ is generated from the copula C.

For instance, for the bivariate Frank copula, we have:

$$C_{2\backslash 1}(u_1, u_2) = \frac{(e^{-\alpha u_2} - 1)e^{-\alpha u_1}}{(e^{-\alpha} - 1) + (e^{-\alpha u_1} - 1)(e^{-\alpha u_2} - 1)}$$

and

$$C^{-1}(u; u_1) = \{u_2 : C_{2\backslash 1}(u_1, u_2) = u\}$$
$$= -\frac{1}{\alpha} \ln\left(1 + \frac{u(e^{-\alpha} - 1)}{u + (1 - u)e^{-\alpha u_1}}\right).$$

The above algorithm may be generalized to the multivariate case:

Generate n independent uniform(0,1) random variates, $(v_1, \ldots, v_n)$;
Set $u_1 = v_1$;
let $C(u_m; u_1, \ldots, u_{m-1}) = C_{m \backslash 1, \ldots, m-1}(u_1, \ldots, u_m)$, $m = 2, \ldots, n$, where

$$C_{m \backslash 1, \ldots, m-1}(u_1, \ldots, u_m) = \Pr\{U_m \leq u_m \backslash$$

$$(U_1, \ldots, U_{m-1}) = (u_1, \ldots, u_{m-1})\}$$
$$= \frac{\partial^{m-1}_{(u_1, \ldots, u_{m-1})} C(u_1, \ldots, u_m, 1, \ldots, 1)}{\partial^{m-1}_{(u_1, \ldots, u_{m-1})} C(u_1, \ldots, u_{m-1}, 1, \ldots, 1)} \quad (12)$$

Set $u_m = C^{-1}(v_m; u_1, \ldots, u_{m-1})$, $m = 2, \ldots, n$;
The vector $(u_1, \ldots, u_n)$ is generated from the copula C.

This algorithm is computationally intensive for high values of n. It is a difficult issue to compute the conditional distribution (12).

## Simulation From the Empirical Copula

The below algorithm permits to generate a vector of random variates from the empirical copula (**Eq. 8**):

[20]Introduced by [12].

Randomly draw a complete observation vector $(x_1^t, \ldots, x_n^t)$ from the historical dataset $\chi$;
Using the empirical distribution functions, $\widehat{F}_i$, to transform each component of the observation vector to a set of uniform variates: $u_i = \widehat{F}_i(x_i^t)$, $i = 1, \ldots, n$;
$(u_1, \ldots, u_n)$ is a vector of non-independent uniforms(0, 1) that are dependent through the empirical copula.

## APPLICATION TO TWO ITALIAN EQUITIES

In this section, we apply the methods of copula function calibration and simulation described before. We use a dataset of 1,012 daily observations of the log-returns of two Italian equities: TIM and Olivetti. In **Table 1**, the principal statistics regarding the two Italian equities are reported. In **Figure 1**, we plot the empirical standardized log-returns of TIM against the standardized log-returns of Olivetti.

We have estimated, with the CML method, the parameters of different types of the bivariate copula, using the dataset of 1,012 historical daily log-returns observations. In this way, we do not consider any particular analytical form for the marginal distributions, and only the copula effects are taken into account.

Therefore, we have selected the copula which better approximates the empirical copula using the $L^2$ norm (**Eq. 11**). The results are shown in **Table 2**.

