Journal of Risk **20**(5), 1–32
DOI: 10.21314/JOR.2018.384

Research Paper

Impact of D-vine structure on risk estimation

Catalina Bolancé, Ramon Alemany and Alemar E. Padilla Barreto

Department of Econometrics, Riskcenter-IREA, University of Barcelona,
Avinguda Diagonal 690, 08034 Barcelona, Spain;
emails: bolance@ub.edu, ralemany@ub.edu, alemarpadillabarreto@ub.edu

(Received September 26, 2016; revised February 27, 2017; accepted March 17, 2017)

ABSTRACT

In this paper, a sensitivity analysis using pair–copula decomposition of multivariate dependency models is performed on estimates of value-at-risk (VaR) and conditional value-at-risk (CVaR). To illustrate the results, we use four financial share portfolios selected to exemplify this purpose. For each share, we calculate filtered log returns using autoregressive moving average–generalized autoregressive conditional heteroscedasticity models and study their dependence. We analyze how selecting pairs of assets to define vines prior to pair–copula decomposition affects the estimated VaR and CVaR. Further, using bootstrap confidence intervals, we compare the results of different risk measures obtained by employing alternative measures of dependence to select the order in which the drawable vine (D-vine) is defined in different portfolios. Moreover, we carry out a simulation study to analyze the finite sample properties of the different criteria for selecting the pair–copula decomposition associated with the D-vine. We find some differences between the results obtained for VaR and CVaR.

Keywords: value-at-risk (VaR); conditional value-at-risk (CVaR); pair–copula; dependence measures; drawable vine (D-vine).

Corresponding author: C. Bolancé

Print ISSN 1465-1211 | Online ISSN 1755-2842

1 INTRODUCTION

Value-at-risk (VaR) is the most popular measure proposed by the financial industry regulator to quantify risk. The estimation of VaR has posed and will continue to pose different challenges in the context of financial analyses. When losses have been generated by a set of dependent risk factors, we must take this dependence into account when estimating VaR. Further, we include conditional value-at-risk (CVaR) in our analysis.

We have two strategies for including dependency. We can either use a multivariate distribution to fit the marginal behavior of risk factors and their dependence together or employ univariate distributions to model each risk factor and use a copula to model dependency. However, when copulas are used, some restrictions exist if the number of dimensions or risk factors is more than two, ie, except for the Gaussian copula, the generalization from the bivariate copula to the multivariate copula assumes some restrictions with regard to the dependence between pairs of risk factors. For example, on the one hand, when using elliptical copulas to estimate the Student *t* copula, we need to assume the degrees of freedom are the same for all pairs. On the other hand, when using Archimedean copulas, it is assumed the value of the parameter of the copula is the same for all pairs (for an introduction to copulas, see, for example, Nelsen (2006)).

A pair–copula decomposition of a multivariate distribution is a flexible form of modeling the dependence relations between pairs of variables, considering different intensities of dependence or even dependence structures. However, such flexibility could lead to some inefficiency and bias in the risk estimation. In response to that, in this work we analyze the effect of selecting pairs, better known as "vines" (Bedford and Cooke 2002), on the risk estimation. In particular, we focus on the selection of the drawable vine (D-vine), although a similar analysis may be carried out on the canonical vine (C-vine) or the regular vine (R-vine), the latter being the more general representation.

The subject of vine copulas is growing more popular in the literature. In fact, there exists a wide variety of financial applications, related not only to the estimation of VaR and CVaR but also to other contexts, such as portfolio optimization and trading strategies (see, for example, Brechmann and Czado 2013; Low *et al* 2013, 2016; Nikoloulopoulos *et al* 2012; Rad *et al* 2016; Weiß and Supper 2013). In this respect, statistical analyses such as the one presented in this paper are fundamental.

Our contribution contains three main results. First, we show how different pair–copula decompositions provide different VaR and CVaR values, and we analyze the magnitude of these differences. The second result is related to the selection of the optimal decomposition, where different criteria, depending on whether we assume the copula is known or not, are analyzed. Finally, we define a specific algorithm for selecting the optimal D-vine; this algorithm is programmed in R and is available from the authors.

As indicated in the previous paragraph, our statistical aim is to analyze the effect of using different D-vines to estimate VaR and CVaR. In relation to this, having greater or lesser dependence between variables is a fundamental characteristic. To exemplify our results, we model the losses obtained from filtered returns associated with four stock portfolios, whose differences lie in their degrees of diversification. We design the four portfolios so that greater diversification implies less dependence, while less diversification implies greater dependence between filtered returns. Further, to have a sufficient number of different D-vines, the portfolios are composed of six stocks. In order to complete our analysis and compare results with different dimensions, we also include a simulation study that allows us to compare the statistical properties of different criteria for selecting optimal decomposition.

In this context, we use the D-vine for pair–copula decomposition because we have no information justifying any hierarchical relationships between stocks. Our study is similar to those of Aas *et al* (2006) and Min and Czado (2010), who analyze the D-vine pair–copula construction applied to filtered financial returns.

When we use the D-vine, the pair–copula decomposition relates to what we decide is the most suitable risk-factor order at the beginning of the process. We can use different criteria for sorting variables. For instance, for the Student t copula, Aas *et al* (2006) estimate the degrees of freedom for each pair of variables and select the order of variables, starting with the pairs that have lower estimated degrees of freedom and ending with those that have higher estimated degrees of freedom, ie, ordering from more to less lower and upper tail dependence. Dissmann *et al* (2013), based on Kendall's tau, use a method for selecting R-vine that they name a "maximum spanning tree" algorithm; typically, this algorithm is described as a minimum spanning tree (Prim's algorithm). Righi and Ceretta (2013, 2015) also use a Kendall's tau dependence matrix to determine the order of risk factors to design the D-vine. Common to most applications is the flexibility with which this order can be chosen, due to the absence of a specific decision rule; for this reason, it is necessary to analyze the effect of selecting the order of stocks to define the D-vine with the aim of estimating the risk of loss.

So, for an analyzed portfolio we conduct two types of analysis. First, we estimate the VaR for all possible orders and analyze its dispersion; with this aim in mind, we use portfolios in our numerical example. Second, we determine the VaR obtained using different criteria to select the order in the D-vine. In general, for a given copula, Kendall's tau, Spearman's rho and, if it exists, tail dependence can be used. These criteria are evaluated supposing that the copula is either known or unknown. We use the Monte Carlo method to estimate every VaR and CVaR (see McNeil *et al* 2015, Chapter 2).

Nowadays, the use of pair–copula decomposition is a topic that is gaining increasingly more followers in different lines of research, but especially in the financial industry. Recently, Weiß and Scheffer (2015) proposed the use of different copulas for different pairs in pair–copula decomposition in order to forecast the VaR of financial portfolios. Min and Czado (2014) analyzed the properties of the maximum likelihood estimator of the copula parameters associated with each pair, using pseudo-data. This estimator is called the pseudo-maximum likelihood estimator, or, as the authors say, the semiparametric maximum likelihood estimator, given that the empirical distribution multiplied by $T/(T+1)$ is used for marginals, where T is the sample size. They show an application to daily log returns of foreign exchange rates (see Bolancé *et al* (2014) for a comparison of semiparametric maximum likelihood estimators using different nonparametric methods for marginal distributions).

To estimate multivariate distribution using pair–copula decomposition, we must have independent and identically distributed (iid) observations. Then, similarly to the papers cited in the preceding paragraphs, we use ARMA(P, Q)–GARCH(p, q) serial models to filter our data, ie, we use the residuals of the fitted serial models to estimate the parameter in the pair–copula decomposition approach.[1]

2 PORTFOLIO RISK QUANTIFICATION

Bedford and Cooke (2002) define the concept of vine and R-vine. Specifically, we can see that an R-vine is a sequence of trees that represents the dependence structure of a multivariate vector of continuous random variables, given a factorization of the multivariate density function. The D-vine is a particular case of the R-vine. In Appendix 1 (available online), we describe the particular case of a D-vine with six variables, which we will analyze in the empirical part of this paper. In practice, the order in T_1 of our D-vine (see Appendix 1, available online, for more details) affects the estimation results and therefore the estimated risk (see Min and Czado (2010) for further information). So, we analyze how, by using a different order, we can obtain different estimates of VaR and CVaR.

To analyze the dispersion of the estimated VaRs and CVaRs, ie, the precision of estimated risk, specifically for the case of six dimensions, we obtain $6!/2 = 360$ possible orders and their estimated VaRs and CVaRs using each resulting D-vine. This procedure, based on Monte Carlo simulation, is described later.

To quantify the portfolio risk, we need to calculate a measure of returns for every asset. Hence, we are interested in using a measure that allows us to reflect the relative changes in the prices. For each share j, we define the log return variable at time t, ie, $R_{jt} = \log(P_{jt}/P_{jt-1}), t = 1, \dots, T$, where T is the total of the observed time points, and P_{jt} is the price of the asset j at time t. We are interested in modeling dependence

[1] ARMA: autoregressive moving average. GARCH: generalized autoregressive conditional heteroscedasticity.

between random time-independent shifts in the returns. For this reason, before starting our pair–copula analysis we filtered the stock returns in the following way. We assume that returns R_{jt} are generated by a time series model ARMA(P, Q)–GARCH(p, q) that can be expressed as

$$\begin{aligned} R_{jt} &= \mu_j + \sum_{i=1}^{P} \phi_{ji} R_{jt-i} - \sum_{i=1}^{Q} \psi_{ji} e_{jt-i} + e_{jt}, \\ e_{jt} &= \sigma_{jt} x_{jt}, \\ \sigma_{jt}^2 &= \alpha_{j0} + \sum_{i=1}^{p} \alpha_{ji} e_{jt-i}^2 + \sum_{i=1}^{q} \beta_{ji} \sigma_{jt-i}^2. \end{aligned} \tag{2.1}$$

The filtered log returns are x_{jt}, $t = 1, \dots, T$, and we can suppose that these are T values independent of and equally distributed from random variable X_j.

