Research Article
Multivariate Option Pricing with Pair-Copulas

Anna Barban[1] and Luca Di Persio[2]

[1] Group Strategic Planning and Control at Generali, Head Office, Via N. Machiavelli 4, 34132 Trieste, Italy
[2] Department of Computer Science, University of Verona, Strada le Grazie 15, 37134 Verona, Italy

Correspondence should be addressed to Luca Di Persio; dipersioluca@gmail.com

We propose a copula-based approach to solve the option pricing problem in the risk-neutral setting and with respect to a structured derivative written on several underlying assets. Our analysis generalizes similar results already present in the literature but limited to the trivariate case. The main difficulty of such a generalization consists in selecting the appropriate vine structure which turns to be of D-vine type, contrary to what happens in the trivariate setting where the canonical vine is sufficient. We first define the general procedure for multivariate options and then we will give a concrete example for the case of an option written on four indexes of stocks, namely, the S&P 500 Index, the Nasdaq 100 Index, the Nasdaq Composite Index, and the Nyse Composite Index. Moreover, we calibrate the proposed model, also providing a comparison analysis between real prices and simulated data to show the goodness of obtained estimates. We underline that our pair-copula decomposition method produces excellent numerical results, without restrictive assumptions on the assets dynamics or on their dependence structure, so that our copula-based approach can be used to model heterogeneous dependence structure existing between market assets of interest in a rigorous and effective way.

1. Introduction

In what follows we will consider a European option written on 4 assets. We will assume that the risk-neutral setting holds true, according to the framework defined in [1]; that is, the price of each of the four considered underlying assets only depends on its history; moreover, it is independent form the others past behaviour. Previous condition allows us to write the price of the above mentioned option as the following discounted value:

$$p_t = e^{-r_f(T-t)} \mathbb{E}_{\mathbb{Q}} \left[G\left(S_1(T), \ldots, S_4(T)\right) \mid \mathcal{F}_t \right], \quad (1)$$

where $T > 0$ is the maturity time of the option, r_f is a real positive parameter usually representing the interest rate given by a bank for our cash deposit and it is assumed to be constant over the whole option's life, $G(S_1(T), \ldots, S_4(T))$ is the payoff of the option written on the four assets $S_1, \ldots, S_4$, whose prices, at any time $t \in [0, T]$, are $S_i(t)$ for $i = 1, \ldots, 4$. We would like to underline that (1) is given directly under the risk-neutral (martingale) measure $\mathbb{Q}$ and it represents

the (fair or no-arbitrage) price of the option with payoff G determined by the so-called martingale approach, see, for example, [2, Section 5.4].

Under suitable assumptions, the price determined in (1) can be rewritten according to the expected value definition as follows:

$$p_t = e^{-r_f(T-t)} \iint_0^{+\infty} G\left(s_1, \ldots, s_4\right) f_{\mathbb{Q}}\left(s_1, \ldots, s_4\right) ds_1 \cdots ds_4, \quad (2)$$

where $f_{\mathbb{Q}}(\cdot)$ is the joint density probability function of the underlying assets with respect to the risk-neutral probability measure $\mathbb{Q}$.

Our aim is to apply a pair-copula decomposition approach to reproduce the joint density of the 4 underlying assets as correlated random variables. Latter goal requires first to find a method able to (statistically) describe the behaviour of each underlying asset. In particular, we assume that the underlying assets returns evolve as generalized autoregressive conditional heteroskedastic processes with parameters

$p = 1$ and $q = 1$, namely, a GARCH(p, q) process. Latter choice is motivated by the fact that the GARCH processes properly describe the evolution of variables that do not have constant volatility over time; see, for example, [3]. Therefore, GARCH processes turn out to be very effective models for heteroskedastic processes, as in the case of financial time series which exhibit structured interrelations.

In particular, in what follows we consider 4 assets, namely, $S_1, \ldots, S_4$ and a discrete set of times whose elements represent a day between day 0, beginning of transactions, and day T which can be taken as the end of our daily observations of the asset's prices. The quantity $S_i(t)$, for $t \in \{0, \ldots, T\}$ and $i \in \{1, \ldots, 4\}$, stands for the *closing price* of the underlying i at the trading day t. The one-day log-return of the ith asset is defined as

$$r_{i,t+1} := \log\left(\frac{S_i(t+1)}{S_i(t)}\right), \tag{3}$$

$$(i, t) \in \{1, \ldots, 4\} \times \{0, \ldots, T-1\}.$$

Note that the objective 4-variate distribution of the log-returns is specified conditional to all return information available at time t, that is conditional to the sigma algebra $\mathscr{F}_t := \sigma((r_{1,s}, \ldots, r_{4,s}) : s \leq t)$.

In order to derive the joint risk-neutral return process from the objective one, we assume that the objective marginal distributions of the assets returns evolve as a GARCH$(1, 1)$ process with Gaussian innovations; see, for example, [4]; namely,

$$r_{i,t+1} = \mu_i + \eta_{i,t+1},$$

$$\sigma_{i,t+1}^2 = \gamma_i + \alpha_i \eta_{i,t+1}^2 + \beta_i \sigma_{i,t}^2, \tag{4}$$

$$\eta_{i,t+1} \mid \mathscr{F}_t \sim_P N\left(0, \sigma_{i,t}^2\right),$$

where, for every $i = 1, \ldots, 4$, μ_i is the expected daily log-return of the asset S_i, while $\eta_{i,t+1}$ is the Gaussian innovation under the objective, or *real world*, probability $\mathbb{P}$ linked to the return $r_{i,t+1}$. In particular, $\eta_{i,t+1}$, conditioned to all returns information available at time t, that is conditioned to $\mathscr{F}_t$, has mean zero and variance $\sigma_{i,t}^2$ which evolves as a GARCH$(1, 1)$ processes with parameters $\gamma_i, \alpha_i, \beta_i > 0$ such that $\alpha_i + \beta_i < 1$.

Since the marginal distributions, at a given time t, are specified conditional to a common set of information, that is, with respect to the σ-algebra $\mathscr{F}_t$, then we are allowed to exploit copula theory techniques to derive the joint conditional distribution. In what follows, inspired by [5, 6], see also [1, 7], we assume that the objective and the risk-neutral copulas are the same.

The main idea behind our option pricing model is that we can use a convenient map to transform each marginal objective distribution to its risk-neutral counterpart, as in [4]. Then we define a proper 4-dimensional copula function, obtained by the pair-copula construction method (see, for example, [8]) instead of deriving the joint risk-neutral distribution directly. Finally, we determine the fair price of the option by taking the discounted expected value of the option's payoff under the risk-neutral measure.

In particular, assuming that the risk-neutral probability $\mathbb{Q}$ satisfies a *local risk-neutral valuation relationship* (LRNVR) (see [4, Def.2.1, Th.2.2]), then the law of the returns under $\mathbb{Q}$ is given by

$$r_{i,t+1} = r_f - \frac{1}{2}\sigma_{i,t}^2 + \eta_{i,t+1}^*,$$

$$\sigma_{i,t+1}^2 = \gamma_i + \alpha_i \left(r_{i,t+1} - \mu_i\right)^2 + \beta_i \sigma_{i,t}^2, \tag{5}$$

$$\eta_{i,t+1}^* \mid \mathscr{F}_t \sim_\mathbb{Q} N\left(0, \sigma_{i,t}^2\right),$$

where $\eta_{i,t+1}^*$ are the Gaussian innovations under $\mathbb{Q}$. Exploiting the transform defined in (5) it is possible to model the marginal behaviour of the marginal log-returns under the measure $\mathbb{Q}$.

