Advances in Industrial Engineering, Autumn 2021, 55(4): 403-432
DOI: 10.22059/aie.2022.336312.1818
RESEARCH PAPER

Providing an Optimal Model in Modeling the Dependence Structure of the Elements of Financial Systems Using an Approach Based on Vine-Copula Functions. (Case Study: Market and Industry Indices at Tehran Stock Exchange)

Mohammad Sadegh Zarrin Nal, Hojjatollah Sadeqi [*]

Department of Accounting and Finance, Faculty of Humanities and Social Sciences,Yazd University, Yazd, Iran.

Received: 27 December 2021, Revised: 06 Febrauary 2022, Accepted: 06 Febrauary 2022

Abstract

Identification of the structure of dependence among different elements of a financial system has long been a hot topic to researchers due to its impact on the financial asset risk assessment. Currently, the capital market is one of the key financial systems in Iran’s economy, making the understanding and identification of its intra-system associations a major concern to investors and investment managers who seek to forecast future conditions. Accordingly, the present research investigates and models the dependence structure of different market indices of the Tehran Stock Exchange (TSE), as a representative of the country’s financial system, and the indices referring to the active industries in the TSE, as a component of the financial system. We herein investigated a total of 10 market indices and 31 other indices referring to the most significant active industries in the TSE. The mentioned industries were clustered based on three distinctive scenarios. Considering the number of components and the abnormal structure of their distributions and also taking into account the importance of marginal distributions in the assessment of the system component dependence structure model, we found the copula functions as a useful tool for expressing the dependence between different variables. In this research, the dependence structure of the market and industry indices of the TSE was investigated using two subroutines of the vine-copula functions, namely C-Vine and R-Vine. The results were then studied using Vuong’s test. The outcomes indicated that the C-Vine functions can generate very good fits to the dependence structures among various industry indices. Moreover, the best fits could be explained using the t-student family of the copula functions. Based on the results of the present study, it is possible to evaluate the relationships between industry indicators and the impact of different market industries from changes in the whole market and make optimal decisions in choosing the portfolio composition.

Keywords:
Vine-Copula Functions;
R-Vine;
C-Vine;
Dependence Structure;
Indices of TSE

Introduction

The understanding and modeling of the dependence of the return on different financial assets play a key role in the allocation of the assets and configuration of investment strategies.

[*] Corresponding author: (H. Sadeqi)
Email: sadeqi@yazd.ac.ir

Financial risk management is particularly affected by the mutual dependence of the financial assets and markets, making it necessary for the investors to model and quantify the dependence structure and intensity [1]. Financial fluctuations can be transferred between different assets in the same market or even different markets. In this way, the nature of the dependence between financial returns and financial market conditions and its impact on investment is a hot topic in financial management. Accordingly, comprehension of the associations among financial assets contributes largely to the decision-making about the investment in such assets. Therefore, identification of the dependence structure between assets and financial markets has long been a field of interest to researchers. Since the 2007 – 2008 financial crisis, many efforts have been made to configure the mutual impacts of the elements of a financial system on one another. The relevant research has been classified under the field of the systemic risk studies and is focused on the probability of the events that lead to action-reaction schemes in the entire financial system and hence incur losses to the practitioners across the system. Quantification of system variations has been among the most important challenges faced by financial theories in the recent past – a challenge that requires an understanding of the dependence structure between different elements of the financial system [2].

In all financial markets around the world, the stock market index serves as a good indicator of the performance of the stock exchange, making it highly regarded by practitioners. This is because such indices represent a collective measure of the trends followed by the prices of all or a particular segment of the listed companies and hence enable the assessment of the direction and intensity of the price movements in the market. Prediction of this index provides some information on the future trend and overall state of the market. On the other hand, investment in the stock exchange represents an important sector of the economy. Thus, the subject matters of prediction and design of robust prediction models for developing countries like Iran are of paramount importance for realizing proper management of the stock exchange toward sustainable development. Such models further smooth the way for executive decision-makers of the stock exchange when making a decision under uncertain conditions [3].

The industries with activities in the capital market can exhibit high degrees of dependence on one another. Therefore, assessing the risk of one industry without considering its dependencies on other relevant industries may not lead to proper results [4]. On this basis, the investigation of the dependency structure of the returns for different industries in a portfolio is highly useful. In the literature on the financial economy, joint distribution modeling is a major challenge against achieving the mentioned objective. Traditional investigations assumed that such dependency structure can be extracted from a linear association model. Although this approach has been implemented in numerous economic contexts, its results are reliable only when the considered variables follow normal distribution [5].