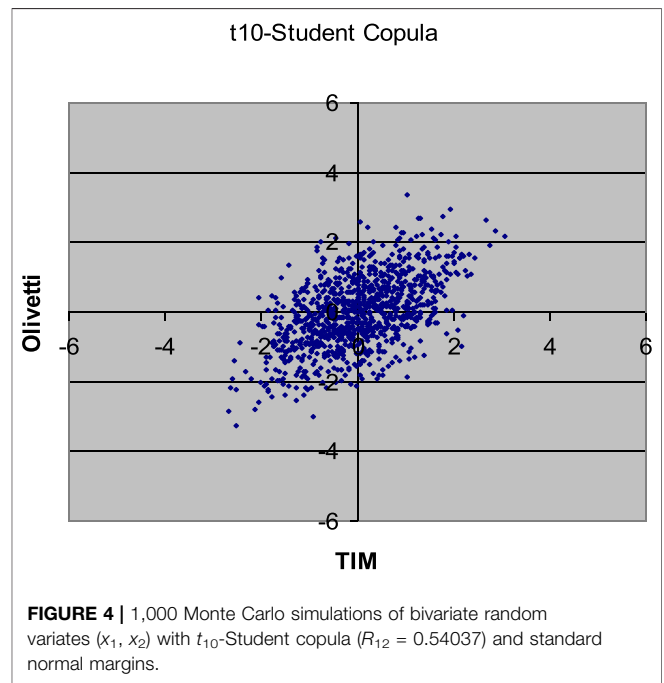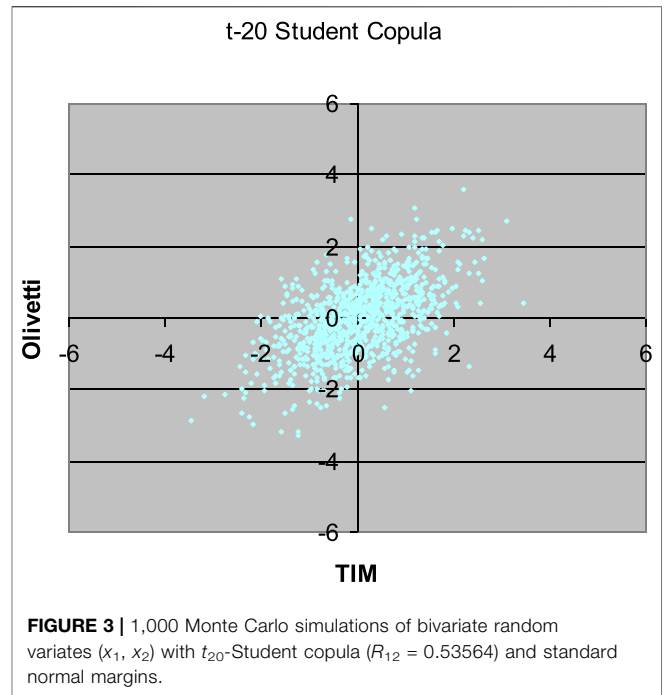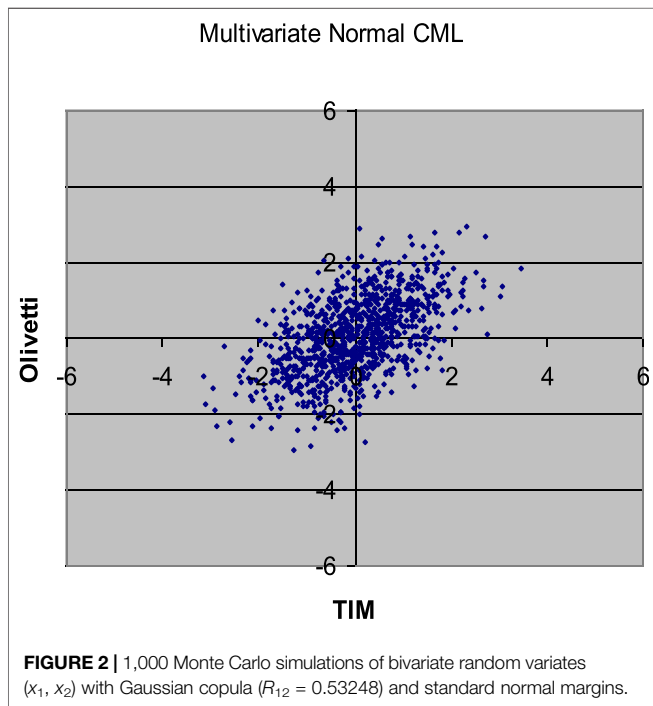**TABLE 1 |** Main statistics of the empirical distribution of the log-returns of TIM and Olivetti.

|  | Mean | Standard deviation |
|---|---|---|
| TIM | 0.000269 | 0.025799233 |
| Olivetti | 0.000919 | 0.031200767 |

| Linear correlation | Spearman's rho | Kendall's tau |
|---|---|---|
| 0.522391 | 0.517832 | 0.359868 |