1.1. Pair-Copula Decomposition.

In this section, we show how to model the interdependence of the underlying assets $S_1, \ldots, S_4$, under the risk-neutral measure $\mathbb{Q}$, by a copula-based approach. We would like to underline that in the d dimensional case for $d > 2$ and differently from the bivariate one, the main difficulty is that of finding the right d-dimensional copula which properly reproduces the dependence structure existing between the single pairs of components. We solve such a problem by a multivariate copula construction method, namely, the pair-copula decomposition approach; see, for example, [8].

Let us start introducing some basic concepts for the pair-copula representation in view of the analysis of the three-dimensional case; see [9].

Definition 1 (tree). A tree $T = \{N, E\}$ is an acyclic graph, where N is its set of nodes and E is its set of edges (unordered pairs of nodes).

Definition 2 (regular vine). A regular vine tree on d variables is a structure of connected trees $T_1, \ldots, T_{d-1}$, with nodes N_i and edges E_i, $i = 1, \ldots, d - 1$, such that the following conditions are satisfied:

(1) T_1 is a tree with a set of nodes $N_1 = \{1, \ldots, d\}$ and a set of edges denoted by E_1;

(2) for $i = 2, \ldots, d - 1$ the tree T_i has node set $N_i = E_{i-1}$ and a set of edges denoted by E_i;

(3) two edges in tree T_i are joined in tree T_{i+1} if they share a common node in tree T_i.

Here we focus on two special cases of regular vines, namely, the *D-vine* and the *canonical vine*; see, for example, [8] for details.

Before going into details, we would like to underline that while from the pair-copula decomposition point of view the three-dimensional case can be easily treated (for example, the D-vine and the canonical vine specification coincide), difficulties arise when the number of correlated dimensions increases.

In three dimensions the construction method proceeds as follows: let f denote the joint density probability function of three random variables X_1, X_2, X_3, and then f can be decomposed as follows:

$$
\begin{aligned}
& f\left(x_1, x_2, x_3\right) \\
& \quad = f_1\left(x_1\right) f_2\left(x_2\right) f_3\left(x_3\right) c_{12}\left(F_1\left(x_1\right), F_2\left(x_2\right)\right) \\
& \qquad \times c_{23}\left(F_2\left(x_2\right), F_3\left(x_3\right)\right) c_{13|2} \\
& \qquad \times \left(F_{1|2}\left(x_1 \mid x_2\right), F_{3|2}\left(x_3 \mid x_2\right)\right),
\end{aligned}
\tag{6}
$$

where, for every $i, j = 1, \ldots, 3$, F_i stands for the marginal probability distribution function of X_i, f_i is the corresponding marginal density function, and c_{ij} is the associated bivariate copula density, while $c_{13|2}$ represents the bivariate copula density of X_1, X_3 conditional to the second component.

In order to reproduce a density function through a pair-copula decomposition in dimension $d > 3$, we have first to select an appropriate decomposition structure. In particular, from a financial point of view, previous statement means that we first have to analyze the economic relationship between the variables of interest in order to find the variable that governs the interactions in the data set. If such a variable exists, then a canonical vine specification should be the best choice; otherwise, we will use the D-vine decomposition structure; see, for example, [8].

Once the vine structure has been chosen, we are left with the selection of the order of the variables, the families of pair-copulas, and the specification of their parameters.

The order of the variables is set according to the following rule: first we compute the dependence measure Kendall's tau for each pair of variables and we order the variables in the first tree on the basis of their dependence structure; this means that the two variables with the strongest dependence are put in the first two nodes of the first tree and so on. The next trees are determined as a consequence; see [10]. Then we select the best-fitting pair-copula family; namely, we have to select the pair-copula specification that guarantees the best fitting with observed data. Such a problem can lead to different solutions depending on different setting. In our specific case, we are justified to consider only the Gaussian, the t-Student, the Clayton and the Gumbel copulas since they are the most used types in financial applications; see, for example, [1, 5–7]. The particular copula is selected exploiting standard information criteria such as Akaike information criterion (AIC) and the Bayesian information criterion (BIC) respectively; see, for example, [11–13] and references therein, for each edge in each tree. In particular, the choice is made in order to minimize the AIC and the BIC coefficient, respectively. Finally, the selected copula's parameters are estimated using the maximum likelihood criterion for each pair.

Once we have obtained the results for the first tree, we need to construct a sample for the conditional bivariate distribution, in order to find the pair-copula associated with the copula density $c_{13|2}$ in (6). Let us consider the following

definition of bivariate conditional distribution function in terms of copula:

$$
h(u, v, \Theta) := F(u \mid v) = \frac{\partial C_{u,v}(u, v, \Theta)}{\partial v},
\tag{7}
$$

where u and v are uniform random variables and Θ is the set of parameters characterizing the copula of the joint distribution of u and v.

We would like to underline that the function h plays a key role in the pair-copula decomposition approach in a d-dimensional setting when $d > 3$, since it allows to reproduce the conditional behaviour of a random vector in terms of a bivariate function. Indeed, by an iterative algorithm, we can rewrite the conditional distribution using a proper choice of h and of the copula written on its marginals.

Once the pair-copula method has been theoretically implemented, thus obtaining the joint distribution of interest, we are left with the need to calibrate the model in order to obtain satisfactory numerical results. The calibration procedure will be the main goal of the next section.

1.2. Calibration of the Model. The pair-copula decomposition and the h-function described in Section 1.1 can be calibrated starting from the definition of the Gaussian innovations $\eta_{i,t}$ of the model, namely, a GARCH$(1, 1)$, we have chosen for the marginal distributions of the underlying assets.

In particular, if we are in a d-dimensional setting, $d \geq 2$, we define the standardized innovations as follows:

$$
\left(Z_{1,s}, Z_{2,s}, \ldots, Z_{d,s}\right)_s := \left(\frac{\eta_{1,s}}{\sigma_{1,s}}, \frac{\eta_{2,s}}{\sigma_{2,s}}, \ldots, \frac{\eta_{d,s}}{\sigma_{d,s}}\right),
\tag{8}
$$
$$
\forall (i, s) \in \{1, \ldots, d\} \times [0, T],
$$

where $\sigma_{i,s}$ is the standard deviation of the underlying i at time s. By the well-known results for the GARCH model (see, for example, [14, Part V, Sec.16]) the $Z_{i,s}$ are i.i.d, standard Gaussian random variables, even if the stochastic processes $Z_1, \ldots, Z_d$ are not independent. By the Sklar's Theorem we have that the joint distribution F (under the objective measure $\mathbb{P}$) of the underlying innovations can be written in terms of its marginals; see, for example, [15, Sec.1]. In particular, since $Z_1, \ldots, Z_d$ are continuous stochastic processes, there exists an unique copula C such that

$$
F\left(z_1, \ldots, z_d\right) = C\left(F_1\left(z_1\right), \ldots, F_d\left(z_d\right)\right),
\tag{9}
$$

for all $z_i \in \mathbb{R}$, $i = 1, \ldots, d$. Moreover, we assume that the copula C in (9) is a parametric copula of parameter $\theta_{\mathbb{P}}$. Latter assumption is justified since the dependence structure existing between the variables z_i is usually given by a copula function that depends on a particular parameter. Given the standardized innovations, we want to infer their dependence structure under $\mathbb{P}$ using the pair-copula decomposition method. Then, we will use the dependence structure, namely, the copula function that better reproduces the joint behaviour of the underlyings, to price an option written on four indexes which will be considered as underlyings.