Let $(X_{(1)}, \dots, X_{(6)})$ be a multivariate vector of six random continuous variables that represent filtered log returns, where the parentheses indicate a given order; their marginal cumulative distribution functions (cdfs) are $F_{(1)}, \dots, F_{(6)}$, and the multivariate cdf is F. In our case, we are interested in estimating the VaR and CVaR, with confidence level α, of the random variable $L = -V_0(w_1 X_1 + \cdots + w_6 X_6)$ (see McNeil *et al* 2015, Chapter 2), where L is the linearized loss variable associated with a share portfolio, w_j is the weight of share j in the portfolio and V_0 is an initial investment value that we can suppose equals 1.

The VaR with confidence level α for a continuous random loss variable L can be defined as

$$\mathrm{VaR}_\alpha(L) = \inf\{l, F_L(l) \geqslant \alpha\} = F_L^{-1}(\alpha), \tag{2.2}$$

where L has a probability distribution function (pdf) of f_L and a cdf of F_L. Given a VaR, the CVaR is (see Denuit *et al* 2005)

$$\mathrm{CVaR}_\alpha(L) = E(L - \mathrm{VaR}_\alpha(L) \mid L > \mathrm{VaR}_\alpha(L)). \tag{2.3}$$

Let us assume that $(X_{(1)t}, \dots, X_{(6)t})$, $t = 1, \dots, T$, denotes a six-dimensional sample of T iid observations from random vector $(X_{(1)}, \dots, X_{(6)})$; for each possible order $o = 1, \dots, (6!/2) = 360$, we use a Monte Carlo procedure to estimate the $\mathrm{VaR}_\alpha^o(L)$ and the $\mathrm{CVaR}_\alpha^o(L)$, which is described below.

Step 1. We obtain the pseudo-data

$$U_{jt} = \frac{T}{T+1} \hat{F}_{jT}(X_{jt}), \quad j = 1, \dots, 6, \tag{2.4}$$

where $\hat{F}_{jT}(x) = (1/T) \sum_{t=1}^{T} I(X_{jt} \leqslant x)$ is the empirical distribution function, and $I(\cdot)$ is the indicator function that takes a value of 1 if the condition in parentheses is true and 0 otherwise.

Step 2. Using pseudo-data $(U_{(1)t}, \dots, U_{(6)t})$, $t = 1, \dots, T$, we calculate the pseudo-loglikelihood associated with the multivariate copula in order to estimate the dependence parameters that maximize this partial likelihood. The pseudo-loglikelihood (see Genest *et al* (1995) and Min and Czado (2014) for a review of estimation procedures based on pseudo-data) is

$$\ln(L(\theta)) = \sum_{t=1}^{T} \ln\{c_\theta(U_{(1)t}, \dots, U_{(6)t})\}, \tag{2.5}$$

where c_θ is the pair–copula density and θ is a vector with $6(6-1)/2 = 15$ copula parameters that depends on the order selected for the risk factors, ie,

$$\begin{aligned}\theta = (&\theta_{(1)(2)}, \theta_{(2)(3)}, \theta_{(3)(4)}, \theta_{(4)(5)}, \theta_{(5)(6)}, \theta_{(1)(3)|(2)},\\ &\theta_{(2)(4)|(3)}, \theta_{(3)(5)|(4)}, \theta_{(4)(6)|(5)}, \theta_{(1)(4)|(2)(3)}, \theta_{(2)(5)|(3)(4)},\\ &\theta_{(3)(6)|(4)(5)}, \theta_{(1)(5)|(2)(3)(4)}, \theta_{(2)(6)|(3)(4)(5)}, \theta_{(1)(6)|(2)(3)(4)(5)}).\end{aligned} \tag{2.6}$$

To evaluate the goodness-of-fit of the estimated pair–copula decomposition, given the bivariate copula used for all pairs, we use the statistics based on the probability integral transform (PIT) method, proposed and described in detail by Aas *et al* (2006).

Step 3. We simulate the vectors $(\tilde{U}_{(1)s}, \dots, \tilde{U}_{(6)s})$, $s = 1, \dots, S$, from the estimated copula, where S is the number of simulated six-dimensional vectors. We use the CDVine package of R. The simulation method implemented in this R package is described by Brechmann and Schepsmeier (2013).

Step 4. We calculate the simulated risk factors as

$$\tilde{X}_{(1)s} = \hat{F}_{(1)}^{-1}(\tilde{U}_{(1)s}), \dots, \tilde{X}_{(6)s} = \hat{F}_{(6)}^{-1}(\tilde{U}_{(6)s}), \quad s = 1, \dots, S,$$

where $\hat{F}_{(j)}^{-1}$ denotes the inverse of the estimated marginal cdf of the random variable $X_{(j)}$. For marginal cdfs, we estimated $j = 1, \dots, 6$ univariate normal distributions with parameters μ_j^{Normal} and σ_j^{Normal}, or six univariate Student t distributions with υ degrees of freedom and parameters μ_j^{Student} and $\sigma_j^{\text{Student}}$. In all cases, to estimate $\hat{F}_{(j)}$ we use the maximum likelihood method. Recent studies by Christoffersen *et al* (2013) and Oh and Patton (2013) show that the distribution of equity returns exhibits skewness and fat tails (see Fernández and Steel 1998). In this way, we perform different tests for asymmetry of the marginal distributions of the filtered returns (see Boos 1982) and, as we will indicate in Section 3, we do not reject the null hypothesis of symmetry.

Step 5. Finally, we simulate the linearized losses as

$$\tilde{L}_s = -V_0(w_{(1)}\tilde{X}_{(1)s} + \cdots + w_{(6)}\tilde{X}_{(6)s}), \quad s = 1, \ldots, S,$$

and we estimate $\mathrm{VaR}_\alpha(L)$ empirically once a large number S of simulated data is available. In our numerical example (Section 3), where the aim is to analyze the sensitivity of VaR and CVaR to D-vine selection, the calculation of weights is irrelevant. So, for simplicity, we assume $w_{(j)} = 1/6$ for all $j = 1, \ldots, 6$.

The empirical estimation of the VaR for order o is

$$\widehat{\mathrm{VaR}}^o_\alpha(L)_S = \inf\{l, \hat{F}^o_{L_S}(l) \geqslant \alpha\}, \tag{2.7}$$

where $\hat{F}^o_{L_S}$ is the empirical estimation of F_L, given the order o, obtained from the S simulated losses. The empirical estimation of CVaR ($\widehat{\mathrm{CVaR}}^o_\alpha(L)_S$) is the mean of differences $\tilde{L}^o_s - \widehat{\mathrm{VaR}}^o_\alpha(L)_S$ for all $\tilde{L}^o_s > \widehat{\mathrm{VaR}}^o_\alpha(L)_S$, where $\tilde{L}^o_s$ are the estimated losses using the D-vine associated with order o.

Once we have estimated $\widehat{\mathrm{VaR}}^o_\alpha(L)_S$ and $\widehat{\mathrm{CVaR}}^o_\alpha(L)_S$ for $o = 1, \ldots, (6!/2) = 360$, in order to analyze the accuracy of our risk estimation, we calculate some dispersion measures based on central moments, which can be the standard deviation or the coefficient of variation. Alternatively, we can also calculate the range (maximum minus minimum) or the interquartile range (quartile 3 minus quartile 1). At this point, it is important to note that using the Monte Carlo procedure causes some dispersion associated with the random process itself. To control this spurious dispersion, we use the same initial seed in each random generation for each pair–copula decomposition.

2.1 Order-selection criteria

An easy way to select the optimal order of risk factors in the D-vine consists of maximizing the pseudo-loglikelihood that was defined in (2.5); however, this criterion has some drawbacks. First, it requires fitting the copula parameters for all possible orders to search for those that maximize pseudo-loglikelihood. Second, maximizing pseudo-loglikelihood does not have to be related to obtaining a better structure of dependence for estimating risk; in fact, pseudo-loglikelihood obtained with different orders allows us to estimate differently conditioned models that give more importance to observations with higher density, which contradicts the fact that the risk is associated with the least likely observations. For this reason, it is necessary to search for different criteria that allow us to select the order before estimating the multivariate copula (see Vaz de Melo *et al* (2010) for a financial example in six dimensions).

Given that we are interested in estimating dependence structure, the most natural criteria for selection order are based on the dependence measures related to copulas, ie, Kendall's tau, Spearman's rho and lower (left) or upper (right) tail dependence (λ_L and

λ_{U}, respectively). These dependence measures can be defined according to the copula or empirically. For this reason, we can say that selecting the order in D-vine may be carried out with or without copula information. To estimate λ_{L} and λ_{U} empirically, we use the nonparametric estimation proposed by Schmidt and Stadtmüller (2006).

In order to select the pairs, we define an algorithm that consists of sorting the edges in the first tree (see Figure 1) from maximum to minimum dependence, considering D-vine structure. The procedure is defined below.

2.1.1 Pair-selection procedure

To select the D-vine, we can use an algorithm based on finding the shortest or longest paths between nodes in a graph (see, for example, Dissmann *et al* 2013). However, in our context, these algorithms consist of maximizing the sum of the dependences between the pairs of returns without considering that we need to select a specific order for defining the D-vine. Our proposed procedure is based on sorting the pairs of returns from higher to lower dependence. In every step, we select a pair conditioned by the pair selected in the previous step.

Let D be a matrix of dependences that is symmetric and positive definite, and let $d_{ij} = d_{ji}$, $i, j = 1, \dots, k$, be the dependence between returns i and j, where $d_{ij} = 1$ if $i = j$; then, if there are no ties between dependences, the following hold.

Step 1. The first pair $[(1), (2)]$ is that for which (the parentheses indicate order)

$$d_{(1)(2)} = \max_{i \neq j}(d_{ij});$$

then, the first pair can be $[(1), (2)] = [i^*, j^*]$ or $[(1), (2)] = [j^*, i^*]$. To select between these pairs, it is necessary to analyze the possible second pairs $[(2), (3)]$. So,

$$d_{i^*(3)} = \max_{j \neq j^*}(d_{i^* j}) \quad \text{and} \quad d_{j^*(3)} = \max_{i \neq i^*}(d_{i j^*});$$

then,

$$\begin{aligned} &\text{if } d_{i^*(3)} > d_{j^*(3)}, \text{ the first pair is } [(1), (2)] = [j^*, i^*], \\ &\text{if } d_{i^*(3)} < d_{j^*(3)}, \text{ the first pair is } [(1), (2)] = [i^*, j^*], \end{aligned}$$

and we eliminate row i^* and column j^* of matrix D. We denote by $D^{[-1]}$ the matrix of dependence with $k-1$ columns and rows. In general, $D^{[-s]}$ is the matrix of dependence with $k-s$ columns and rows.