A very popular measure of dependency between two variables is the linear correlation coefficient, which is based on the very fundamental assumption of the Gaussian normality of the distributions of the two variables. As has been indicated by Lanchin and Solnic (2001), Karman and Herra (2014), and some other researchers in the field of financial dependency modeling, the financial data do not follow a Gaussian (*i.e.*, normal) distribution in many cases. Accordingly, using the correlation coefficient alone as a measure of the dependency between financial variables oftentimes leads to misleading results. In addition, the normal correlation coefficient measures the dependency between assets and/or financial markets in terms of intensity rather than providing information on their dependency structure (*i.e.,* how the assets are associated with financial markets under different conditions [1].

In this respect, the search for multivariate and flexible distributions has made copula modeling a popular approach in many contexts (Berchman and Shepsmear, 2012). The copula function is a useful yet flexible tool for generating a multivariate distribution function that links a group of marginal distributions to one another to come up with a *joint distribution* (Nelson, 2005). This key secret behind the attractiveness of copulas is that they can model the behavior of marginal univariate distributions corresponding to each random variable independently from the dependency between the random variables [1].

Although the copula function has outperformed other indices in describing the dependency structure yet the studies published by Berchman and Shepsmear (2012), Dimen *et al.* (2013), and Gazado *et al.* (2014) showed that these functions cannot perform as successfully for larger numbers of variables. In many cases, multivariate data exhibit such complex patterns as asymmetry and dependency at margins. Accordingly, implementation of simple copula tends to impose some structural limitations in the model, questioning the reliability of the outcomes. Generalization of these models to hierarchical copula functions can end up with some improvements but at the same time increases the complexity of the dependence structure and brings about such limitations as the parametric constraints [6]. Considerable effort has been made to build flexible models, with the vine copulas acknowledged as one of the best instances of such efforts (Gazado *et al.*, 2014). The vine copulas were first introduced by Joe (1996) and then expanded by Bedford and Kodak (2001, 2002) and then Korovebca and Cock (2006). The vine copulas enable proper identification of dependence structure between variable pairs as they are capable of capturing the symmetry/asymmetry, association intensity, and tail dependence [7].

As mentioned, the basic premise of classical financial models in identifying structures and estimating the dependency pattern of variables has been to have a normal distribution. Based on the above, in this study, contrary to the classical financial assumptions, Copula functions have been used to identify the structure of dependence. A distinctive feature of these functions is the non-consideration of assumptions regarding the distribution of survey data, which distinguishes it from classical financial models. The present study seeks to examine the relationship between dependence between market indicators and industries active in the Tehran Stock Exchange and identify patterns of dependence.

Based on what was mentioned so far, the present research aims at investigating the dependency between market indices and active industries in the Tehran Stock Exchange (TSE) to identify the dependency models. For this purpose, considering the relatively large number of the industry indices, we began by clustering the entire pool of industries under several scenarios. Next, the dependence structure of industry indices and market indices was carried out using two subroutines of vine copulas, namely C-Vine and R-Vine. The modeling results indicate the effects of different industries and indices on one another and clearly expressed the dependency structure in the financial system of the capital market. The current study was further aimed at selecting the best copula function for explaining the dependence structure of indices. Application of these functions to account for the tail dependence between different industries and the financial market of the country is of paramount importance in investment portfolio management. In other words, the objectives of this study include identifying the best statistical distribution in the financial data under study, fitting the dependency structure between distributions based on Copula functions, and finally recognizing and mapping the dependency pattern between individual industry indices and the relationship between market indices and indices. Is an industry.

The present contribution is composed of sections on introduction, theoretical foundations and research background, methodology, analysis of the results and experimental findings, and finally, a discussion on the results and conclusions.

Theoretical foundations and research background

Theoretical foundations

Copula functions and Sklar's theorem

According to Kurowicka *et al.* (2012), the credit of presenting the copula models as an efficient and popular tool in financial science goes to Joe (1997) and Nelssen (2006). In the field of mathematics and statistics, the term *copula* was coined by Sklar (1959) as he presented his theorem called Sklar's theorem [8].

As a fundamental theorem in copula modeling, Sklar's theorem (1959) shows how a multivariate distribution of a dataset can be decomposed into copulas and marginal distributions of single data points. This theorem is of profound importance in statistical modeling of multivariate distributions functions and serves as a foundation for building pair copula functions. According to Sklar's theorem, letting H be a two-dimensional distribution function with marginal distributions F and G, there is a copula function C for which $H(x,y) = C(F(x),G(y))$. Moreover, for each distribution function (*e.g.*, F and G) and copula function C, the joint distribution function H has two marginal distribution functions F and G. Should the functions F and G be continuous, then function C is unique.

$$H(F^{-1}(u),G^{-1}(v)) = C(u,v) \tag{1}$$

In the above equation, F^{-1} is the inverse RCLL[†] of F. If the random variables X and Y are continuous with the above distribution functions, then the function C expresses the joint distribution function of the uniformly distributed random variables F(X) and G(Y) [9].