In order to use the pair-copula decomposition, and in particular the h-function defined in (7), we compute

the corresponding estimated standardized innovations in the interval $(0, 1)$, applying the cumulative distribution function Φ of the standard normal distribution $N(0, 1)$. Thus, the corresponding uniform variables U_i, for $i = 1, \ldots, 4$ are defined as follows:

$$(U_{1,s}, \ldots, U_{4,s}) := \left(\Phi\left(\frac{\widehat{\eta}_{1,s}}{\widehat{\sigma}_{1,s}} \right), \ldots, \Phi\left(\frac{\widehat{\eta}_{4,s}}{\widehat{\sigma}_{4,s}} \right) \right). \tag{10}$$

Hereafter, we will work using the random vector $(U_{1,s}, \ldots, U_{4,s})$, with $s \in [0, T]$, where $T > 0$ is the maturity date of our *investment*.

The next step is that of finding the copula on the joint distribution of the random vector $U_1, \ldots, U_4$. In particular, we have to choose the bivariate copulas that best fit our data. Such a copula structure will be then exploited to reproduce the 4-variate copula density and thus the joint distribution density, for the random vector $U_1, \ldots, U_4$ under the objective measure $\mathbb{P}$.

1.3. Option Pricing. In this section, we will consider (2) to solve the associated option pricing problem exploiting results obtained through previous sections. To reproduce the joint density function $f_{\mathbb{Q}}$ we use, as for the density under the objective probability $\mathbb{P}$, a multivariate copula constructed through the pair-copula approach. In particular, we assume that the copula under $\mathbb{Q}$ belongs to the same family as the one determined under $\mathbb{P}$, possibly being characterized by different parameters.

The option pricing problem will be solved in two steps. First we simulate the underlying assets $S_1, \ldots, S_4$ under the risk-neutral measure $\mathbb{Q}$; then, we simulate the option price and we calibrate the model to determine the set $\Theta_{\mathbb{Q}}$ of the copulas' parameters under $\mathbb{Q}$. The set $\Theta_{\mathbb{Q}}$ is composed of all the parameters of the pair-copulas used to create the proper vine specification. For example, in the 3-dimensional case, we have $\Theta_{\mathbb{Q}} = (\theta_{12}^{\mathbb{Q}}, \theta_{13}^{\mathbb{Q}}, \theta_{23|1}^{\mathbb{Q}})$, for such a particular vine structure.

In order to simulate the assets $S_1, \ldots, S_4$ under $\mathbb{Q}$ we consider the risk-neutral transformation of the returns defined in (5); see [4] for details. This turns out to be a recursive method to estimate both the returns and the volatility terms under the risk-neutral measure. Moreover, in order to maintain the dependence structure obtained from the market data, we simulate, using the vine specification previously defined, the variables U_i, $i = 1, \ldots, 4$, according to the algorithm described in [10].

Concerning the option price estimation, let us recall that at time t the option price is given by the following equation:

$$p_t = e^{-r_f(T-t)} \mathbb{E}_{\mathbb{Q}} \left[G\left(S_1(T), \ldots, S_4(T)\right) \mid \mathcal{F}_t \right]. \tag{11}$$

We first estimate the price, defined as p_t^M, inserting the prices $S_i(t)$, observed in the market at time t, for $i = 1, \ldots, 4$, into the option payoff equation. Then we use a Monte Carlo approach to estimate the price, defined as $\widehat{p}_t^{mc}(\Theta_{\mathbb{Q}})$, by (1) and using the observation simulated from the vine structure.

Finally, we calculate the set $\Theta_{\mathbb{Q}}$ minimizing the sum of the quadratic error; namely,

$$\min_{\Theta_{\mathbb{Q}}} \sum_{i=1}^{n} \left(\widehat{p}_{t_i}^{mc} - p_{t_i}^{M} \right)^2, \tag{12}$$

where $\{t_i\}_i$ are the past observations of the option prices. The question about the choice of the parameter set $\Theta_{\mathbb{Q}}$ corresponds to the calibration of the pricing model.

2. Numerical Implementation

In our implementation, we work with a dataset that comes from the U.S market. In particular, we consider the following indexes: the Standard and Poor's (S&P) 500 index, the Nasdaq 100 index, the Nasdaq Composite index, and the New York Stock Exchange (Nyse) Composite index; see, for example, the related Wikipedia occurrences for detailed definitions of these market indexes. We have considered such indexes between January 1, 2012 and December 28, 2012, for a total of 249 days resulting in the same amount of *closing levels* for each index. On these indexes we write two option contracts, X_1 and X_2, respectively, which are defined by the following payoffs at maturity $T > 0$:

$$X_1(T) := \max\left(\max\{S_i(T)\} - K, 0\right), \tag{13}$$

$$X_2(T) = \max\{S_i(T)\} - \min\{S_i(T)\}, \tag{14}$$

respectively, where $K > 0$ is the strike price of the option X_1 and $i = 1, \ldots, 4$. Note that $K = 1000$ in our numerical implementation. In Figure 1 we plot the four data sets.

Notice that the plots of the closing prices of the four indexes are not comparable, since they do not fluctuate around a common mean value. Hereafter, we will consider the log-return of each index instead of time series of the prices, as in (3). By an analysis of the time series of the returns, in Figure 2, we can infer an heteroskedasticity characteristic for the considered setting; hence, a GARCH model turns out to be considered as a natural choice.

The GARCH parameters estimation is the first step of our implementation. In particular, once the time series of the underlying assets prices are given, we compute the returns estimation. Before analyzing the results obtaining using the GARCH model, let us underline some statistics for each return series together with the response given by the Jarque-Bera normality test; see, for example, [16], for each series. From Table 1 we see that the Jarque-Bera statistics are quite high, especially for the S&P 500 series; thus we can infer that the returns are not normally distributed.

Moreover, both the kurtosis coefficients (greater than 3 which is the characteristic kurtosis value for a Gaussian distribution) and the quantile-quantile plot of the returns, as in Figure 3, strengthen previous deduction.