Step 2. The following pairs $[(s+1)(s+2)]$ are selected using the criterion

$$d^{-s}_{(s+1)(s+2)} = \max_{h \neq (s)}(d_{(s+1)h}) = d_{(s+1)h^*}, \tag{2.8}$$

where $(s+2) = h^*$.

If there are ties between dependences, we apply step 1 for each pair with the same dependence and select the initial pairs that provide the higher $d_{i^*(3)}$ or $d_{j^*(3)}$. If there are ties between dependences in step 2, similarly to step 1, we apply the criterion defined in (2.8) for each pair with the same dependence until we reach a tiebreaker. This procedure is implemented in R by the authors.

2.2 The analyzed copulas

We compare the results obtained with Student t, Gumbel, Clayton and Frank bivariate copulas for the pair decomposition. For each multivariate model, we use the same copula for all pairs, although, as Weiß and Scheffer (2015) suggest, the copula with the best fit for each pair could be used. With these four copulas, a wide range of structures of dependence alternatives to the Gaussian benchmark model are taken into account.

The bivariate Student t copula is an implicit and elliptical copula belonging to the family of extreme value copulas (see Bahraoui *et al* (2014) for a revision of the family of extreme value copulas and its inference). Its functional form is equal to the standard Student t bivariate cdf, with υ degrees of freedom and a correlation coefficient ρ. This copula represents symmetric dependence structures, having heavier tails than those of the Gaussian copula. In addition, the Gaussian copula does not present tail dependence, while the Student t copula has both lower and upper tail dependence. For a given pair, the bivariate Student t copula is

$$\begin{aligned} C_{\rho_{12},\upsilon_{12}}(u_1,u_2) &= \int_{-\infty}^{t^{-1}_{\upsilon_{12}}(u_1)} \int_{-\infty}^{t^{-1}_{\upsilon_{12}}(u_2)} \frac{1}{2\pi\sqrt{1-\rho_{12}^2}} \\ &\quad \times \left(1 + \frac{\begin{array}{c} t^{-1}_{\upsilon_{12}}(s_1)^2 + t^{-1}_{\upsilon_{12}}(s_2)^2 \\ -2\rho_{12} t^{-1}_{\upsilon_{12}}(s_1) t^{-1}_{\upsilon_{12}}(s_2) \end{array}}{\upsilon_{12}(1-\rho_{12}^2)}\right)^{-(\upsilon_{12}+2)/2} \mathrm{d}t^{-1}_{\upsilon_{12}}(s_1)\,\mathrm{d}t^{-1}_{\upsilon_{12}}(s_2), \end{aligned} \tag{2.9}$$

where t_{υ}^{-1} is the quantile function of the univariate Student t.

From the explicit and Archimedean family of copulas, we use the most popular: the Gumbel (1960), Clayton (1978) and Frank (1979) copulas. This class of bivariate copulas has a simple closed form, and its structure depends only on the dependence parameter.

The Gumbel is an extreme value copula with functional form given by

$$C_{\theta_{12}}(u_1,u_2) = \exp(-[(-\ln(u_1))^{\theta_{12}} + (-\ln(u_2))^{\theta_{12}}]^{1/\theta_{12}}), \tag{2.10}$$

where $\theta_{12} \in [1, +\infty)$ is the parameter controlling the dependence structure. Finally, the dependence is perfect when $\theta_{12} \to \infty$, and we have independence when $\theta_{12} = 1$.

For the Gumbel copula, it is well known that $\lambda_{\mathrm{L}} = 0$ and $\lambda_{\mathrm{U}} = 2 - 2^{1/\theta_{12}}$, ie, the Gumbel copula has upper tail dependence.

The Clayton copula has a functional form equal to

$$C_{\theta_{12}}(u_1, u_2) = (u_1^{-\theta_{12}} + u_2^{-\theta_{12}} - 1)^{-1/\theta_{12}}, \tag{2.11}$$

where $\theta_{12} > 0$. In this case, the perfect dependence structure is achieved when $\theta_{12} \to \infty$, and independence is achieved when $\theta_{12} \to 0$. In contrast to the Gumbel copula, the Clayton copula has lower tail dependence, in this case $\lambda_{\mathrm{L}} = 2^{-1/\theta_{12}}$ and $\lambda_{\mathrm{U}} = 0$.

The Frank copula is defined by the parameter $\theta_{12} \in (-\infty, 0[\cup]0, +\infty)$ and is given by

$$C_{\theta_{12}}(u_1, u_2) = -\frac{1}{\theta_{12}} \ln\left(1 - \frac{(1 - \mathrm{e}^{\theta_{12} u_1})(1 - \mathrm{e}^{\theta_{12} u_2})}{1 - \mathrm{e}^{-\theta_{12}}}\right).$$

The upper and lower tail dependence coefficients of the Frank copula are 0. This copula is characterized by its higher dependence in the central quantiles.

3 RESULT OF EMPIRICAL ANALYSIS USING FINANCIAL DATA

To study the effect on the risk quantification of the selection order of risk factors in D-vine, we analyze four portfolios (P1, P2, P3 and P4). We select these portfolios, changing the companies that form every one (see Table 10 in the online appendix to see the companies that form each portfolio). The most diversified portfolios (P1 and P2) are composed of shares of companies belonging to at least five different economic sectors. Alternatively, the less diversified portfolios (P3 and P4) are composed of shares from companies belonging to the same economic sectors or, at most, two different sectors. In our example, we use only stocks, but the results can be extrapolated directly to portfolios composed using other assets. The different log return diversifications of the portfolios are evaluated using different empirical dependence matrixes of filtered log returns, eg, Kendall's tau and Spearman's rho (see Table 11 in the online appendix to see the estimated ARMA(P, Q)–GARCH(p, q) models). The log returns are calculated using the daily prices from January 2011 to December 2013 obtained from Yahoo Finance. All prices are checked and expressed in US dollars.

In Figures 1, 2, 3 and 4, we show the scatter plots of filtered returns in portfolios P1, P2, P3 and P4, respectively. From Figures 2 and 4, we eliminate the three plots that have already been included in Figures 1 and 3. In general, we observe an elliptical shape that reflects the dependence between filtered returns. This elliptical shape is more pronounced in Figures 3 and 4, which correspond to less diversified portfolios. In general, we also observe fairly symmetric shapes and some extreme points that reflect the tail dependence between filtered returns.

In Table 1, we show the empirical Kendall's tau and Spearman's rho for each portfolio; later, in Table 2 we present empirical upper and lower tail dependences (R1, ..., R6 refer to filtered log returns). We also calculate the determinants (in parentheses) of each dependence matrix: the smaller this determinant, the greater the dependence between returns.

We fit pair–copula decomposition using Student t, Gumbel, Clayton and Frank copulas, obtaining the best fit with the Student t copula. The Clayton copula does not pass the PIT goodness-of-fit test at a 5% significance level for any equity portfolio and any of the 360 possible D-vines, so we eliminate the results for this copula. For the rest of the copulas, the PIT test is passed for all the D-vines in the four portfolios, although the significance level is the highest for the Student t copula.

For the univariate marginal, using filtered returns associated with each share, we tested the hypothesis of symmetry using a statistic based on the Hodges–Lehmann estimator (see Boos 1982). In all cases, we cannot reject the null hypothesis of symmetry. Given these results, we analyze the fit of normal and Student t distributions, and it turns out that the Student t presents the best fit for all filtered log returns.

3.1 Dispersion analysis

To analyze the shape of the distribution of the estimated VaR and CVaR using the 360 different initial orders in the tree T_1 of the D-vine, which represent 360 different multivariate dependence models, we use kernel density estimation with rule-of-thumb bandwidth and the Gaussian kernel (see Silverman 1986). This is a nonparametric method that allows us to easily smooth the shape of our histogram. In Figures 5, 6 and 7, we plot the densities obtained using the Student t, Frank and Gumbel copulas, respectively, with Student t marginal distributions. The Student t and Frank copulas assume symmetrical dependence structures between filtered log returns two by two. Alternatively, the Gumbel copula supposes asymmetry in the dependence structures and, in our example, it provides greater risk estimates.

The differences between the copulas can also be observed in Tables 3 and 4, where we include the mean (Mean), the standard deviation (STD), the coefficient of variation (CV), the interquartile range (IQR) and the range (Range) of the estimated VaR and CVaR at the 99% and 99.5% confidence levels, using the 360 different initial orders. When analyzing dispersion measures, we observe that, with the Gumbel copula, the results differ from those obtained with the Student t and Frank copulas. The highest differences between estimated VaR and CVaR are obtained with the former, followed by the Student t and finally the Frank copula.

FIGURE 1 Scatter plots of filtered returns in P1.