As expressed by Fischer (1997) in his encyclopedia of statistics, copulas became popular for two main reasons. First, they provide a dimensionless measure of dependence and association between random variables. And second, copulas serve as a starting point for producing a family of bivariate and multivariate distributions [8]. The usefulness of the copula functions as a tool for modeling cross-sectional dependence structures between random variables is sourced from the capability of these functions in distinguishing between marginal distributions and joint dependence between variables. Pair copulas can be defined as follows: letting X and Y be continuous random variables with distribution functions $F(x) = P(X \leq x)$ and $G(y) = P(Y \leq y)$, respectively, and the joint distribution function $H(x, y) = P(X \leq x, Y \leq y)$, then for each pair ($x$, y) in the $[-\infty,\infty]^2$ space, there is a corresponding point in the $(I = [0,1])I^3$ space with the coordinates$(F(x), G(y), H(x,y))$. This mapping from the I^2 space to the I space is referred to as the copula. In other words, a copula is a function with I^2 as its domain and I as its range, by which the following equations hold true for all $x \in I$.

$$C(0,x) = C(x,0) = 0 \tag{2}$$

$$C(1,x) = C(x,1) = x \tag{3}$$

Furthermore, considering the copula equations, the following relationship is always satisfied for all values of $a,b,c,d \in I$ provided $a \leq b, c \leq d$:

$$V_c([a,b]) \times [c,d]) = C(b,d) - C(a,d) - C(b,c) + C(a,c) \geq 0 \tag{4}$$

[†] right continuous with left limits

In the above equation, the function V_C represents the volume of the rectangle $[a, b] \times [c, d]$ under the function C [10].

In a multivariate case, the copula functions can be defined as those which combine the information about the dependence structure between $n > 2$ random variables (*i.e.*, $x_1, x_2, \ldots, x_n$) [9].

Families of copula functions

The most popular copula functions with applications in the modeling of dependence structure in financial systems are described in the following.

Normal or Gaussian copula

A multivariate normal copula function has a correlation matrix Σ for the parameters. Since the correlation coefficient has long been perceived as an important index in the financial arena, normal copulas have been widely applied in finance. Although they are most commonly used to simplify different procedures. A normal copula function is developed by combining a standard univariate normal copula function with a standard *n*-variable function, which is usually denoted by Φ and Φ, respectively. The following expression characterizes a normal copula function [11]:

$$C(u_1, \ldots, u_n; \Sigma) = \Phi(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n)) \tag{5}$$

Expressing a copula function in deterministic form is usually difficult, and it rather can be expressed using integrals. This makes it easier to work with the copula densities rather than the copula functions. The density function of the previous equation is written as follows:

$$c(u_1, \ldots, u_n; \Sigma) = |\Sigma|^{-\frac{1}{2}} exp\left(-\frac{1}{2}\xi'(\Sigma^{-1} - 1)\xi\right) \tag{6}$$

in which $(\xi = \xi_1, \ldots, \xi_n)'$ and ξ_i refers to u_i quantiles of the standard normal variable X_i:

$$u_i = P(X_i < \xi_i) \tag{7}$$

$$X_i \sim N(0,1) \tag{8}$$

It is worth mentioning that a normal copula is a function of $(u_1, \ldots, u_n)$ rather than $(\xi_1, \ldots, \xi_n)$ because we have $\xi = (\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n))$. One may go through the following steps to obtain a joint distribution function when the copula is normal but the marginal distributions are either normal or abnormal [11]:

1. For the marginal functions, we set .
2. We use inverse of the normal distribution function.
3. We use the correlation matrix $\sum$ and the vector ξ in Equation (--).