We would also like to underline the presence of fat tail for the S&P 500 time series of returns, in agreement with the (non-Gaussianity) results given by the Jarque-Bera statistics. We point out that using the GARCH(1, 1) model we obtain a highly satisfactory description of our market behaviour;
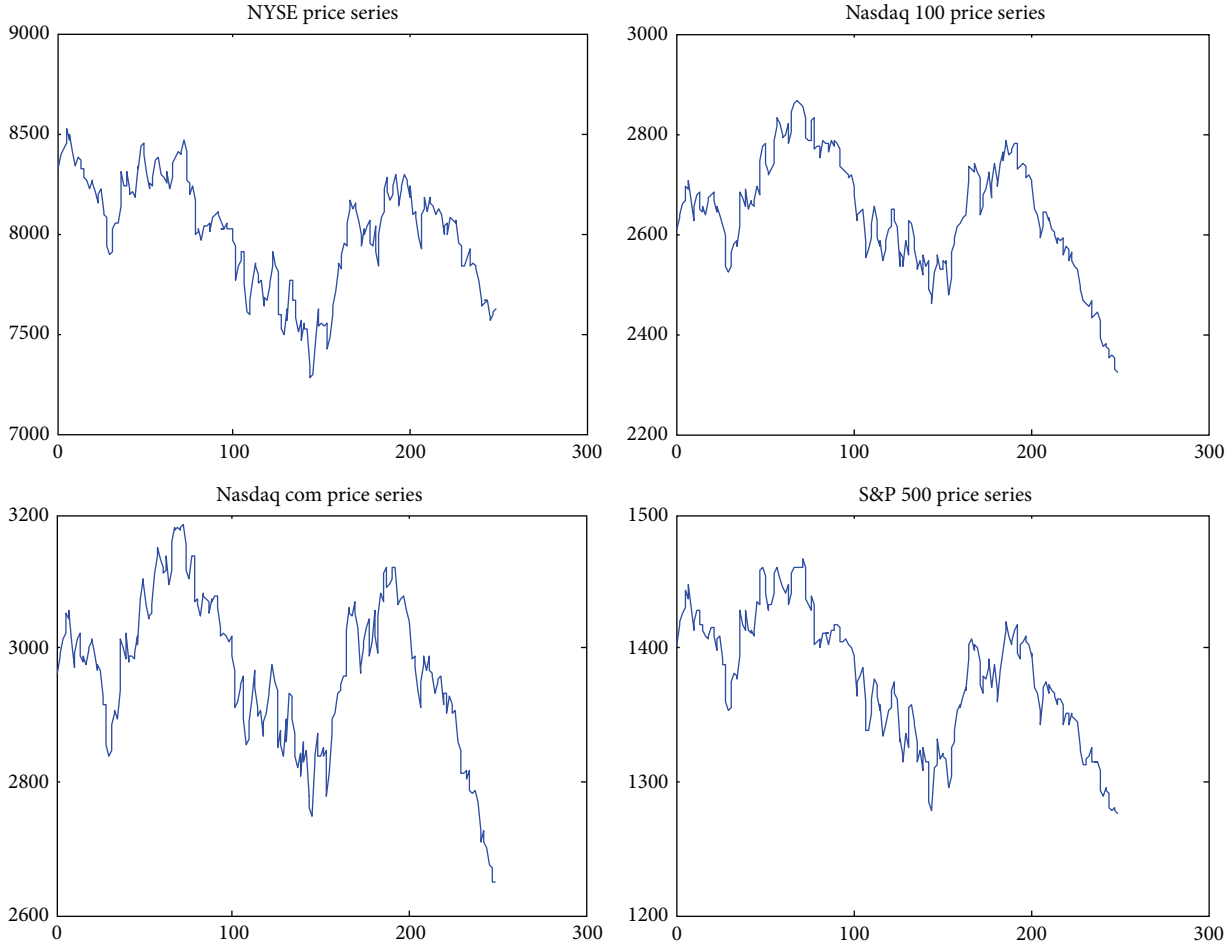
FIGURE 1: Plot of the data sets.

TABLE 1: Descriptive statistics.

	NYSE	Nasdaq 100	Nasdaq Com	S&P 500
Mean	0.350	0.466	0.448	0.377
Standard deviation	0.0085	0.0096	0.0093	0.0079
Kurtosis	3.549	3.685	3.629	3.915
JB statistics	3.18	5.07	4.09	8.68

TABLE 2: Estimated coefficients and corresponding standard errors.

	NYSE	Nasdaq 100	Nasdaq Com	S&P 500
$\widehat{\mu}$	$5.03e - 04$	$9.38e - 04$	$7.47e - 04$	$5.29e - 04$
	$(5.43e - 04)$	$(6.03e - 04)$	$(5.98e - 04)$	$(5.13e - 04)$
$\widehat{\gamma}$	$3.64e - 06$	$2.93e - 06$	$3.19e - 06$	$3.02e - 06$
	$(4.53e - 06)$	$(3.57e - 06)$	$(4.24e - 06)$	$(3.60e - 06)$
$\widehat{\alpha}$	0.0471	0.0488	0.0419	0.0415
	(0.0372)	(0.0315)	(0.0324)	(0.0318)
$\widehat{\beta}$	0.9035	0.9220	0.9229	0.9127
	(0.0899)	(0.0592)	(0.0710)	(0.0796)

TABLE 3: Sample Kendall's tau.

	NYSE	Nasdaq 100	Nasdaq Com	S&P 500
NYSE	1			
Nasdaq 100	0.5224	1		
Nasdaq Com	0.6018	0.8985	1	
S&P 500	0.7286	0.7320	0.7548	1

indeed all the processes have a log-likelihood value rather high, between 800 and 850. Moreover, it results that the most significant variable in the model is the one corresponding to the variance past value, with the latter being in agreement with the use of a heteroskedastic model. In Table 2 we present the estimated coefficients and the corresponding estimation errors of the coefficients of our GARCH(1, 1) process for each index.

In order to construct a proper pair-copula structure for the multivariate density function, we first have to find the vine structure that best fits our data, and then we will properly order the variables in the chosen vine. We first investigate the dependence structure between pairs out of {Nyse, Nasdaq 100, Nasdaq Com, S&P 500}, and the result, obtained using the sample Kendall's tau, is given in Table 3.

Using the data collected in Table 3, we can infer that there is not a variable that governs the interactions. Thus
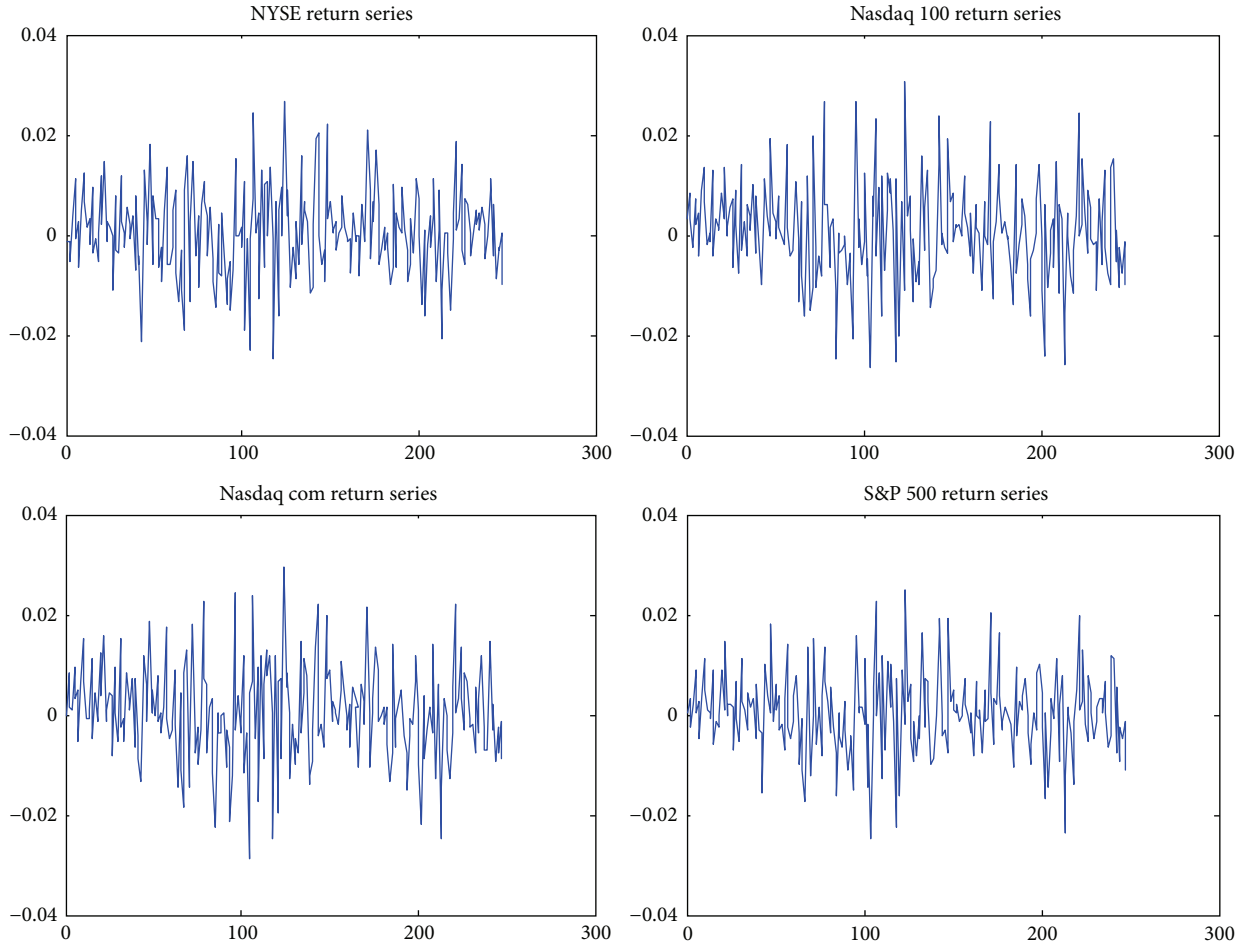
FIGURE 2: Plot of the indexes returns.