**FIGURE 1 |** Plot of the empirical standardized log-returns TIM/Olivetti.
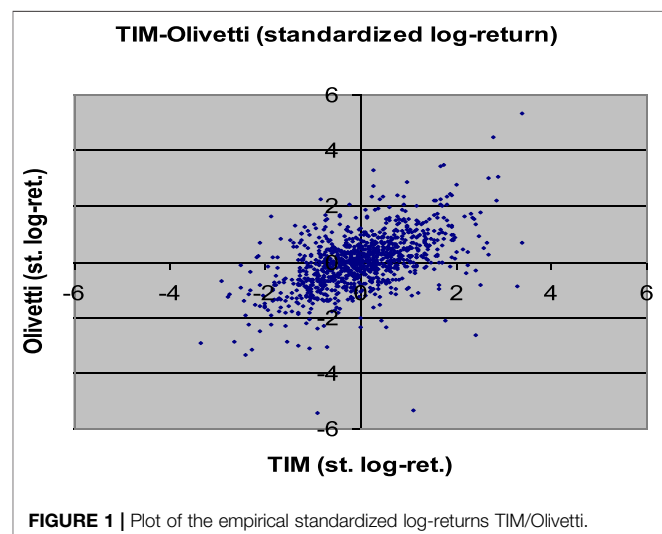
**TABLE 2 |** CML estimation of the parameters ($\alpha$ or $R_{12}$) and calculation of the $L^2$ norm for different copula types.

| Copula | Parameter estimation | $\overline{d}_2(\hat{C}, C)/1012$ |
|---|---|---|
| Gaussian | 0.53248 | 0.00451 |
| $t_5$-student | 0.53953 | 0.00460 |
| $t_{10}$-student | 0.54037 | 0.00432 |
| $t_{20}$-student | 0.53564 | 0.00446 |
| FGM | 1.55349 | 0.00595 |
| Gumbel | 1.56218 | 0.00839 |
| Frank | 3.82211 | 0.00507 |
| Clayton | 1.12436 | 0.01583 |



**FIGURE 3 |** 1,000 Monte Carlo simulations of bivariate random variates ($x_1$, $x_2$) with $t_{20}$-Student copula ($R_{12}$ = 0.53564) and standard normal margins.



**FIGURE 2 |** 1,000 Monte Carlo simulations of bivariate random variates ($x_1$, $x_2$) with Gaussian copula ($R_{12}$ = 0.53248) and standard normal margins.



**FIGURE 4 |** 1,000 Monte Carlo simulations of bivariate random variates ($x_1$, $x_2$) with $t_{10}$-Student copula ($R_{12}$ = 0.54037) and standard normal margins.

Observing the results in **Table 2**, the $t_{10}$-Student copula seems to be the one that better approximates the empirical copula of the dataset. However, the difference between the $t_{10}$-Student copula and the Normal copula is very low. So the Gaussian copula could be appropriate. We remember that the use of the Gaussian copula permits us to construct algorithms to simulate scenarios from a multivariate distribution with different margins. The commonly used multivariate Normal is only a particular case where all the margins are Gaussians too.

The simulation algorithms shown in *Selecting the Right Copula* Section are applied for simulating 1,000 scenarios for the standardized log-returns of the equities TIM and Olivetti, using the parameter estimations in **Table 2**. In all the cases, we have used standardized Gaussian margins, because our aim is only to compare the different copulas. In **Figures 2–8** we have plotted the results.

Comparing the plots in the above Figures with the empirical distribution obtained from the historical data and represented in **Figure 1**, we can see some differences. These deviations may be caused by an inadequate choice of the margins[21]. Our aim was only to compare different types of copulas without assumptions about the analytical form of the marginal distributions.
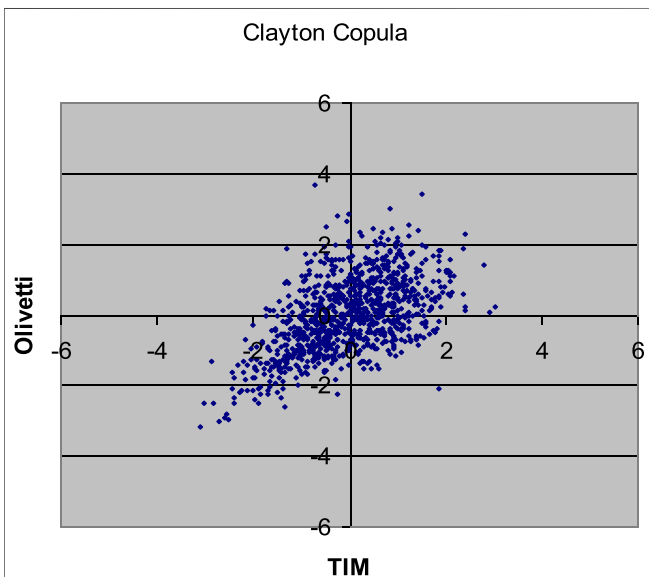
---