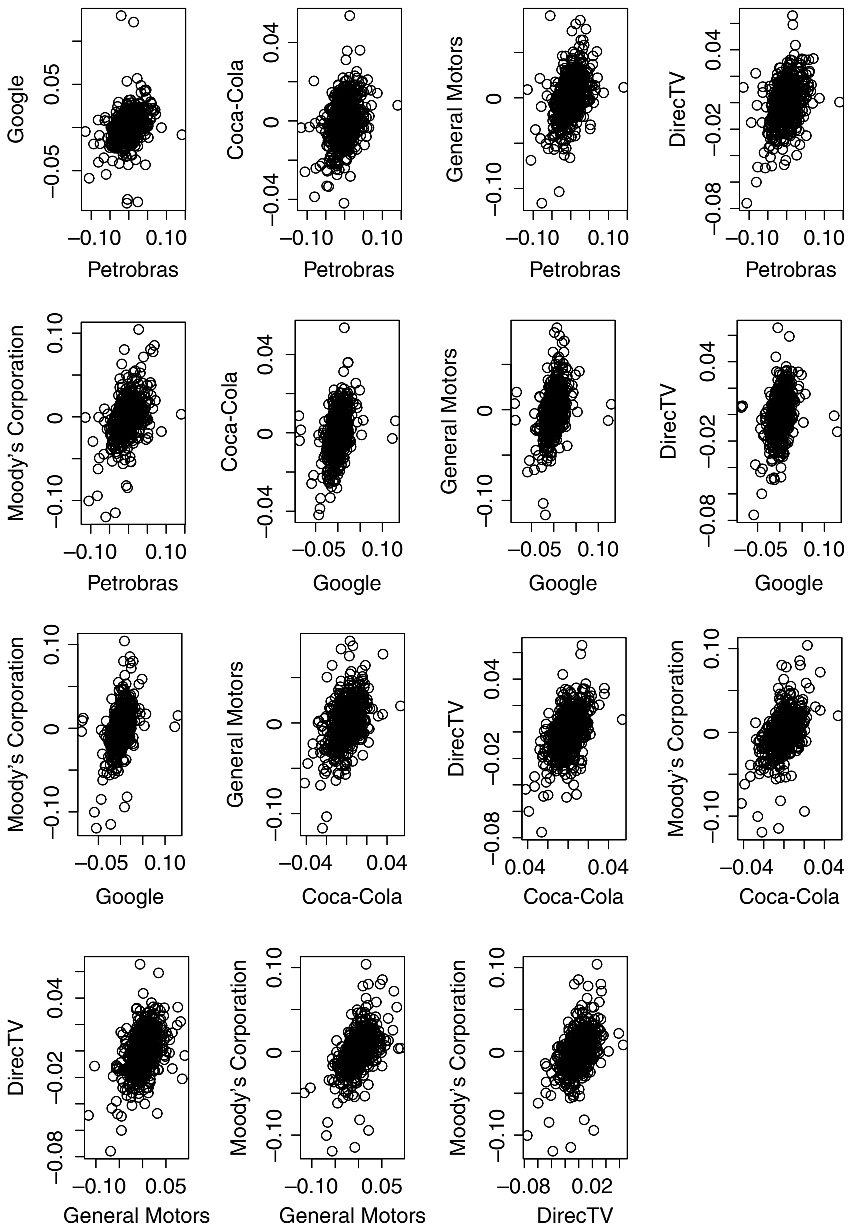


FIGURE 2 Scatter plots of filtered returns in P2 that are not included in P1.

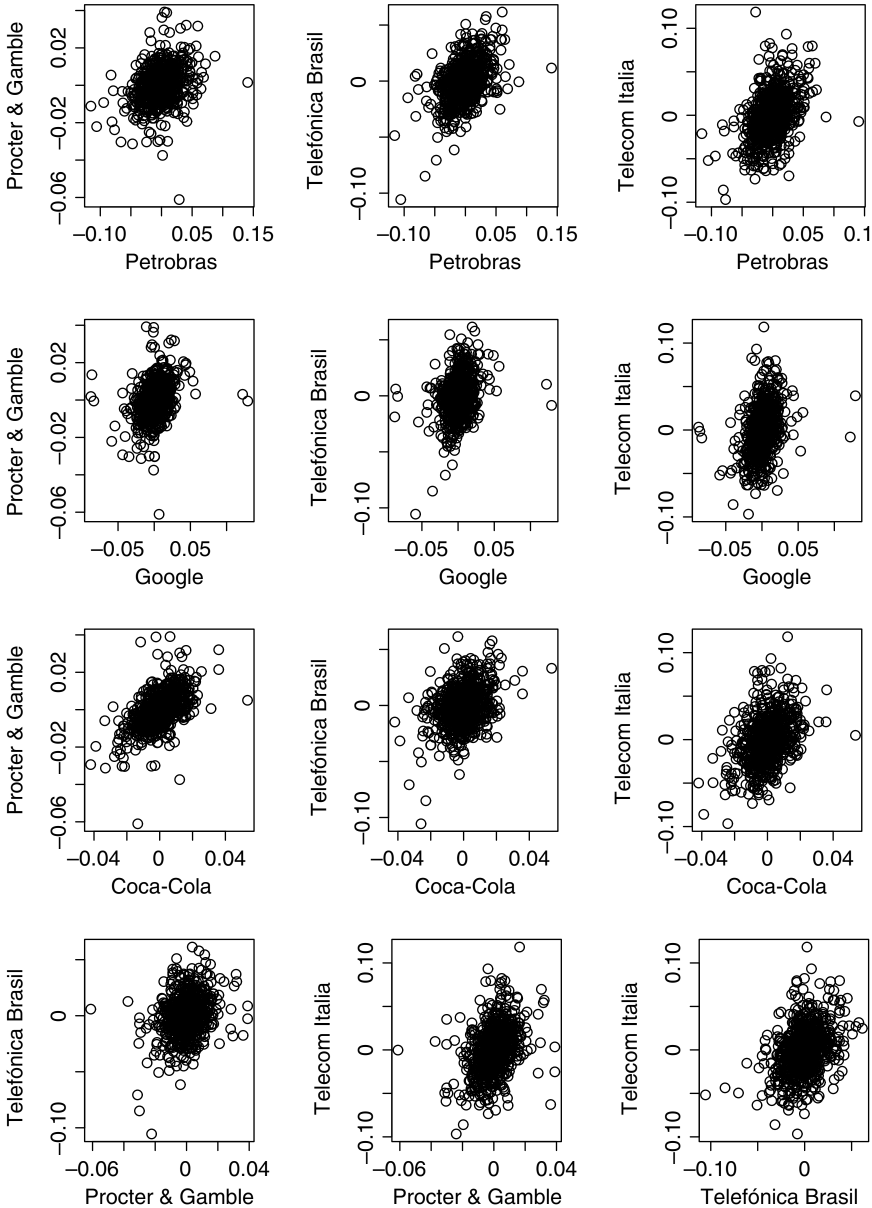


FIGURE 3 Scatter plots of filtered returns in P3.

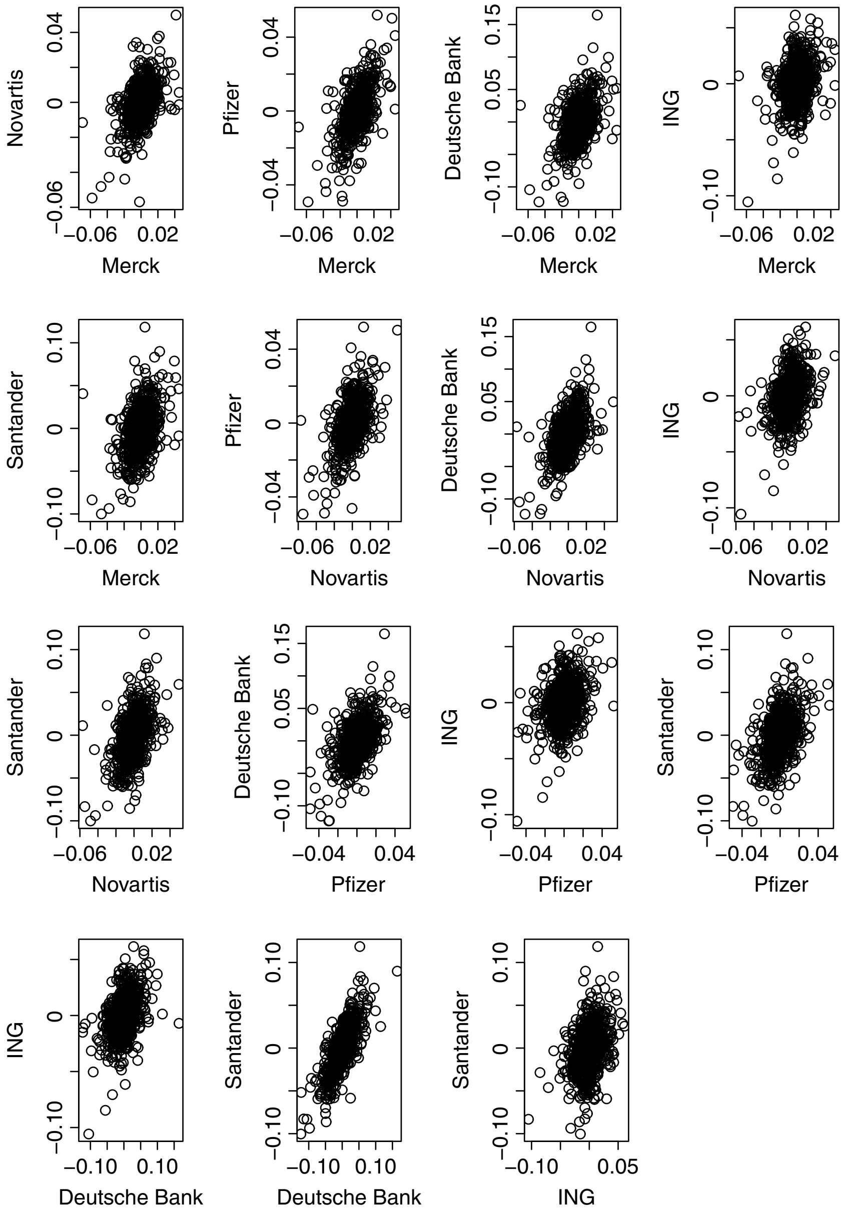


FIGURE 4 Scatter plots of filtered returns in P4 that are not included in P3.