we will consider a D-vine structure for the pair-copula decomposition. Moreover, we can see that the strongest dependence appears to be, as expected, between the Nasdaq 100 Index and the Nasdaq Composite Index, followed by the one between the S&P 500 Index and the Nasdaq Composite Index. The D-vine structure for our data set is given in Figure 4.

Each edge of the D-vines tree corresponds to a pair-copula density; therefore, in our implementation, the pair-copula decomposition of the joint density reads as follows:

$$
\begin{aligned}
f &\left(x_N, x_{Nc}, x_S, x_{Ny} \right) \\
&= f_N \left(x_N \right) \cdot f_{Nc} \left(x_{Nc} \right) \cdot f_S \left(x_S \right) \cdot f_{Ny} \left(x_{Ny} \right) \\
&\quad \cdot c_{N,Nc} \left(F_N \left(x_N \right), F_{Nc} \left(x_{Nc} \right) \right) \cdot c_{Nc,S} \left(F_{Nc} \left(x_{Nc} \right), F_S \left(x_S \right) \right) \\
&\quad \cdot c_{S,Ny} \left(F_S \left(x_S \right), F_{Ny} \left(x_{Ny} \right) \right) \\
&\quad \cdot c_{N,S|Nc} \left(F_{N|Nc} \left(x_N \mid x_{Nc} \right), F_{S|Nc} \left(x_S \mid x_{Nc} \right) \right) \\
&\quad \cdot c_{Nc,Ny|S} \left(F_{Nc|S} \left(x_{Nc} \mid x_S \right), F_{Ny|S} \left(x_{Ny} \mid x_S \right) \right) \\
&\quad \cdot c_{N,Ny|Nc,S} \left(F_{N|Nc,S} \left(x_N \mid x_{Nc}, x_S \right), \right. \\
&\qquad\qquad\qquad \left. F_{Ny|Nc,S} \left(x_{Ny} \mid x_{Nc}, x_S \right) \right).
\end{aligned}
$$

$$(15)$$

Given the pair-copula decomposition, we further investigate the dependence structure constructing the bivariate scatter plots of the copula data U_i, for $i = 1, \ldots, 4$, where U_i is the uniform vector in (10). Graphs in Figures 5 and 6 show the existence of tail dependence between almost all the pairs. In particular, we can easily notice that there is significant dependence between different time series concerning their tails. The latter fact is rather obvious since all the used financial indexes belong to the USA market, and hence, even if their behaviour is also influenced by non-USA assets, a quite strong link between them is exactly what we have to take into account. We would also like to underline that we may expect similar results, namely, the existence of tail dependence between almost all the pairs, even in the case of pairs constituted by indexes belonging to different economies, for example, DAX and NYSE, at least when rather global financial events occur, as in the case of the worldwide economic crisis of 2007/2008. It is worth to mention that, in oder to study such type of dependencies, one can make use of the so-called upper/lower dependence indexes, which can be considered as indicators of the dependence values in the tails of the bivariate distributions.

Previous consideration is the first step in the copula choosing process. The next step, actually the most important
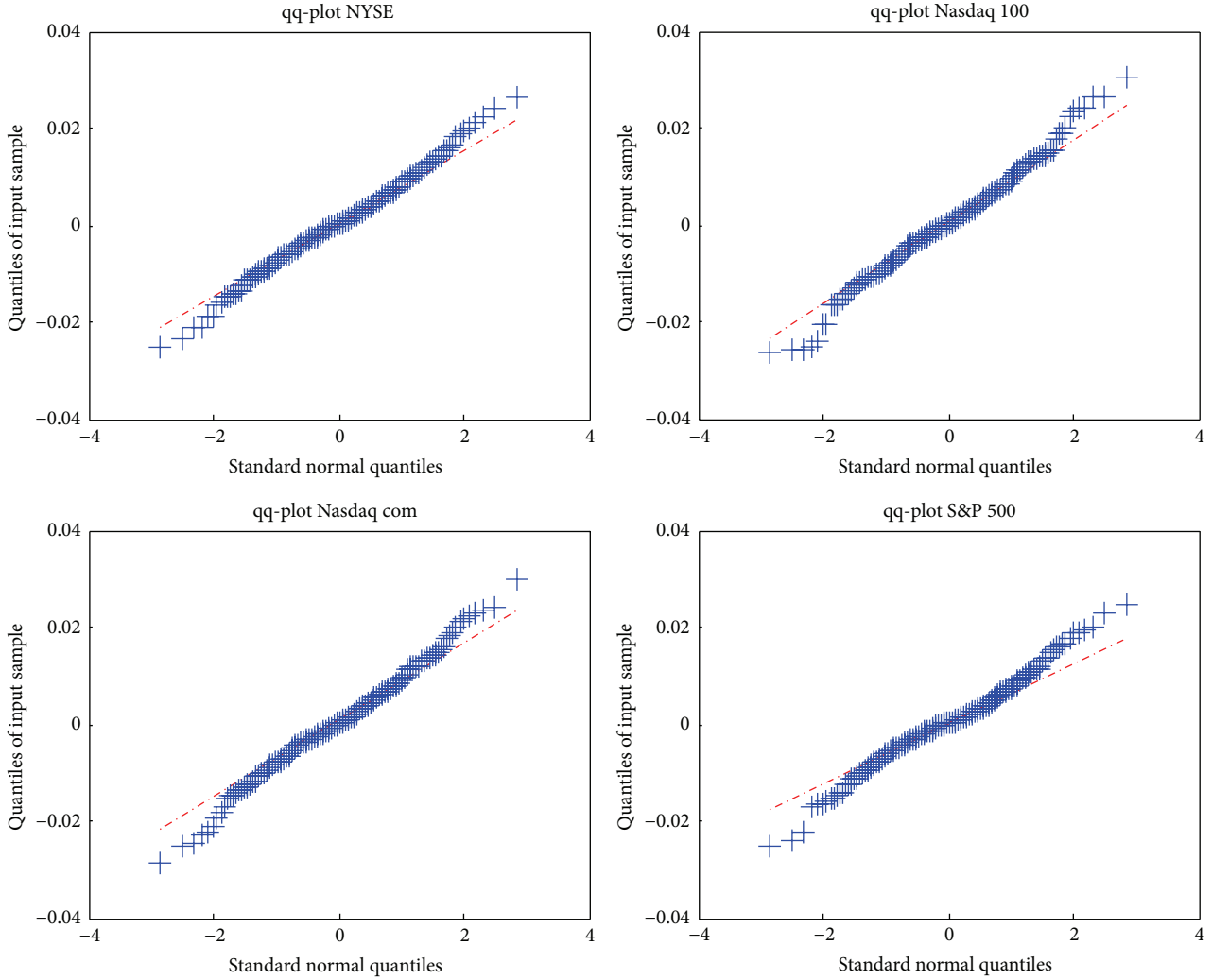
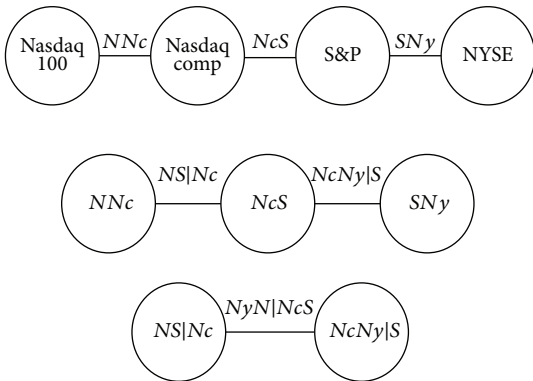FIGURE 3: Quantile-quantile plot of the indexes returns.