[21]Standardized Gaussians.
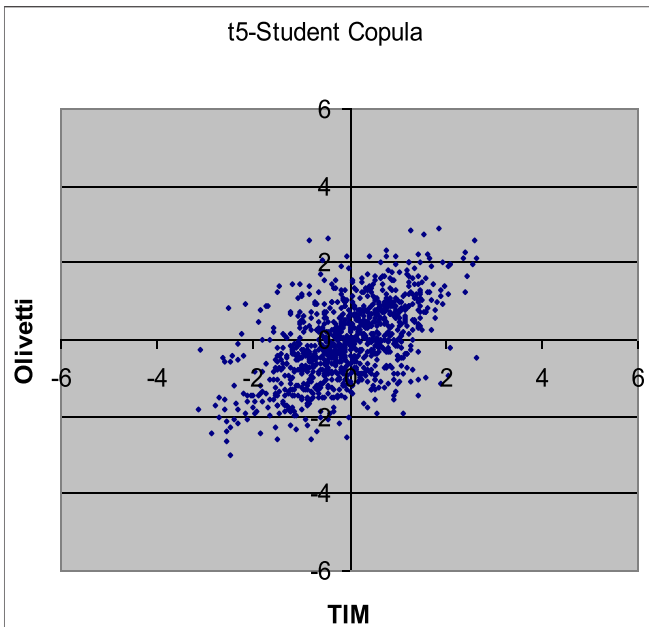
## t5-Student Copula



**FIGURE 5 |** 1,000 Monte Carlo simulations of bivariate random variates $(x_1, x_2)$ with $t_5$-Student copula ($R_{12} = 0.53953$) and standard normal margins.

## FGM Copula
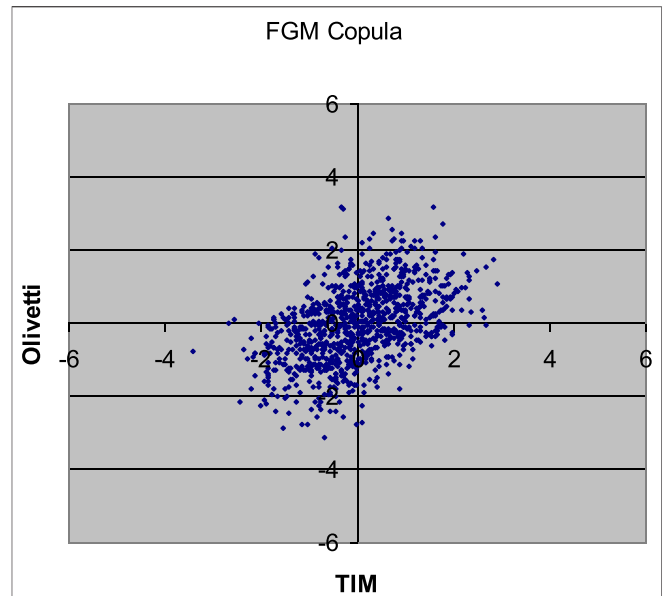


**FIGURE 7 |** 1,000 Monte Carlo simulations of bivariate random variates $(x_1, x_2)$ with Farlie-Gumbel-Morgenstern copula ($\alpha = 1.55349$) and standard normal margins.

## Clayton Copula



**FIGURE 6 |** 1,000 Monte Carlo simulations of bivariate random variates $(x_1, x_2)$ with Clayton copula ($\alpha = 1.12436$) and standard normal margins.
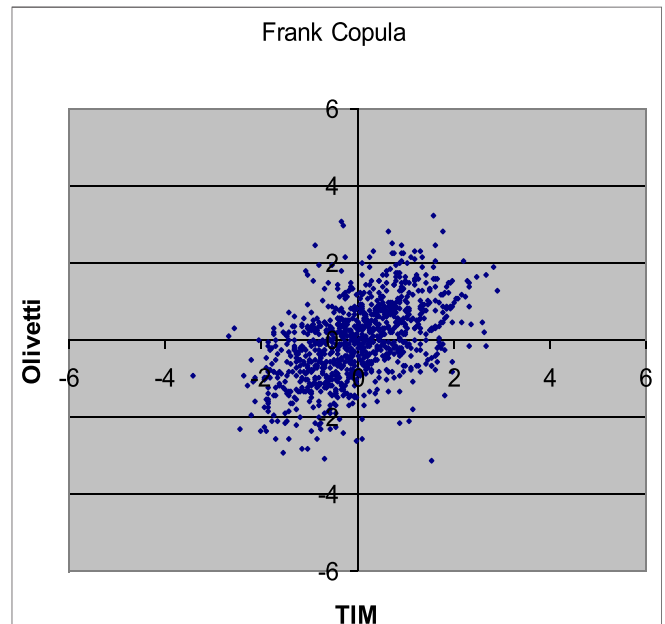
## Frank Copula



**FIGURE 8 |** 1,000 Monte Carlo simulations of bivariate random variates $(x_1, x_2)$ with Frank copula ($\alpha = 3.82211$) and standard normal margins.