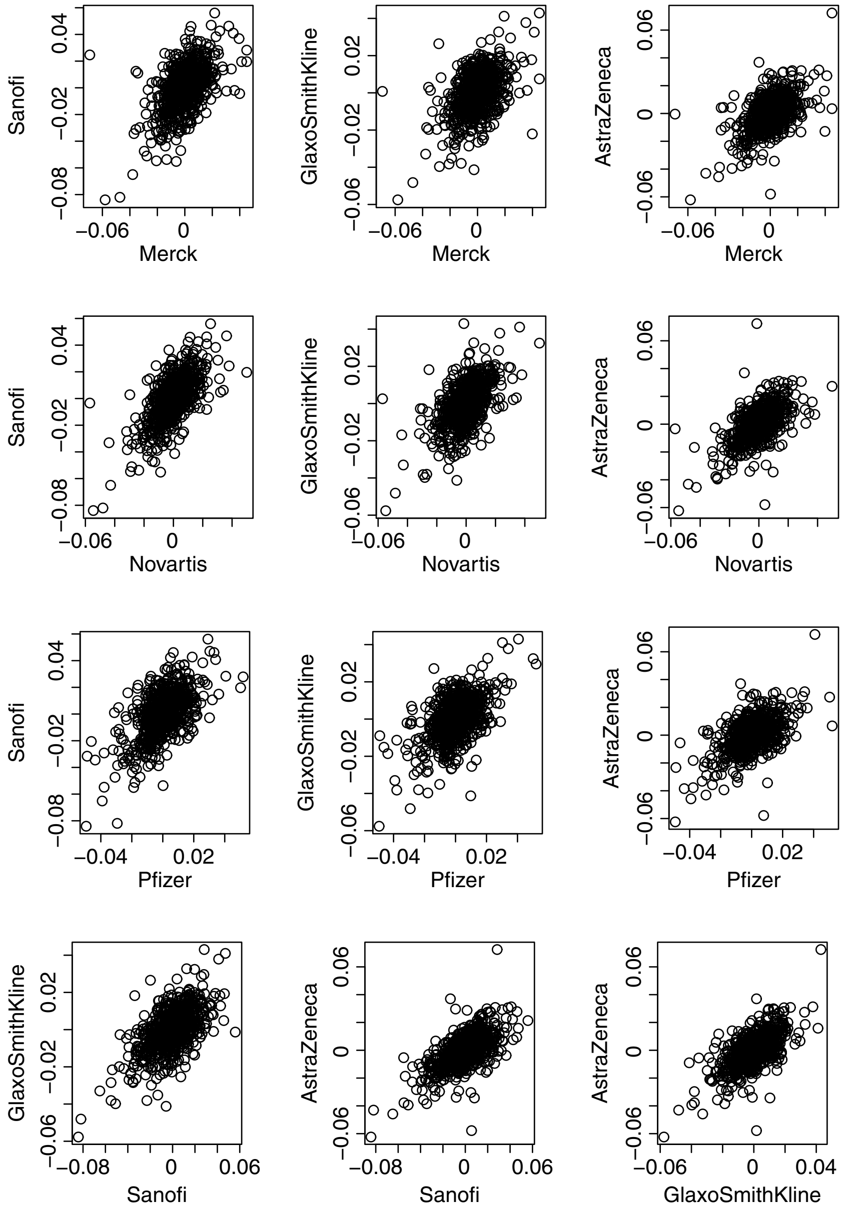


TABLE 1 Dependence measures: empirical Spearman's rho (above diagonal) and empirical Kendall's tau (below diagonal).

	P1: τ (0.43), ρ (0.19)							**P2: τ (0.50), ρ (0.25)**					
	R1	**R2**	**R3**	**R4**	**R5**	**R6**		**R1**	**R2**	**R3**	**R4**	**R5**	**R6**
R1	1.000	0.371	0.315	0.404	0.373	0.427	**R1**	1.000	0.371	0.315	0.276	0.427	0.421
R2	0.256	1.000	0.396	0.465	0.392	0.455	**R2**	0.256	1.000	0.396	0.392	0.264	0.377
R3	0.215	0.274	1.000	0.377	0.433	0.462	**R3**	0.215	0.274	1.000	0.548	0.292	0.374
R4	0.278	0.323	0.261	1.000	0.390	0.538	**R4**	0.189	0.270	0.389	1.000	0.250	0.374
R5	0.261	0.270	0.302	0.270	1.000	0.446	**R5**	0.295	0.178	0.201	0.171	1.000	0.339
R6	0.298	0.318	0.322	0.381	0.311	1.000	**R6**	0.292	0.261	0.258	0.259	0.231	1.000
	P3: τ (0.33), ρ (0.11)							**P4: τ (0.25), ρ (0.071)**					
	R1	**R2**	**R3**	**R4**	**R5**	**R6**		**R1**	**R2**	**R3**	**R4**	**R5**	**R6**
R1	1.000	0.449	0.573	0.411	0.239	0.388	**R1**	1.000	0.449	0.573	0.465	0.399	0.480
R2	0.312	1.000	0.482	0.477	0.321	0.469	**R2**	0.312	1.000	0.482	0.630	0.557	0.583
R3	0.418	0.338	1.000	0.460	0.287	0.407	**R3**	0.418	0.338	1.000	0.521	0.441	0.489
R4	0.289	0.335	0.324	1.000	0.388	0.766	**R4**	0.327	0.456	0.373	1.000	0.554	0.622
R5	0.162	0.217	0.196	0.265	1.000	0.333	**R5**	0.279	0.395	0.308	0.395	1.000	0.630
R6	0.270	0.329	0.282	0.585	0.226	1.000	**R6**	0.338	0.421	0.346	0.450	0.458	1.000

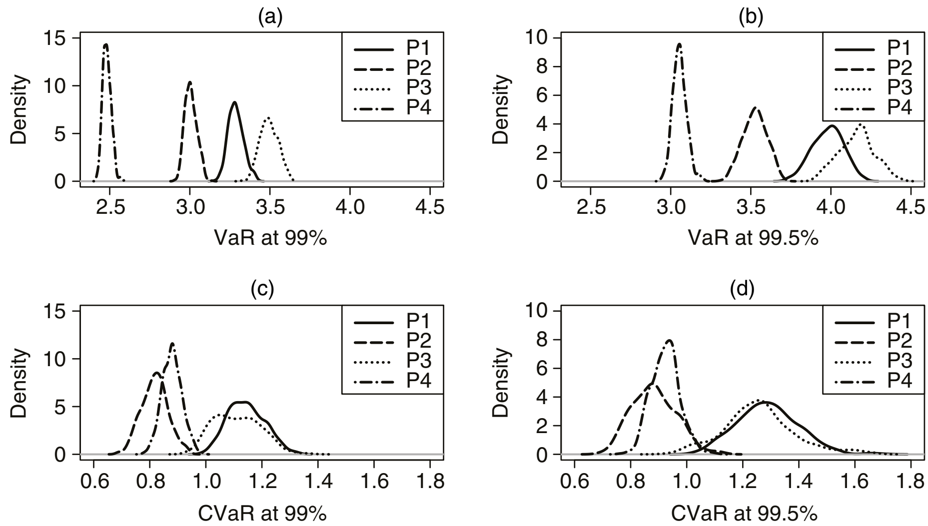
P1, ..., P4 indicate portfolio and R1, ..., R6 indicate filtered log returns.

TABLE 2 Dependence measures: empirical λ_L (above diagonal) and empirical λ_U (below diagonal).

	P1: λ_U (0.66), λ_L (0.34)							P2: λ_U (0.70), λ_L (0.41)					
	R1	**R2**	**R3**	**R4**	**R5**	**R6**		**R1**	**R2**	**R3**	**R4**	**R5**	**R6**
R1	1.000	0.222	0.333	0.370	0.333	0.370	**R1**	1.000	0.222	0.333	0.222	0.296	0.222
R2	0.222	1.000	0.333	0.296	0.259	0.370	**R2**	0.222	1.000	0.333	0.222	0.148	0.185
R3	0.111	0.185	1.000	0.333	0.370	0.407	**R3**	0.111	0.185	1.000	0.481	0.296	0.333
R4	0.074	0.148	0.148	1.000	0.296	0.333	**R4**	0.148	0.222	0.222	1.000	0.296	0.333
R5	0.148	0.185	0.333	0.148	1.000	0.296	**R5**	0.259	0.185	0.185	0.074	1.000	0.222
R6	0.222	0.185	0.222	0.185	0.222	1.000	**R6**	0.185	0.111	0.111	0.148	0.148	1.000
	P3: λ_U (0.41), λ_L (0.34)							P4: λ_U (0.36), λ_L (0.21)					
	R1	**R2**	**R3**	**R4**	**R5**	**R6**		**R1**	**R2**	**R3**	**R4**	**R5**	**R6**
R1	1.000	0.370	0.333	0.444	0.259	0.259	**R1**	1.000	0.370	0.333	0.370	0.259	0.407
R2	0.222	1.000	0.444	0.444	0.185	0.333	**R2**	0.222	1.000	0.444	0.444	0.333	0.407
R3	0.407	0.296	1.000	0.370	0.185	0.259	**R3**	0.407	0.296	1.000	0.370	0.296	0.481
R4	0.370	0.185	0.296	1.000	0.259	0.370	**R4**	0.296	0.296	0.259	1.000	0.519	0.481
R5	0.222	0.148	0.185	0.259	1.000	0.148	**R5**	0.222	0.370	0.444	0.333	1.000	0.370
R6	0.259	0.259	0.259	0.481	0.148	1.000	**R6**	0.296	0.259	0.222	0.333	0.333	1.000

P1, . . . , P4 indicate portfolio and R1, . . . , R6 indicate filtered log returns.

FIGURE 5 Kernel density estimation of VaR (plots (a) and (b)) and CVaR (plots (c) and (d)), estimated with all possible D-vines using the Student t pair–copula.



The densities of VaR plotted in Figures 5 and 6 (plots (a) and (b)), corresponding to Student t and Frank copulas, have only a principal mode, although with the Frank copula all of the distributions have more kurtosis (compare the vertical axes), smaller dispersion and smaller centers than with the Student t copula (see also Table 3). Both copulas sort the portfolios from higher to lower risk in the same way: P3, P1, P2 and P4, respectively, using both analyzed risk measures, VaR at the 99% and 99.5% confidence levels. The densities of CVaR (plots (c) and (d)) show a lot of similarities between P1 and P3 and a reduction in dissimilarities between P2 and P4.

The densities shown in Figure 7, which were obtained using the Gumbel copula in the pair decomposition of the multivariate distribution, have a bimodal shape, more dispersion and larger centers than those obtained using the Student t copula. In general, the estimated risks using the Gumbel copula are much higher, and the differences between these estimated risks using different initial orders are also much higher. Again, the positions of P1 and P3 are similar and the same occurs with P2 and P4.

In general, we observe that the distributions of estimated VaR and CVaR with different D-vines depend on the selected copula and can have different shapes. Because of this, it is fundamental to analyze the results provided by different statistical criteria when selecting D-vine.

FIGURE 6 Kernel density estimation of VaR (plots (a) and (b)) and CVaR (plots (c) and (d)), estimated with all possible D-vines using the Frank pair–copula.