When we have only two random variables, the normal copula function takes the following form:

$$C(u_1, u_2; \varrho) = \Phi(\Phi^{-1}(u_1), \Phi^{-1}(u_2)) \tag{9}$$

where Φ is a bivariate normal function. Using the formula for the bivariate normal function, the above equation can be rewritten as follows:

$$C(u_1,u_2;\varrho) = \int_0^{\Phi^{-1}(u_1)} \int_0^{\Phi^{-1}(u_2)} (2\pi)^{-1} - \varrho^2)^{-\frac{1}{2}} exp(- \frac{[x_1^2 - 2\varrho x_1 x_2 + x_2^2]}{2(1-\varrho^2)} dx_1 dx_2 \tag{10}$$

Density function of the normal copula is as follows:

$$c(u_1,u_2;\varrho) = (1-\varrho^2)^{-\frac{1}{2}} exp\left(-\frac{\varrho^2\xi_1^2 - 2\xi_1\xi_2 + \varrho^2\xi_2^2}{2(1-\varrho^2)}\right) \tag{11}$$

The fact that this equation has the correlation coefficient as its only parameter makes its estimation an easy task to do. The family of normal copulas has been categorized under symmetric copulas, implying that $c(u_1,u_2) = c(u_2,u_1)$. Moreover, the normal copulas exhibit zero or near-zero dependence in the distribution tail except for the cases where the correlation coefficient is 1. This characteristic is not desired when the modeling of return on financial assets is concerned [11].

t-student copula

Similar to normal copulas, the *n*-dimensional t-student copula function is derived from a multivariate distribution function. A t-student copula function can be expressed as follows [11]:

$$C_v(u_1,\dots,u_n;\Sigma) = t_v(t_v^{-1}(u_1),\dots,t_v^{-1}(u_n)) \tag{12}$$

where t_v and t_v^{-1} are a multivariate t-student distribution function and a univariate t-student distribution function with v degrees-of-freedom (DOFs), respectively. Similar to the normal copula, here we have Σ as the correlation matrix. The t-student copula is also difficult to express in a deterministic form. This is usually done with the help of a multivariate t-student density function. A multivariate t-student density function is expressed as follows:

$$f(X) = k|\Sigma|^{-\frac{1}{2}}(1 + v^{-1}X'\Sigma^{-1}X)^{-\frac{(v+n)}{2}} \tag{13}$$

$$k = \Gamma(\frac{v}{2})^{-1}\Gamma(\frac{v+n}{2})(v\pi)^{-\frac{n}{2}} \tag{14}$$

Accordingly, one can obtain the following expression for an *n*-variate t-student copula function:

$$C_v(u_1,\dots,u_n;\Sigma) = \int_0^{t_v^{-1}(u_1)} \dots \int_0^{t_v^{-1}(u_n)} k|\Sigma|^{-\frac{1}{2}}(1 + v^{-1}X'\Sigma^{-1}X)^{-\frac{(v+n)}{2}} dx_1 \dots dx_n \tag{15}$$

Taking derivative of the above equation, one can develop the corresponding t-student copula density function, as follows:

$$C_v(u_1,\dots,u_n;\Sigma) = k|\Sigma|^{-\frac{1}{2}}(1 + v^{-1}\xi'\Sigma^{-1}\xi)^{-\frac{(v+n)}{2}} \prod_{i=1}^{n}(1 + v^{-1}\xi_i^2)^{\frac{(v+n)}{2}} \tag{16}$$

$$\xi = (t_v^{-1}(u_1),\dots,t_v^{-1}(u_n)) \tag{17}$$

$$K = \Gamma(\frac{v}{2})^{n-1}\Gamma(\frac{v+1}{2})^{-n}\Gamma(\frac{v+n}{2}) \quad (18)$$

With only two random variables, the t-student copula function and density function can be expressed as follows:

$$C(u_1,u_2;\varrho) = \int_0^{\Phi^{-1}(u_1)}\int_0^{\Phi^{-1}(u_2)} (2\pi)^{-1}(1-\varrho^2)^{-\frac{1}{2}}[1+v^{-1}(x_1^2 - 2\varrho x_1 x_2 + x_2^2)]^{-\frac{(v+2)}{2}} dx_1 dx_2 \quad (19)$$

$$c(u_1,u_2;\varrho) = K(1-\varrho^2)^{-\frac{1}{2}}[1+v^{-1}(1-\varrho^2)^{-1}(\xi_1^2 - 2\varrho\xi_1\xi_2 + \xi_2^2)]^{-\frac{(v+2)}{2}} \times [(1+v^{-1}\xi_1^2)(1+v^{-1}\xi_2^2)]^{-\frac{(v+1)}{2}} \quad (20)$$

Archimedean copulas

Similar to the normal and t-student copulas, the ellipsoidal copulas are based on multivariate distribution functions. Another approach to developing a copula is the use of a generator function. In this research, the generator function is denoted by $\Psi(u)$. Using the generator function Ψ, the corresponding Archimedean copulas are defined as follows [12]:

$$C(u_1,\dots,u_n) = \Psi^{-1}\big(\Psi(u_1)+\cdots+\Psi(u_n)\big) \quad (21)$$

For which the density function is described as follows:

$$C(u_1,\dots,u_n) = \Psi_{(n)}^{-1}(\Psi(u_1)+\dots+\Psi(u_n))\prod_{i=1}^{n}\Psi'(u_i) \quad (22)$$

in which $\Psi_{(n)}^{-1}$ is the n^{th}-order derivative of the inverse of the generator function. Since there are numerous generator functions, one can come up with different Archimedean copulas. Nelson (2006) alone has defined some 22 univariate Archimedean copulas in his book.