## CONCLUDING REMARKS

This paper is a brief review of the existing methodologies for calibrating, choosing, and simulating different types of copula functions. The described methods are applied to a historical log-returns dataset of two Italian equities. We have seen how copula functions are a useful tool for implementing efficient simulation algorithms. Practical algorithms for generating

Monte Carlo scenarios from a multivariate distribution with a fixed copula and different margins are easily implemented to simulate financial asset returns. The traditional models use the multinormal distribution[22] to simulate asset log-returns. We can

---

[22]i.e. Gaussian copula and margins.

choose different marginal distributions for building more efficient algorithms also using a normal copula. The choice of the margins seems to have a more significant impact than the choice of the type of the copula on the results of the simulation. In this paper, only the copula effects are taken into account.

As clearly underlined in the paper, the application to a return time series of two Italian equities has only a demonstrative scope. The copula functions may be implemented to a portfolio of n financial assets traded on different stock markets such as the American, European and Asian ones.

In the field of portfolio risk management, the copula functions are precious tools for developing advanced financial risk measurement models. Traditionally, these quantitative instruments have been utilized for measuring more adequately the market, credit, and operational risks of financial institutions. Successively, copula functions have been also implemented in the field of the integrated measurement of the different financial risks by modeling the dependence structure among the market, credit,

and operational losses. Currently, the copula functions are used for estimating the marginal contribution of each financial institution to the systemic risk, that is the instability of the global financial system. The implications in terms of macro-prudential policy and supervisory choices on financial institutions are evident.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

## AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

## REFERENCES

1. Li DX. On default correlation: a Copula function approach. *J Fixed Income* (2000). 9:43–54. doi:10.3905/jfi.2000.319253
2. Embrechts P, McNeil AJ, Straumann D. Correlation and dependence in risk management: properties and pitfalls In MAH Dempster, editor. *Value at risk and beyond*. Cambridge, UK: Cambridge University Press (2002). p. 176–223.
3. Alexander C. *Market models: a guide to financial data analysis*. London: Wiley (2001).
4. Bouyé E, Durrleman V, Nikeghbali A, Riboulet G, Roncalli T. Copulas for finance—a reading guide and some applications, Groupe de Recherche Opérationnelle, Crédit Lyonnais, Working Paper (2001).
5. Bouyé E, Durrleman V, Nikeghbali A, Riboulet G, Roncalli T. Copulas: an open field for risk management. Groupe de Recherche Opérationnelle, Crédit Lyonnais, Working Paper (2000).
6. Cook RD, Johnson ME. A family of distributions for modeling non-elliptically symmetric multivariate data, *J R Stat Soc B* (1981). 43:210–8. doi:10.1111/j.2517-6161.1981.tb01173.x
7. Durrleman V, Nikeghbali A, Roncalli T. Which copula is the right one?, Groupe de Recherche Opérationnelle, Crédit Lyonnais, Working Paper (2000).
8. Embrechts P, Hoeing A, Juri A. *Using copulae to bound the Value-at-Risk for functions of dependent risks, Finance and Stochastics*, New York: Springer (2003).
9. Embrechts P, Lindskog F, McNeil AJ. Modelling dependence with copulas and applications to risk management. In: "*Handbook of heavy-tailed distributions in finance, Chapter 8*", S Rachev, editor. Amsterdam: Elsevier (2003). p. 329–84.
10. Frees EW, Valdez EA. Understanding relationships using copulas, *North Am Actuarial J.* (1998). 2:1–25. doi:10.1080/10920277.1998.10595667
11. Frey R, McNeil AJ (2001). Modelling dependent default, ETH Zurich preprint.
12. Genest C, MacKay J. The joy of copulas: bivariate distributions with uniform marginals. *Am Stat* (1986). 40:280–3. doi:10.2307/2684602
13. Genest C, Rivest L-P. Statistical inference procedures for bivariate Archimedean copulas, *J Am Stat Assoc* (1993). 88:1034–43. doi:10.1080/01621459.1993.10476372
14. Joe H, Xu JJ. The estimation method of inference functions for margins for multivariate models. Department of statistics, University of British Columbia, Technical Report N (1996).
15. Nelsen R. *An introduction to Copulas*. New York: Springer (1998).
16. Oakes D. (1982). A model for association in bivariate survival data. *J R Stat Soc B* 44:414–22. doi:10.1111/j.2517-6161.1982.tb01222.x
17. Roncalli T. Gestion des Risques Multiples, Cours ENSAI de 3ème année, Groupe de Recherche Opérationnelle, Crédit Lyonnaise (2002).
18. Wang SS. Aggregation of correlated risk portfolios: models & algorithms, CAS Committee on Theory of Risk, Working Paper (1999).