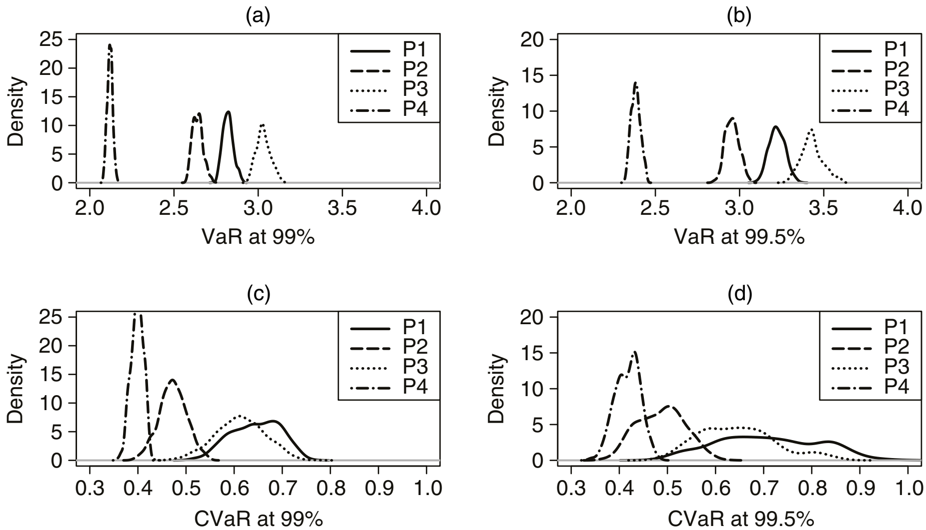


FIGURE 7 Kernel density estimation of VaR (plots (a) and (b)) and CVaR (plots (c) and (d)), estimated with all possible D-vines using the Gumbel pair–copula.

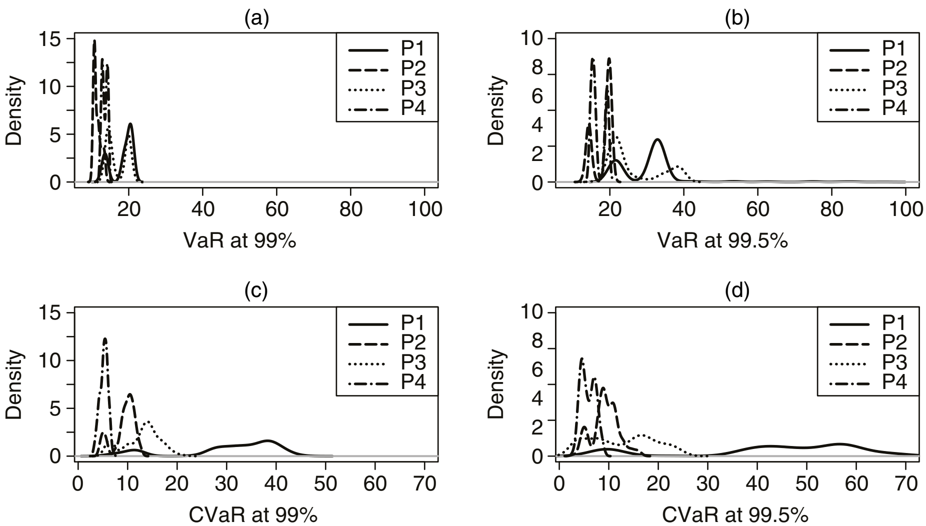


TABLE 3 Descriptive statistics of estimated VaR for all orders; the VaR is multiplied by 100.

	P1						P2					
	Student t		Gumbel		Frank		Student t		Gumbel		Frank	
	99%	99.5%	99%	99.5%	99%	99.5%	99%	99.5%	99%	99.5%	99%	99.5%
Mean	3.287	3.981	18.286	30.514	2.820	3.223	3.006	3.526	11.449	17.975	2.641	2.956
STD	0.048	0.098	2.986	9.015	0.031	0.048	0.038	0.075	0.993	2.619	0.031	0.041
CV	0.015	0.025	0.163	0.295	0.011	0.015	0.013	0.021	0.087	0.146	0.012	0.014
IQR	0.066	0.140	6.256	11.228	0.041	0.068	0.055	0.105	1.388	5.349	0.043	0.058
Range	0.277	0.510	9.487	73.181	0.177	0.281	0.232	0.412	4.542	8.302	0.173	0.229
	P3						P4					
	Student t		Gumbel		Frank		Student t		Gumbel		Frank	
	99%	99.5%	99%	99.5%	99%	99.5%	99%	99.5%	99%	99.5%	99%	99.5%
Mean	3.495	4.160	16.933	26.408	3.031	3.430	2.482	3.056	13.512	16.862	2.119	2.383
STD	0.056	0.111	2.733	7.177	0.042	0.064	0.027	0.043	0.742	1.879	0.016	0.027
CV	0.016	0.027	0.161	0.272	0.014	0.019	0.011	0.014	0.055	0.111	0.008	0.012
IQR	0.083	0.144	5.408	13.169	0.055	0.080	0.036	0.056	1.325	3.833	0.021	0.040
Range	0.308	0.561	8.854	23.810	0.214	0.338	0.173	0.269	2.815	5.475	0.092	0.136

TABLE 4 Descriptive statistics of estimated CVaR for all orders; the CVaR is multiplied by 100.

	P1						P2					
	Student t		Gumbel		Frank		Student t		Gumbel		Frank	
	99%	99.5%	99%	99.5%	99%	99.5%	99%	99.5%	99%	99.5%	99%	99.5%
Mean	1.144	1.289	30.757	43.502	0.645	0.709	0.818	0.883	9.352	9.324	0.473	0.487
STD	0.066	0.110	10.138	17.728	0.051	0.106	0.047	0.080	2.174	2.680	0.029	0.049
CV	0.058	0.085	0.330	0.408	0.079	0.149	0.058	0.091	0.233	0.287	0.060	0.100
IQR	0.096	0.148	10.432	18.585	0.079	0.168	0.063	0.114	2.181	2.995	0.038	0.074
Range	0.342	0.687	39.541	68.379	0.249	0.519	0.255	0.452	8.877	13.996	0.157	0.264
	P3						P4					
	Student t		Gumbel		Frank		Student t		Gumbel		Frank	
	99%	99.5%	99%	99.5%	99%	99.5%	99%	99.5%	99%	99.5%	99%	99.5%
Mean	1.111	1.258	12.799	12.789	0.613	0.645	0.879	0.924	5.189	5.815	0.398	0.418
STD	0.084	0.129	3.691	6.509	0.051	0.080	0.035	0.051	0.818	1.468	0.013	0.026
CV	0.075	0.103	0.288	0.509	0.083	0.124	0.040	0.055	0.158	0.252	0.033	0.062
IQR	0.128	0.144	4.808	10.690	0.071	0.112	0.049	0.065	1.173	2.492	0.018	0.038
Range	0.448	0.766	16.219	23.390	0.299	0.390	0.207	0.346	4.226	6.848	0.079	0.145

3.2 Selecting optimal order

In Section 3.1, we described alternative criteria for selecting the optimal order between the 360 possible orders in the six-dimensional D-vine. In this section, we show the results obtained with the better fit that, in our case, is the Student t copula with Student t marginal distributions.

First, in Table 5, we show the optimal order obtained using the different criteria estimated empirically (left-hand side) and supposing a Student t copula (right-hand side). In the latter case, we add the criterion based on degrees of freedom (changing maximum for minimum in the procedure defined in Section 2.1.1) as well as that based on the dependence structure, which maximizes the likelihood. Second, in Tables 6 and 7, we show the estimated VaRs and CVaRs at the 99% and 99.5% confidence levels, respectively, using the different orders quoted in Table 5. The results show how different criteria provide different orders. In general, however, there is not a lot of difference between the estimated VaR and CVaR. Although the orders differ, the results of estimated VaR and CVaR obtained with empirical criteria and Student t copula-based criteria are, at first glance, very similar.

In Tables 6 and 7, under the estimations of VaR and CVaR, and given each order, we include bootstrap confidence intervals (CIs) at the 90% confidence level (in parentheses), ie, 5% of estimated values are above the upper limit and 5% are below the lower limit. The CIs allow us to evaluate to what extent the differences between results obtained with different orders are statistically significant. To calculate bootstrap confidence intervals, we extract 500 samples with replacements from the data set associated with each portfolio. For each sample, we estimate the VaR and CVaR at the 99% and 99.5% confidence levels using the same selection order as in Table 5. Therefore, we do not consider the dispersion associated with the selection criteria; we evaluate this in the simulation study shown in Section 4.

In general, we observe that CIs overlap to a greater or lesser extent. There is only one case in which the CIs do not overlap: this corresponds to P1 when we compare the VaR obtained using the λ_{U} criterion, estimated empirically, with the VaR obtained from the D-vine that maximizes the likelihood. For all other cases, we can conclude that the values of risk estimated with different D-vines are not significantly different at the 5% level if the test is to one tail, or at the 10% level if it is to two tails.

4 SIMULATION STUDY

We now summarize the results of our simulation study. We compare the mean squared errors (MSEs) of VaR and CVaR obtained using the different criteria for selecting D-vine described in Section 2.1.1. We analyze whether significant differences exist between these criteria.

TABLE 5 Selection orders.

Portfolio	Empirical criterion				Student t copula-based criterion					
	ρ	τ	λ_U	λ_L	ρ	τ	λ_U	λ_L	DF	Likelihood
P1	4,6,3,5,2,1		5,3,6,1,2,4	6,3,5,1,4,2	6,4,2,3,5,1		1,5,3,6,4,2		1,5,3,4,6,2	6,1,5,3,4,2
P2	3,4,2,6,1,5		1,5,2,4,3,6	4,3,1,5,6,2	3,4,2,1,5,6		4,3,2,1,6,5			5,1,2,6,4,3
P3	4,6,2,3,1,5		4,6,1,3,2,5	4,1,2,3,6,5	4,6,2,3,1,5		4,6,1,3,2,5		1,3,4,6,2,5	5,4,6,3,1,2
P4	4,2,6,5,3,1	6,5,4,2,3,1	3,5,2,4,6,1	4,5,6,3,2,1	4,2,6,5,3,1		3,1,6,5,4,2		3,1,5,6,2,4	6,5,4,3,1,2

Here and below, “DF” denotes degrees of freedom.