In this subsection, we introduce the three most popular Archimedean copulas in the field of risk management. These three copulas include Clayton copula, Gumbel copula, and Frank copula. The Clayton and Gumbel copulas exhibit asymmetric dependence in their distribution tail – a fact that makes them significant to us. Clayton copula shows lower-tail dependence while Gumbel copula exhibits upper tail dependence [12].

In the following, characteristics of different families of copula functions are presented in brief [13].

Table 1. Family of ellipsoidal copula functions.

#	Copula function	Domain of parameters	Kendall's τ coefficient (2)	Upper/lower tail dependence
1	Gaussian	$\rho \in (-1,1)$	$\frac{2}{\pi} arcsin(\rho)$	0
2	t-student	$\rho \in (-1,1), v > 2$	$\frac{2}{\pi} arcsin(\rho)$	$2t_{v+1}\left(-\sqrt{v+1}\sqrt{\frac{1-\rho}{1+\rho}}\right)$

Table 2. Family of univariate and bivariate Archimedean copula functions.

#	Copula function	Generator function	Domain of parameters	Kendall's τ coefficient (2)	Upper/lower tail dependence
1	Clayton	$\frac{1}{\theta}(t^{-\theta}-1)$	$\theta > 0$	$\frac{\theta}{\theta+2}$	$(2^{\frac{-t}{\theta}}, 0)$
2	Gumbel	$(-\log t)^{\theta}$	$\theta \geq 1$	$1-\frac{1}{\theta}$	$(0, 2-2^{\frac{1}{\theta}})$
3	Frank	$-\log\left[\frac{e^{-\theta t}-1}{e^{-\theta}-1}\right]$	$\theta \in R\{0\}$	$1-\frac{4}{\theta}+4\frac{D_1(\theta)}{\theta}$	(0,0)
4	Joy	$-\log[1-(1-t)^{\theta}]$	$\theta > 1$	$1+\frac{4}{\theta^2}\int_0^1 t\log(t)(1-t)^{\frac{2(1-\theta)}{\theta}}.dt$	$(0, 2-2^{\frac{1}{\theta}})$
5	BB1	$(t^{-\theta}-1)^{\delta}$	$\theta \geq 0, \delta \geq 1$	$1-\frac{2}{\delta(\theta+2)}$	$(2^{\frac{1}{\theta\delta}}, 2-2^{\frac{1}{\delta}})$
6	BB6	$(-\log(1-(1-t)^{\theta})^{\delta}$	$\theta \geq 1, \delta \geq 1$	$1+\frac{4}{\theta\delta}\int_0^1 \{-\log(1-(1-t)^{\theta}) \times (1-t)(1-(1-t)^{-\theta})\}.dt$	$(0, 2-2^{\frac{1}{\theta\delta}})$
7	BB7	$(1-(1-t)^{\theta})^{-\delta}-1$	$\theta \geq 1, \delta > 0$	$1+\frac{4}{\theta\delta}\int_0^1 (-(1-(1-t)^{\theta})^{\delta+1} \times \frac{(1-(1-t)^{\theta})^{-\delta}-1}{(1-t)^{\theta-1}}).dt$	$(2^{-\frac{1}{\varepsilon}}, 2-2^{\frac{1}{\theta}})$
8	BB8	$-\log\left[\frac{1-(1-\delta t)^{\theta}}{1-(1-\delta)^{\theta}}\right]$	$\theta \geq 1, \delta \in (0,1]$	$1+\frac{4}{\theta\delta}\int_0^1 (-\log(\frac{(1-\delta t)^{\theta}-1}{(1-\delta)^{\theta}-1}) \times (1-t\delta)(1-(1-t\delta)^{-\theta})).dt$	(0,0)

Table 3. Summary of the characteristics of selected bivariate copula functions.