19. Markowitz H. Portfolio selection. *J Finance* (1952). 7:77–91. doi:10.1111/j.1540-6261.1952.tb01525.x
20. Di Clemente A, Romano C. Measuring and optimizing portfolio credit risk: a copula-based approach. *Econ Notes by Banca Monte dei Paschi di Siena SpA* (2004). 33(no. 3):325–57. doi:10.1111/j.0391-5026.2004.00135.x
21. Breymann W, Dias A, Embrechts P. Dependence structures for multivariate high-frequency data in finance. *Q. Finance* (2003). 3(2):1–14. doi:10.1080/713666155
22. Daul S, De Giorgi E, Lindskog F, McNeil AJ. The grouped t-copula with an application to credit risk. *Risk* (2003). 16:73–6.
23. Cherubini U, Luciano E, Vecchiato W. *Copula methods in finance*. New York: John Wiley (2004).
24. Demarta S, McNeil AJ. *The t Copula and related Copulas*, working paper. Zurich: ETH Zentrum (2004).
25. Bouyé E, Salmon M. Dynamic copula quantile regressions and tail area dynamic dependence in Forex markets. *Eur J Finance* (2009). 15:721–50. doi:10.1080/13518470902853491
26. Hakwa B, Jager-Ambrozewicz M, Rudiger B. Analysing systemic risk contribution using a closed Formula for conditional valueat-risk through copula. *Commun Stoch Anal* (2015). 9(1):131–58. doi:10.31390/cosa.9.1.08
27. Karimalis EN, Nomikos NK. Measuring systemic risk in the European banking sector: a Copula CoVaR approach. *Eur J Finance* (2018). 24(11):944–75. doi:10.1080/1351847x.2017.1366350
28. Di Clemente A. Estimating the marginal contribution to systemic risk by a CoVaR-model based on copula functions and ExtremeValue Theory. *Finance Monet Econ* (2018). 47(1):69–112. doi:10.1111/ecno.12095
29. Gauthier C, Lehar A, Souissi M. Macroprudential capital requirements and systemic risk. *J Financial Intermediation* (2012). 21(4):594–618. doi:10.1016/j.jfi.2012.01.005
30. Di Clemente A. Comparing different systemic risk measures for European banking system. *Int Bus Res* (2019). 12:35–53.
31. Sklar A. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris* (1959). 8:229–31.
32. Sklar A. (1996). : *Random variables, distribution functions, and copulas – a personal look backward and forward*, in *Distributions with fixed marginals and related topics*, ed. By L Rüschendorff, B Schweizer, M Taylor. Hayward, CA: Institute of Mathematical Statistics, pp. 1–14.
33. Clayton DG. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* (1978). 65:141–51. doi:10.1093/biomet/65.1.141
34. Gumbel EJ. Bivariate exponential distributions. *J Am Stat Assoc* (1960). 55:698–707. doi:10.1080/01621459.1960.10483368
35. Hougaard P. A class of multivariate Failure time distributions. *Biometrika* (1986). 73:671–8. doi:10.2307/2336531

36. Frank MJ. On the simultaneous associativity of F(x, and x+y-F(x,y). *Aequationes Mathematicae* (1979). 19:194–226. doi:10.1007/bf02189866

37. Mashal R, Zeevi A. "*Beyond Correlation: extreme Co-movements between financial assets*". Working Paper: Columbia University (2002).

38. Deheuvels P. 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. N.* (1979). 65:274–92. doi:10.3406/barb.1979.58521

39. Marshall AW, Olkin I. Families of multivariate distributions. *J Am Stat Assoc* (1988). 84:487–93.

40. Johnson ME. *Multivariate statistical simulation*. New York: Wiley (1987).

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.