TABLE 6 VaR and CVaR at the 99% confidence level, estimated using the different selection orders in Table 5: confidence intervals are in parentheses.

(a) VaR

	Empirical criterion				Student t copula-based criterion					
Portfolio	ρ	τ	λ_U	λ_L	ρ	τ	λ_U	λ_L	**DF**	**Likelihood**
P1	3.255 (3.014,3.399)		3.364 (3.257,3.644)	3.250 (3.017,3.372)	3.253 (3.005,3.373)		3.338 (3.180,3.563)		3.346 (3.207,3.588)	3.193 (2.856,3.245)
P2	3.071 (2.963,3.282)		2.957 (2.705,3.038)	3.017 (2.812,3.138)	3.034 (2.873,3.198)		3.030 (2.905,3.217)			3.026 (2.851,3.171)
P3	3.551 (3.538,3.908)		3.494 (3.441,3.797)	3.487 (3.430 3.805)	3.551 (3.538,3.908)		3.494 (3.424,3.794)		3.489 (3.410,3.758)	3.460 (3.262,3.628)
P4	2.473 (2.265,2.540)	2.499 (2.317,2.592)	2.464 (2.255,2.524)	2.533 (2.403,2.663)	2.473 (2.265,2.540)		2.501 (2.335,2.599)		2.522 (2.373,2.639)	2.518 (2.369,2.635)

(b) CVaR

	Empirical criterion				Student t copula-based criterion					
Portfolio	ρ	τ	λ_U	λ_L	ρ	τ	λ_U	λ_L	**DF**	**Likelihood**
P1	1.210 (0.906,1.481)		1.052 (0.626,1.161)	1.260 (1.002,1.587)	1.211 (0.959,1.471)		1.120 (0.755,1.286)		1.156 (0.822,1.342)	1.270 (1.071,1.599)
P2	0.819 (0.591,0.944)		0.861 (0.713,1.043)	0.800 (0.551,0.895)	0.840 (0.622,0.973)		0.836 (0.634,0.987)			0.785 (0.527,0.879)
P3	1.066 (0.748,1.239)		1.125 (0.867,1.365)	1.103 (0.861,1.357)	1.066 (0.748,1.239)		1.125 (0.866,1.357)		1.085 (0.781,1.294)	1.093 (0.774,1.292)
P4	0.921 (0.740,1.125)	0.905 (0.708,1.093)	0.884 (0.683,1.070)	0.853 (0.630,1.000)	0.921 (0.740,1.125)		0.898 (0.689,1.087)		0.857 (0.604,1.006)	0.865 (0.645,1.021)

TABLE 7 VaR and CVaR at the 99.5% level, estimated using the different selection orders in Table 5: confidence intervals are in parentheses.

(a) VaR

	Empirical criterion				Student t copula-based criterion					
Portfolio	ρ	τ	λ_U	λ_L	ρ	τ	λ_U	λ_L	**DF**	**Likelihood**
P1	4.019 (3.753,4.317)		4.019 (3.776,4.348)	4.068 (3.856,4.415)	4.127 (3.997,4.512)		4.076 (3.856,4.413)		4.041 (3.788,4.357)	3.995 (3.673,4.248)
P2	3.613 (3.450,3.893)		3.533 (3.287,3.725)	3.495 (3.153,3.614)	3.641 (3.493,3.937)		3.655 (3.561,4.008)			3.497 (3.175,3.652)
P3	4.135 (3.940,4.488)		4.164 (4.011,4.575)	4.221 (4.146,4.704)	4.135 (3.940,4.488)		4.164 (3.998,4.546)		4.193 (4.054,4.575)	4.073 (3.703,4.233)
P4	3.068 (2.855,3.271)	3.099 (2.917,3.333)	3.046 (2.829,3.234)	3.038 (2.816,3.222)	3.068 (2.855,3.271)		3.050 (2.809,3.238)		3.062 (2.832,3.255)	3.087 (2.883 3.320)

(b) CVaR

	Empirical criterion				Student t copula-based criterion					
Portfolio	ρ	τ	λ_U	λ_L	ρ	τ	λ_U	λ_L	**DF**	**Likelihood**
P1	1.315 (0.691,1.637)		1.207 (0.609,1.428)	1.355 (0.779,1.722)	1.239 (0.647,1.486)		1.145 (0.457,1.301)		1.286 (0.696,1.554)	1.408 (1.001,1.820)
P2	0.856 (0.442,1.018)		0.865 (0.563,1.063)	0.917 (0.599,1.138)	0.758 (0.270,0.822)		0.779 (0.309,0.861)			0.895 (0.565,1.098)
P3	1.292 (0.875 1.671)		1.276 (0.866,1.658)	1.162 (0.712,1.458)	1.292 (0.875,1.671)		1.276 (0.843,1.639)		1.136 (0.603,1.365)	1.303 (0.917,1.686)
P4	1.013 (0.695,1.277)	0.966 (0.601,1.183)	0.907 (0.482,1.088)	0.955 (0.603,1.195)	1.013 (0.695,1.277)		0.996 (0.624,1.250)		0.964 (0.551,1.188)	0.922 (0.524,1.097)

We generate 500 samples of size 500 from four pair–copula decompositions, two with a dimension of four (dim $=$ 4) and two with a dimension of six (dim $=$ 6). In total, then, we have four models, named Model 1 (dim $=$ 4), Model 2 (dim $=$ 6), Model 3 (dim $=$ 4) and Model 4 (dim $=$ 6). We use the D-vine based on the given orders defined as $(1, 2, 3, 4)$ for dim $=$ 4 and defined as $(1, 2, 3, 4, 5, 6)$ for dim $=$ 6. For each pair we use a Student t copula. The theoretical parameters used in the simulated models are shown in Table 12 of the online appendix. These theoretical parameters are defined based on those obtained for the analyzed portfolios in Section 3, ie, assuming greater (Model 3 and Model 4) or lesser (Model 1 and Model 2) dependency between random variables in the multivariate vector. Moreover, in Table 13 (see the online appendix) the true VaR and CVaR that we use to obtain the MSE associated with alternative selection criteria are shown.

The simulation study is carried out in the following way. As a first step, using each of the four theoretical models, we generate 500 samples of size 500, ie, we have four simulated data sets. The second step consists of applying the pair-selection procedure described in Section 2.1.1 to each sample in each data set. The third step involves estimating the parameters of the D-vines associated with each order and using a Student t pair–copula. The fourth and final step consists of using the Monte Carlo method to estimate the VaR and CVaR of each sample. To obtain these results, we assume the marginals are uniform (0, 1) distributed.

The MSE estimations of model j's estimated VaR_j and CVaR_j using criterion k are obtained, respectively, using

$$\widehat{\mathrm{MSE}}_k(\widehat{\mathrm{VaR}}_j) = \frac{1}{500}\sum_{i=1}^{500}(\mathrm{VaR}_{ij}^k - \mathrm{VaR}_j)^2 \tag{4.1}$$

and

$$\widehat{\mathrm{MSE}}_k(\widehat{\mathrm{CVaR}}_j) = \frac{1}{500}\sum_{i=1}^{500}(\mathrm{CVaR}_{ij}^k - \mathrm{CVaR}_j)^2, \tag{4.2}$$

where k refers to, on the one hand, the criterion based on empirical ρ, τ, λ_{U} and λ_{L}, and, on the other hand, the criterion based on ρ, τ, λ_{U}, λ_{L} and the degree of freedom estimated using the Student t copula.

In practice, to eliminate the size effect given the different theoretical values, we calculate a relative MSE, which is the square root of the MSE divided by the corresponding theoretical VaR or CVaR. The results for VaR and CVaR are shown in Table 8. We compare all cases, using the statistic for testing the value of the estimated mean of VaR and CVaR obtained from the 500 simulated samples, with the null hypothesis, which assumes that estimated values are equal to theoretical values. In both parts of Table 8, we indicate using underlined italics the cases in which a statistic does not reject the null hypothesis.

TABLE 8 Square root of MSE divided by the corresponding theoretical values of VaR and CVaR obtained with the different criteria of selection orders using a Student t copula.

(a) Empirical criterion

	VaR								CVaR							
	τ		ρ		λ_U		λ_L		τ		ρ		λ_U		λ_L	
	99%	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**
Model 1	0.0175	0.0145	0.0184	0.0137	0.0177	0.0144	0.0175	0.0144	0.2515	0.3599	0.2592	0.3510	0.2632	0.3551	0.2603	0.3539
Model 2	0.0106	0.0307	*0.0092*	0.0293	*0.0092*	0.0298	*0.0093*	0.0297	0.6942	0.4836	0.6945	0.4816	0.6856	0.4801	0.6891	0.4806
Model 3	0.0111	0.0131	0.0112	0.0130	0.0113	0.0131	0.0121	0.0121	0.2670	0.5157	0.2702	0.5117	0.2741	0.5143	0.2798	0.4927
Model 4	0.0135	0.0139	0.0134	0.0137	0.0133	0.0140	0.0135	0.0137	*0.2708*	0.4323	*0.2670*	0.4306	*0.2679*	0.4359	*0.2696*	0.4351

(b) Student t copula-based criterion

	VaR								CVaR							
	τ		ρ		λ_U, λ_L		DF		τ		ρ		λ_U, λ_L		DF	
	99%	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**
Model 1	0.0186	0.0136	0.0186	0.0136	0.0169	0.0149	0.0187	0.0135	0.2606	0.3467	0.2606	0.3467	0.2586	0.3545	0.2593	0.3521
Model 2	*0.0093*	0.0297	*0.0093*	0.0297	0.0097	0.0305	0.0089	0.0285	0.6881	0.4796	0.6881	0.4796	0.6857	0.4812	0.6957	0.4792
Model 3	0.0115	0.0127	0.0115	0.0127	0.0115	0.0127	0.0127	0.0113	0.2755	0.5052	0.2755	0.5052	0.2764	0.5021	0.2818	*0.4714*
Model 4	*0.0134*	*0.0135*	*0.0134*	*0.0135*	0.0136	0.0137	0.0151	0.0121	*0.2681*	*0.4259*	*0.2681*	*0.4259*	*0.2715*	*0.4337*	0.2744	0.4077

TABLE 9 Square root of MSE divided by the corresponding theoretical values of VaR and CVaR obtained with the different criteria of selection orders using the Gumbel copula.