FIGURE 4: D-vine structure with four variables.

the corresponding copula functions' parameters. From Figure 4 we infer that we have to compute the following copulas: $C_{N,Nc}$, $C_{Nc,S}$, $C_{S,Ny}$, $C_{N,S|Nc}$, $C_{Nc,Ny|S}$, and $C_{N,Ny|Nc,S}$.

The first three copulas are estimated using the copula data of the vector U and comparing the AIC values for each possible choice of the copula function. The results are as follows:

  (i) $C_{N,Nc}$ is a Clayton copula;

  (ii) $C_{Nc,S}$ is a Clayton copula too;

  (iii) $C_{S,Ny}$ is a Gumbel copula.

Once the copula functions have been estimated, we can compute the corresponding parameters, which are reported in Table 4.

From the second tree of Figure 4 it turns out that we have to compute $C_{N,S|Nc}$ and $C_{Nc,Ny|S}$. Since they are conditional pair copula we have to infer first the proper h-function. Then, we can estimate the conditional copula using the sample created by the h-functions. In a four-variable D-vine structure we have to compute three h-functions, since
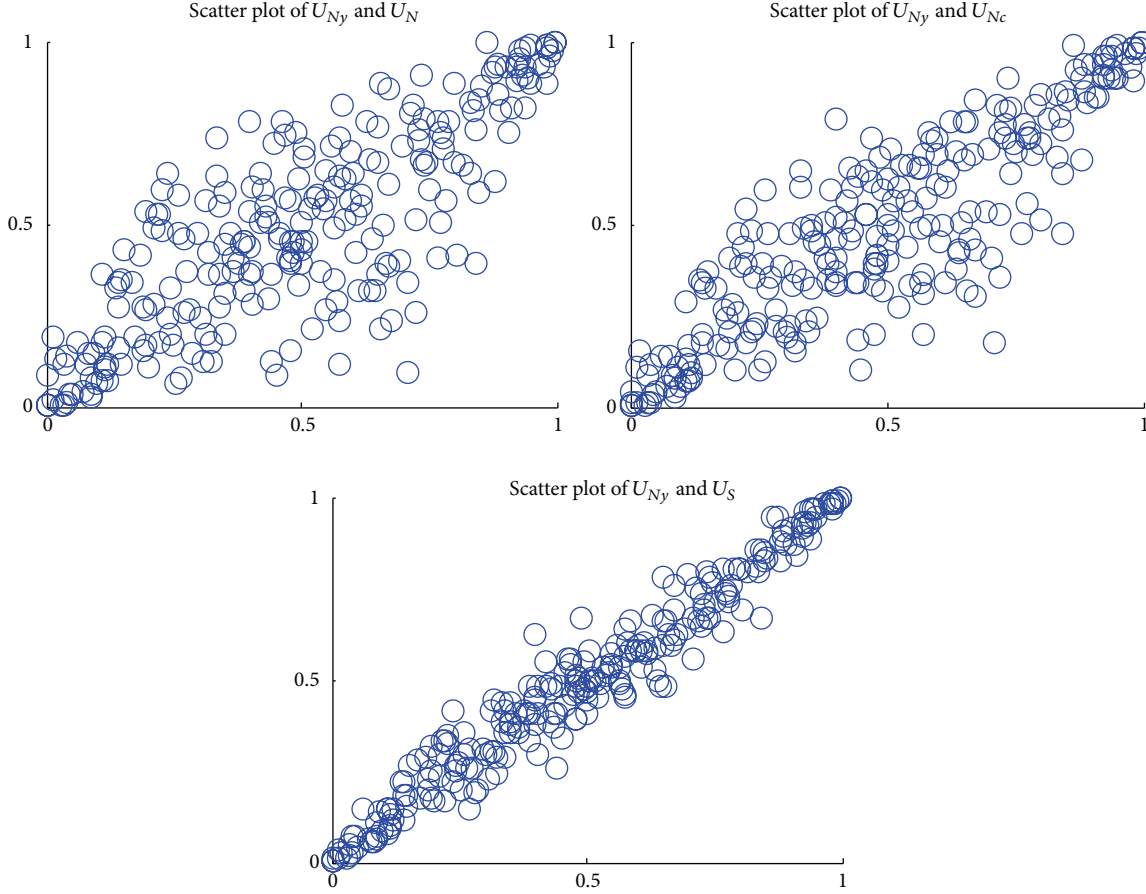
one, is that of comparing the AIC values for the four possible copula alternatives, that is, the Gaussian, the T-Student, the Clayton, and the Gumbel copula. Let us first analyze the bivariate copulas of the first tree, and then we will estimate

Scatter plot of U_{Ny} and U_N

Scatter plot of U_{Ny} and U_{Nc}

Scatter plot of U_{Ny} and U_S

FIGURE 5: Scatter plots of the copula data sets.

TABLE 4: Copula parameters estimation.

	$\hat{\theta}$
$C_{N,Nc}$	9.2084
$C_{Nc,S}$	3.5953
$C_{S,Ny}$	7.2419
$C_{N,S|Nc}$	1.3708
$C_{Nc,Ny|S}$	1.00
$C_{N,Ny|Nc,S}$	1.00

we need to estimate the following conditional distribution functions: $F(x_{Nc} \mid x_S)$, $F(x_{Ny} \mid x_S)$ and $F(x_N \mid x_{Nc})$. Hence we simulate $F(x_{Nc} \mid x_S)$ using the Clayton h-function $hC(x_{Nc}, x_S)$, while for $F(x_{Ny} \mid x_S)$ we use the Gumbel h-function $hGu(x_{Ny}, x_S)$, and finally the corresponding h-function for $F(x_N \mid x_{Nc})$ is again the Clayton one. Once the conditional distributions have been simulated, we can estimate the best-fitting copula functions for the pairs $(F(x_N \mid x_{Nc}), F(x_S \mid x_{Nc}))$ and $(F(x_{Nc} \mid x_S), F(x_{Ny} \mid x_S))$, and we have that the selected copula is for both of the pairs the Gumbel copula, with parameters as in Table 4. Finally, the copula $C_{N,Ny|Nc,S}$ has to be computed. Again we have to use the h-function to simulate a sample for the conditional distribution functions $F(x_N \mid x_{Nc}, x_S)$ and $F(x_{Ny} \mid x_{Nc}, x_S)$.