Copula function	Positive dependence	Negative dependence	Distribution tail symmetry	Lower tail dependence	Upper tail dependence
N	✓	✓	✓	-	-
t	✓	✓	✓	✓	✓
C	✓	-	-	✓	-
G	✓	-	-	-	✓
F	✓	✓	✓	-	-
J	✓	-	-	-	✓
RC	-	✓	-	-	-
RG	-	✓	-	-	-
RJ	-	✓	-	-	✓
BB1	✓	-	-	✓	✓
BB6	✓	-	-	✓	✓
BB7	✓	-	-	✓	✓
BB8	✓	-	-	✓	✓

Vine structures

Besides the attention paid to the bivariate copula functions for modeling the relationship between financial and economic variables during the past decades, the hierarchically-structured pair copulas have also been regarded by the researchers of finance and economics recently. Moreover, the application of copula to time-series models have been developed in the recent past. The main challenge hindering the multivariate modeling by means of copula functions is the identification of appropriate copula functions for a specific multivariate modeling task.

Standard multivariate copulas (*e.g.*, Gaussian, t-student, and Archimedean) suffer from the lack of flexibility for modeling the dependence structure between a large number of variables. Despite the superior performance of simple copulas over other dependence structure measurement methods, these copulas are still limited when large-dimension problems are encountered, where multivariate data exhibits a complex dependence structure in many cases. Standard multivariate copulas apply a structural constraint to the dependence structure. generalization of these models (*e.g.*, hierarchical model) leads to some improvements but usually complicates their structure and generates some limitations like the parametric limitations [14]. As an alternative method for modeling multivariate dependence structures, the development of pair copulas in the form of vine structures was first introduced by Joe (1996). It was further expanded into more detail by Bedford and Cooke (2001, 2002) and later on by Kurowicka and Cooke (2006). The vines are flexible graphical structures that use a series of pair copulas to describe multivariate joint distributions. In a key innovative approach to modeling multivariate dependence structures, Aas *et al.* (2009) presented a sequence of par copulas called C-Vine and D-Vine structures. These two vine structures are members of a wider family of vine structures known as R-Vine structures [1]. The vine structures determine, in the form of theoretical graphical models, which variables at which positions must be used for building pair copulas in such a way to come up with the best representation of the dependence structure and joint variation of the considered set of variables. In other words, vines are tools for labeling the limitations in the configuration of random variables in large-scale distributions. A vine structure V is composed of several interconnected trees $T_1, .. T_{n-1}$, where the branches of the j^{th} tree initiate the next tree, *i.e.* $(j+1)^{th}$ tree. A regular vine (R-vine) structure is the one where, for n variables, $E(V) = EU \dots UE_{n-1}$ represents the set of branches of the V such that (1) $V = \{T_1, \dots, T_{n-1}\}$ (a vine structure is a set of subsequent trees). T_1 is a tree with nodes $N_1 = \{1,2, \dots, n\}$ and branches E_1. For $i = 2, \dots, n-1$, T_1 is a tree with nodes $N_i = E_{i-1}$ (except for the very first tree, branches of the previous tree serve as a node to the next tree). (closeness rule) For each $\{a, b\} \epsilon E_i, \dots, n-1$ we will have $E_i \# a \Delta b$, where Δ and $\#$ denote symmetric difference and cardinal number of the set, respectively. Each branch in the tree T_j is a non-sorted pair of the nodes T_j on the same tree or, say, a non-sorted pair of the branches on the tree T_{j-1}. The order of each node on each tree T_j is equal to the number of branches connected to that node. An R-Vine structure is referred to as a conical vine (C-Vine) if it has a tree T_j with a unique node of the order $n-1$. A vine structure for which all nodes comprising the trees exhibit an order of no higher than 2 is referred to as a drawable vine (D-Vine) [15]. An R-Vine structure over n random variables includes $\frac{n(n-1)}{2}$ branches. In order to form the very first tree of the vine structure, we need to identify n – 1 unconditional bivariate copulas. Density of an R-Vine copula structure can be expressed by assigning appropriate bivariate copulas to the branches of the R-Vine structure [16].