(a) Empirical criterion

	VaR								CVaR							
	τ		ρ		λ_U		λ_L		τ		ρ		λ_U		λ_L	
	99%	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**
Model 3	0.0129	0.0231	0.0129	0.0231	0.0126	0.0228	0.0126	0.0228	0.2472	0.8422	0.2473	0.8433	0.2426	0.8300	0.2417	0.8337
Model 4	0.0188	0.0313	0.0187	0.0313	0.0178	0.0304	0.0180	0.0305	0.2489	0.9158	0.2481	0.9157	0.2386	0.8792	0.2408	0.8848

(b) Gumbel copula-based criterion

	VaR								CVaR							
	τ		ρ		λ_U		DF		τ		ρ		λ_U		DF	
	99%	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**	**99%**	**99.5%**
Model 3	0.0128	0.0230	0.0128	0.0230	0.0128	0.0230	—	—	0.2452	0.8345	0.2452	0.8345	0.2452	0.8345	—	—
Model 4	0.0189	0.0313	0.0189	0.0313	0.0189	0.0313	—	—	0.2521	0.9146	0.2521	0.9146	0.2521	0.9146	—	—

The unmarked cases are those in which the estimated VaR or CVaR are biased. This occurs in Model 1 with dim $= 4$, using any of the criteria for selecting the D-vine, as well as in Model 3, also with dim $= 4$, except for the CVaR at 99.5%, which is estimated with the criterion based on DF. For Model 2 and Model 4 with dim $= 6$, the results improve when we assume higher dependency and use criteria based on the known copula. In practice, the copula is not known and is based on statistical goodness-of-fit criteria.

To evaluate the effect on the MSE when we use a different copula, we calculate results on Model 3 and Model 4 using the Gumbel copula. These are shown in Table 9. In this case, we reject the null hypothesis that estimated values are equal to theoretical values for VaR and CVaR at the 99% and 99.5% levels. In this case, the MSE of estimated VaR increases, especially for the 99.5% confidence levels. The results for CVaR show that at the 99% level the relative MSE is similar to that obtained using the Student t copula, while at the 99.5% level the relative MSE increases considerably.

In those cases where the null hypothesis is rejected, we need to assume some bias in the estimations of risk, which can be positive or negative. In our simulation study, this bias is quite small for the VaR; it does not reach 4% of the theoretical value in any case. However, for the CVaR the bias much bigger. When the confidence level is 99.5%, for Model 2 (dim $= 6$) it may exceed 65% of the theoretical value.

5 CONCLUSIONS

We showed that different orders in the first tree of the D-vine provide different estimations of VaR and CVaR. The higher or lower dispersion of these different estimations depends on the selected copula, ie, the model for pair dependence has an important role.

The dispersion of the estimated VaR and CVaR highlights the problem of selecting the optimal order. Maximum likelihood estimation requires an estimation of all possible D-vines. Further, this estimation may not be the best choice if the objective is to estimate the VaR or CVaR; alternative criteria based on dependence measures should be used instead. In our empirical analysis, in general, we found that, by using different criteria, we obtained different pair–copula decompositions. However, the estimated measures of VaR and CVaR do not show significant differences when we use a copula that offers a better fit: in our case, the Student t copula.

In our simulation study, we compared models with four and six dimensions and obtained better results for the latter. In some cases, VaR and CVaR estimations are biased, and this bias can increase considerably if the copula is not true. Given this bias, it is advisable to use different criteria for selecting the D-vine and to evaluate the differences between the different estimations of risk.

DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

ACKNOWLEDGEMENTS

We thank the Spanish Ministry of Economy and Competitiveness for its support through ECO2016-76203-C2-2-P. We also thank the members and affiliates of Riskcenter at the University of Barcelona.

REFERENCES

Aas, K., Czado, C., Figessi, A., and Bakken, H. (2006). Pair–copula constructions of multiple dependence. *Insurance: Mathematics and Economics* **44**, 182–198 (https://doi.org/10.1016/j.insmatheco.2007.02.001).

Bahraoui, Z., Bolancé, C., and Pérez-Marín, A. (2014). Testing extreme value copulas to estimate the quantile. *SORT-Statistics and Operations Research Transactions* **38**, 89–102.

Bedford, T., and Cooke, R. M. (2002). Vines: a new graphical model for dependent random variables. *Annals of Statistics* **30**, 1031–1068 (https://doi.org/10.1214/aos/1031689016).

Bolancé, C., Bahraoui, Z., and Artís, M. (2014). Quantifying the risk using copulae with nonparametric marginals. *Insurance: Mathematics and Economics* **58**, 46–56 (https://doi.org/10.1016/j.insmatheco.2014.06.008).

Boos, D. D. (1982). A test for asymmetry associated with the Hodges–Lehmann estimator. *Journal of the American Statistical Association* **77**, 647–651 (https://doi.org/10.1080/01621459.1982.10477864).

Brechmann, E. C., and Czado, C. (2013). Risk management with high-dimensional vine copulas: An analysis of the Euro Stoxx 50. *Statistics and Risk Modeling* **30**, 307–342 (https://doi.org/10.1524/strm.2013.2002).

Brechmann, E. C., and Schepsmeier, U. (2013). CDvine: modeling dependence with C- and D-vine copulas in R. *Journal of Statistical Software* **52**, 1–27 (https://doi.org/10.18637/jss.v052.i03).

Christoffersen, P., Jacobs, K., Jin, X., and Langlois, H. (2013). Dynamic dependence in corporate credit. Technical Report, Bauer College of Business, University of Houston.

Clayton, D. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* **65**, 141–151 (https://doi.org/10.1093/biomet/65.1.141).

Denuit, M., Dhaene, J., Goovaerts, M., and Kaas, R. (2005). *Actuarial Theory for Dependent Risks: Measures, Orders and Models*. Wiley (https://doi.org/10.1002/0470016450).

Dissmann, J., Brechmann, E., Czado, C., and Kurowicka, D. (2013). Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics and Data Analysis* **59**, 52–69 (https://doi.org/10.1016/j.csda.2012.08.010).

Fernández, C., and Steel, M. F. (1998). On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association* **93**, 359–371 (https://doi.org/10.1080/01621459.1998.10474117).

Frank, M. (1979). On the simultaneous associativity of and $x + y - F(x, y)$. *Aequationes Math* **19**, 194–226 (https://doi.org/10.1007/BF02189866).

Genest, C., Ghoudi, K., and Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* **82**, 543–552 (https://doi.org/10.1093/biomet/82.3.543).

Gumbel, E. J. (1960). Bivariate exponential distributions. *Journal of the American Statistical Association* **55**, 698–707 (https://doi.org/10.1080/01621459.1960.10483368).

Low, R. K. Y., Alcock, J., Faff, R., and Brailsford, T. (2013). Canonical vine copulas in the context of modern portfolio management: are they worth it? *Journal of Banking and Finance* **37**, 3085–3099 (https://doi.org/10.1016/j.jbankfin.2013.02.036).

Low, R. K. Y., Faff, R., and Aas, K. (2016). Enhancing mean–variance portfolio selection by modeling distributional asymmetries. *Journal of Economics and Business* **85**, 49–72 (https://doi.org/10.1016/j.jeconbus.2016.01.003).

McNeil, A. J., Frey, R., and Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press.

Min, A., and Czado, C. (2010). Bayesian inference for multivariate copulas using pair–copula construction. *Journal of Financial Econometrics* **8**, 511–546 (https://doi.org/10.1093/jjfinec/nbp031).

Min, A., and Czado, C. (2014). Scomdy models based on pair–copula constructions with application to exchange rates. *Computational Statistics and Data Analysis* **76**, 523–535 (https://doi.org/10.1016/j.csda.2012.08.003).

Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer Science and Business Media, New York.

Nikoloulopoulos, A. K., Joe, H., and Li, H. (2012). Vine copulas with asymmetric tail dependence and applications to financial return data. *Computational Statistics and Data Analysis* **56**, 3659–3673 (https://doi.org/10.1016/j.csda.2010.07.016).

Oh, D. H., and Patton, A. (2013). Time-varying systemic risk: evidence from a dynamic copula model of cds spreads. Technical Report, Duke University, Department of Economics.

Rad, H., Low, R. K. Y., and Faff, R. (2016). The profitability of pairs trading strategies: distance, cointegration and copula methods. *Quantitative Finance* **16**, 1541–1558 (https://doi.org/10.1080/14697688.2016.1164337).

Righi, M. B., and Ceretta, P. S. (2013). Analyzing the dependence structure of various sectors in the Brazilian market: a pair copula construction approach. *Economic Modelling* **35**, 199–206 (https://doi.org/10.1016/j.econmod.2013.06.012).

Righi, M. B., and Ceretta, P. S. (2015). Forecasting value at risk and expected shortfall based on serial pair–copula constructions. *Expert Systems with Applications* **42**, 6380–6390 (https://doi.org/10.1016/j.eswa.2015.04.023).

Schmidt, R., and Stadtmüller, U. (2006). Non-parametric estimation of tail dependence. *Scandinavian Journal of Statistics* **33**, 307–335 (https://doi.org/10.1111/j.1467-9469.2005.00483.x).

Silverman, B. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman & Hall/CRC Finance Series, London (https://doi.org/10.1007/978-1-4899-3324-9).

Sklar, A. (1959). Fonctions de répartition a n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris* **8**, 229–231.

Vaz de Melo Mendes, B., Mendes Semeraro, M., and Câmara Leal, R. P. (2010). Pair–copulas modeling in finance. *Financial Markets and Portfolio Management* **24**, 193–213 (https://doi.org/10.1007/s11408-010-0130-1).

Weiß, G., and Scheffer, M. (2015). Mixture pair–copula-constructions. *Journal of Banking and Finance* **54**, 175–191 (https://doi.org/10.1016/j.jbankfin.2015.01.008).

Weiß, G. N., and Supper, H. (2013). Forecasting liquidity-adjusted intraday value-at-risk with vine copulas. *Journal of Banking and Finance* **37**, 3334–3350 (https://doi.org/10.1016/j.jbankfin.2013.05.013).