The last pair copula is a Gumbel copula function too. Note that for both $C_{Nc,Ny|S}$ and $C_{N,Ny|Nc,S}$ we have $\hat{\theta} = 1$, which implies that the variables in the distribution function are conditionally independent so that, for the copula $C_{Nc,Ny|S}$ and given the S&P Index, the Nasdaq Composite and the Nasdaq 100 are independent. Similarly, given the Nasdaq Composite and the S&P, the Nasdaq 100 and the NYSE are conditionally independent. Moreover, we have that the density copula function $c_{N,Ny|Nc,S}$ is equal to 1; see [10]; therefore, such copula components do not play a role in our pair-copula decomposition structure.

Exploiting previous considerations about the estimated parameters, we obtain that the joint density function can be decomposed as follows:

$$
\begin{aligned}
f\left(x_N, x_{Nc}, x_S, x_{Ny}\right) \\
= f_N\left(x_N\right) \cdot f_{Nc}\left(x_{Nc}\right) \cdot f_S\left(x_S\right) \cdot f_{Ny}\left(x_{Ny}\right) \\
\cdot c_{N,Nc}^{\text{Clayton}}\left(F_N\left(x_N\right), F_{Nc}\left(x_{Nc}\right)\right) \\
\cdot c_{Nc,S}^{\text{Clayton}}\left(F_{Nc}\left(x_{Nc}\right), F_S\left(x_S\right)\right) \\
\cdot c_{S,Ny}^{\text{Gumbel}}\left(F_S\left(x_S\right), F_{Ny}\left(x_{Ny}\right)\right) \\
\cdot c_{N,S|Nc}^{\text{Gumbel}}\left(F_{N|Nc}\left(x_N \mid x_{Nc}\right), F_{S|Nc}\left(x_S \mid x_{Nc}\right)\right).
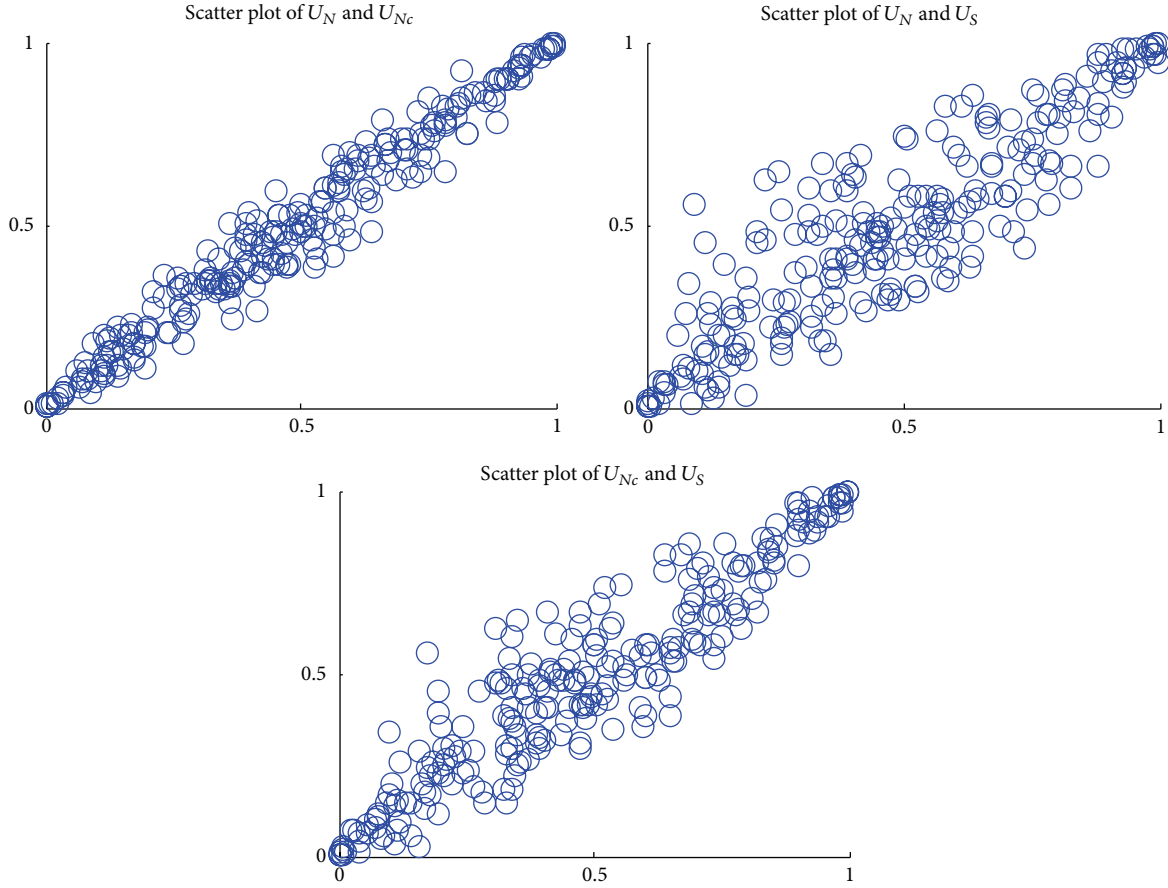\end{aligned}
\tag{16}
$$

FIGURE 6: Scatter plots of the copula data sets.

2.1. Option Pricing with Four Underlying Assets. In this section we analyze how the option pricing problem can be solved using our pair-copula decomposition.

We first consider the option pricing problem for an option written on four underlyings whit payoff given as in (13). Then we consider the same problem, but for an option with payoff as in (14). Following what we have developed in the previous sections from the theoretical model, we first compute the option price using the market prices at the final time and then we compare it with the prices under the risk-neutral probability $\mathbb{Q}$ using the same pair-copula decomposition, but with proper parameters. To find the optimal parameters set we use a sort of confidence interval for the possible values $\theta_i^{\mathbb{Q}}$. In particular, starting from the corresponding $\theta_i^{\mathbb{P}}$ computed under the objective probability measure $\mathbb{P}$, we look for the values that belong to a predetermined interval such that they minimize the difference between the market price and the estimated one, as in (12). Hereafter, we will face the problem of finding the copula parameters which allow obtaining an option price, computed under the risk-neutral measure $\mathbb{Q} \sim \mathbb{P}$, as close as possible to the market option price computed exploiting the observed data. We consider first an option with payoff given by (13) and we compute its market price at the end of our investment period, say $p^M(T)$. Then we compute the option price under the risk-neutral probability

measure $\mathbb{Q}$, and related numerical results (prices) are shown in Table 6. Before computing the option price under the risk-neutral measure $\mathbb{Q}$, we have to derive the assets prices under the probability $\mathbb{Q}$. Let us note that, by assumption, the dependence structure under $\mathbb{Q}$ is the same as the one under the objective probability $\mathbb{P}$. The first step in the method used to construct the proper prices processes is that of simulating a 4-dimensional, uniformly distributed vector in $[0, 1]^4$, say $U^{\mathbb{Q}} = [U_1^{\mathbb{Q}}, U_2^{\mathbb{Q}}, U_3^{\mathbb{Q}}, U_4^{\mathbb{Q}}]$. Such a simulation is iterated up to 248 times, in order to create a sample of data as long as the data set of observations. Then we construct the returns series according to the specification given in (5). To do this we use the coefficients computed with the GARCH(1, 1) model in Table 2, we set an initial volatility level equal to $\gamma_i/(1 - \alpha_i - \beta_i)$, and we exploit the standardized innovations $Z_i^{\mathbb{Q}}$ given by the following identity:

$$Z_{i,t}^{\mathbb{Q}} = \Phi^{-1}\left(U_{i,t}^{\mathbb{Q}}\right). \tag{17}$$

The next step consists in recursively computing the returns series under $\mathbb{Q}$. This procedure allows us to obtain the fact (see [4]) that the final asset price under the risk-neutral probability is defined as follows:

$$S_{i,T} = S_{i,t} \exp\left[(T - t)\,r_f - \frac{1}{2}\sum_{s=t+1}^{T}\sigma_{i,s}^2 + \sum_{s=t+1}^{T}\eta_{i,s}\right], \tag{18}$$