From another point of view, a “vine copula” is a flexible graphical model for describing a multivariate copula structure using bivariate copulas called “pair-copula constructions (PCCs)”. A PCC decomposes a multivariate probability density to bivariate copulas. In this way, the vine copulas combine the advantages of multivariate copula modeling with the flexibility of bivariate copulas to come up with more efficient dependence modeling. The vine copulas enable proper modeling of the behavior of different variable pair structures. This implies that the dependence modeling takes into account the symmetry, dependence and tail dependence. This flexibility requires the selection of a good design model for realizing the potentials of the vine copulas in dependent models [17]. According to Aas *et al.* (2009), the following expression gives the joint multivariate density function for d random variables:

$$f(x_1, x_2, \dots, x_d) = f_d(x_d).f(x_{d-1}|x_d).f(x_{d-2}|x_{d-1}, x_d) \dots f(x_1|x_2, \dots, x_d) \tag{23}$$

Considering a trivariate case $X = (X_1, X_2, X_3)^T \sim F$ with the marginal distribution functions F_3, F_2, F_1 and the densities *f*, the recursive mode gives:

$$f(x_1,x_2,x_3)=f_1(x_1).f(x_2|x_1).f(x_3|x_1,x_2) \tag{24}$$

Application of the Sklar theorem gives:

$$f(x_1|x_2) = \frac{f(x_1,x_2)}{f_1(x_1)} = \frac{c_{12}(F_1(x_1),F_2(x_2)).f_1(x_1).f_2(x_2)}{f_1(x_1)}=c_{12}(F_1(x_1),F_2(x_2)).f_2(x_2) \tag{25}$$

And

$$f(x_3|x_1,x_2) = \frac{f(x_2,x_3|x_1)}{f(x_2,x_1)} = \frac{c_{231}(F(x_2|x_1),F(x_3|x_1)).f(x_2|x_1).f(x_3|x_1)}{f(x_2|x_1)} \tag{26}$$

$$= c_{231}(F(x_2|x_1),F(x_3|x_1)).f(x_3|x_1) \tag{27}$$

$$= c_{231}(F(x_2|x_1),F(x_3|x_1))c_{13}(F_1(x_1),F_3(x_3)).f_3(x_3) \tag{28}$$

The above equation can be rewritten as follows:

$$\begin{aligned} f(x_1,x_2,x_3) = {} & f_1(x_1).\overbrace{f_2(x_2)}^{\text{M arg inal}}.f_3(x_3) \times c_{12}(F_1(x_1),\overbrace{(F_2(x_2)}^{\text{Unconditional Pairs}},(F_3(x_3)) \\ & \times c_{23|1}(F(x_2\overbrace{|x_1),}^{\text{Conditional Pairs}}F(x_3|x_1)) \end{aligned} \tag{29}$$

Therefore, the 3D joint density of the above equation can be expressed by marginal density functions of bivariate copulas C12, C13, and C23|1 with the densities c12, c13, and c23|1, respectively, which are known as pair copulas. Since the decomposition in Eq. 29 is not a one-by-one procedure, numerous recursive PCCs exist. In order to classify these PCCs, Bedford and Cooke (201, 2002) introduced some graphical models called "regular vine copula".

In general, three choices must be determined before an R-Vine copula can be implemented: (1) vine structure selection, (2) pair copula selection, and (3) parameterization of the selected pair copulas.

Based on what was mentioned so far, it seems that the application of the vine copula approach to describe the structure of the dependency between different financial markets in Iran can end up with pretty reliable results. Accordingly, in this research, the dependence structure of Iranian financial markets is evaluated using the vine copula functions [17].

Research background

In the following, a brief review of the literature elaborating on the application of copula functions in the finance is presented [18].

Sklar (1959) was the first to propose the application of the copula functions for measuring nonlinear dependency between variables. Following his research, the use of copula functions in the research works expanded, making it an important highly applied method for building multivariate joint distributions and describing the dependence structure between variables. In finance, Amirtex et al. (1999) were the first to highlight potential applications of the copula functions. These applications were later properly classified by Cherubini et al. (2004). Huta and Paralo (2005) used various copula models with GARCH marginal distributions to calculate the value at risk (VaR), and further compared these models to conventional solutions for evaluating the VaR [1].

Jondi and Rekinger (2006) proposed the copula-GARCH model and used to obtain the dependence structure between stock exchange markets. Huang *et al.* (2009) utilized the conditional copula-GARCH models to estimate the portfolio VaR, and compared the copulas to other classical methods. The results showed that the t-copulas with GARCH as marginal distribution function could predict the VaR more efficiently than the other copula models and classic approaches considered in this research, not to mention their superior performance in describing the dependence structure of the portfolio of assets [18].

In their study, Wang *et al.* used the EVT-copula-GARCH model for optimizing a FOREX portfolio and concluded that the t-copula and Clayton copula tend to offer better description of the dependence structure between the assets in the portfolio. They further found that the t-copula performs better when estimating the VaR. Later on, Wang et al. (2011) investigated the dependence between the Chinese stock market and other international stock markets around the world. In their study, they used the Gaussian, Clayton, Gumbel, and SJC copulas, where the marginal distribution was modeled using the GJR-GARCH(1,1) [19].