TABLE 5: Copula parameters under $\mathbb{Q}$.

	$\widehat{\theta}_1$	$\widehat{\theta}_2$
$C_{N,Nc}$	10.2084	10.2084
$C_{Nc,S}$	3.5953	2.5953
$C_{S,Ny}$	7.2419	6.2419
$C_{N,S|Nc}$	1.3708	0.8708
$C_{Nc,Ny|S}$	1.00	1.00
$C_{N,Ny|Nc,S}$	1.00	1.00

TABLE 6: Market price and Monte Carlo price.

	Option 1	Option 2
Market price $p^M(T)$	$7.32e + 003$	$6.91e + 003$
Monte Carlo estimate $p^{MC}(T)$	$7.27e + 003$	$6.67e + 003$

where T is the final observation time, r_f is the constant risk-free rate, and $\sigma_{i,t}^2$ is the volatility of the asset i at time t, computed using the second line of (5).

The final step consists in an option price estimation made using both Monte Carlo method, with 10000 simulated trajectories for each underlying asset, and the sample data series previously constructed. The estimate procedure is then completed finding the copula parameters that minimize the differences between the market option price at time T, that is, $p^M(T)$, and the Monte Carlo simulated price, say $p^{MC}(T)$. In Table 5, we present the estimated copula parameters that minimize the difference between p^M and p^{MC} both for an option with payoff as in (13), say $\widehat{\theta}_1$, and for an option with payoff as in (14), say $\widehat{\theta}_2$.

Table 6 reports the prices computed for an option with payoff as in (13), say Option 1, respectively, for a second option with payoff as in (14), say Option 2.

Using data collected in Table 6, we can infer that the price computed using the Monte Carlo simulation is rather similar to the market price. This means that the copula-based option pricing is a good estimation method to find the option price. The difference between the two prices can be seen as the witnessing of the lack of arbitrage free hypotheses. An alternative explanation relies on the fact that even if $\mathbb{Q} \sim \mathbb{P}$ we do not have that the copula decomposition structure remains the same under these two probability measures.

We would also like to underline that, from a purely computational point of view, the simulation of a sample from the D-vine dependence structure in four dimensions is rather more complicated than in three dimensions because of the increased algorithm complexity. Previous results show that also for the four-dimensional option pricing problem the copula approach gives satisfactory results. Latter approach is quite general and it can be adapted to a larger set of assets distribution. In fact we can use different specification of the copula functions and we can exploit the flexibility given by the pair-copula decomposition.

3. Conclusions

In this paper we have faced a multivariate option pricing problem considering options written on 4 underlying assets by copula decomposition techniques. In particular, inspired by the method proposed in [9, 10], we have strictly generalized similar results already present in literature, see, for example, [1, 5, 6], which are limited to the trivariate case.

The main difficulty to overcome passing from the three-dimensional to the four-dimensional setting is that in the latter case we have to select the appropriate vine structure which turns to be of D-vine type, contrary to what happens in the trivariate setting where the canonical vine is sufficient.

The latter means that a deeper analysis of the relationship between the 4 underlying assets has to be done in order to give correct vine specification.

We have also shown, by accurate simulations, that the particular pair-copula decomposition method which we proposed to solve option pricing problem produces excellent numerical results, without any particular restrictive assumption on the assets behaviour or on the dependence structure existing between them.

Moreover, a copula-based approach has the necessary flexibility needed to face the most different scenarios characterizing present financial markets, and indeed such a technique allows us to model almost every dependence structure existing between the market assets in a precise and effective way.

Conflict of Interests

The authors declare that they have no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] R. W. van den Goorbergh, C. Genest, and B. J. Werker, "Bivariate option pricing using dynamic copula models," *Insurance: Mathematics & Economics*, vol. 37, no. 1, pp. 101–114, 2005.

[2] S. E. Shreve, *Stochastic Calculus for Finance. II*, Continuous-Time Models, Springer Finance. Springer, New York, NY, USA, 2004.

[3] C. Francq and J.-M. Zakoïan, "Inconsistency of the MLE and inference based on weighted LS for LARCH models," *Journal of Econometrics*, vol. 159, no. 1, pp. 151–165, 2010.

[4] J.-C. Duan, "The GARCH option pricing model," *Mathematical Finance*, vol. 5, no. 1, pp. 13–32, 1995.

[5] C. Bernard and C. Czado, "Multivariate option pricing using copulae," *Applied Stochastic Models in Business and Industry*, vol. 29, no. 5, pp. 509–526, 2013.

[6] U. Cherubini and E. Luciano, "Bivariate option pricing with copulas," *Applied Mathematical Finance*, vol. 9, no. 2, pp. 69–85, 2002.

[7] U. Cherubini, E. Luciano, and W. Vecchiato, *Copula Methods in Finance*, Wiley Finance Series, John Wiley & Sons, Chichester, UK, 2004.

[8] C. Czado, "Pair-copula constructions of multivariate copulas," in *Copula Theory and Its Applications*, pp. 93–109, Springer, Berlin, Germany, 2010.

[9] E. C. Brechmann, C. Czado, and K. Aas, "Truncated regular vines in high dimensions with application to financial data," *The Canadian Journal of Statistics*, vol. 40, no. 1, pp. 68–85, 2012.

[10] K. Aas, C. Czado, A. Frigessi, and H. Bakken, "Pair-copula constructions of multiple dependence," *Insurance: Mathematics & Economics*, vol. 44, no. 2, pp. 182–198, 2009.

[11] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974.

[12] H. Akaike, "Likelihood and the Bayes procedure," in *Bayesian Statistics (Valencia, 1979)*, pp. 143–166, University Press, Valencia, Spain, 1980.

[13] K. P. Burnham and D. R. Anderson, "Multimodel inference: understanding AIC and BIC in model selection," *Sociological Methods & Research*, vol. 33, no. 2, pp. 261–304, 2004.

[14] H. Lutkepohl, *New Introduction to Multiple Time Series Analysis*, Springer, Berlin, Germany, 2005.

[15] R. B. Nelsen, *An Introduction to Copulas*, Springer Series in Statistics, Springer, New York, NY, USA, 2nd edition, 2006.

[16] G. G. Judge, R. C. Hill, W. E. Griffiths, H. Lütkepohl, and T.-C. Lee, *Introduction to the Theory and Practice of Econometrics*, John Wiley & Sons, New York, NY, USA, 1983.