In the followings, a summary of the research on the vine functions is presented.

Scogland et al. (2013) explained hybrid methods for using the copula functions for risk compilation. They showed that the hybrid copula functions for risk compilation enable one to compile the risk using the minimum data availability, although inadequate data availability prevents accurate determination of dependence structure and joint distribution of the risks [20].

Bridgman and Sezader (2013) presented a solid study where they elaborated on the application of vine-copula functions to portfolio management. They used the C-Vine and R-Vine copulas with the daily data over a period of 4 successive years (2006 – 2009) to model the dependence structure between the returns of 50 top listed companies in the European Union. They further forecasted the VaR of the return of these companies for two coming years (2010 and 2010) and determined an optimum portfolio of stocks using the average variance method for the two years. Perichman and Sezado compared the results of VaR estimation of the return on asset and optimum stock portfolio determination among different methods, including C-Vine and R-Vine copulas, normal copula functions, t-student copula function, and the DCC-GARCH method. Results of this comparison showed that the R-Vine copula method tended to produce more accurate results, as compared to other methods [21].

Gojan and Joyad (2012) used pair copula models to compile market risks. They utilized the daily data during 1999 – 2009 to estimate the associated risks with the stock market, FOREX market, development, and interest rate – a set of risks that significantly contribute to the market risk. They then calculated the required economic capital to cover probable loss and expected loss/benefit due to the risk diversity. The economic capital requirement calculation was based on the Solvency II standard and the Swiss Solvency Test (SST) or calculation of VaR 99.5% and ES 99.5% [22].

In a research performed by Sookcharon and Lisam (2017), application of vine-copula functions to estimate multiple-period risk coverage ratio for covering undesired risk of refinery companies was studied. For this purpose, they considered the current and future prices of the WTI crude oil, gasoil, and fuel oil during the 1986 – 2015 period. They believed that the main advantage of the vine-copula functions is their ability to express unique characteristics of petroleum products, such as skewness and fat-tail marginal distribution of each petroleum product as well as the diverse tail dependence between these products, which could be clearly explained by the vine-copula functions [23]. In another piece of research, McAleer, Powell and Singh (2017) utilized an R-Vine structure to study mutual dependencies across the large European finance market in the form of individual indices for each market and the combined index of EURO STOXX 50 and the Dow Jones Industrial Average. Results of this study showed that the dependence between financial markets varies in a complex way [1]. Zhang *et al.* (2014) modeled the dependence structure between stock markets during 2006 – 2013. For this purpose, they applied D-Vine, R-Vine, and C-Vine approaches onto the data related to 10 international stocks indices. Then, the Monte Carlo simulation was practiced to estimate the VaR and CVaR values for equivalent portfolio of stocks based on the estimated dependence structure using the copula functions and vine-copula structures, with the accuracy of the estimations evaluated using different Christopherson statistics. Reburdo and Aguilini (2015) studied multivariate dependence structure between four precious metals (gold, silver, platinum, and palladium) to check for desired and undesired price overflows using the vine-copula constructions and calculate desired and undesired VaRs. According to the findings of these researchers, different metals exhibited different dependence structures, with each of them showing a particular tail and average dependencies. Based on the results of this research, some evidence indicating desired and undesired price overflows between the precious metals was observed, with different magnitudes and significance for different metals [24]. Riyadh Alvi and Muhammed Bin-Ayesha (2016) followed a vine-copula approach to investigate the dynamic relationship between energy, stock prices, and FOREX rate. Using a mixed sample of 10-year return on crude oil. Dow Jones index, and trade-weighted US dollar index, these researchers achieved some evidence indicating a significant and symmetric association between these three assets. Moreover, splitting the entire study period to two segments, namely before and after the financial crisis, they figured out that these three assets do not follow a constant dependency structure throughout time and that the dependence structure has been affected by the 2007 – 2009 financial crisis. Berkhman and Sezado (2013) combined the GARCH family of models to describe marginal distributions using R-Vine copulas and present a factor model for explaining the return on assets. Designated as R-Vine market segment (RVMS) model by the authors, this model was inspired by the CAPM model and offered particular advantages over the CAVA model presented by Hinen and Valdsogov (2009) [1]. Bin Sida (2017) used the symmetric copula functions of Joe and Clayton, which could identify right and left tail dependencies, as well as the C-Vine and R-Vine structures to model the Markov's switching process. He then implemented his model on 12 government-issued bonds, including those issued by the US government and 11 European states. The results showed that the switching-regime copula models can outperform the fixed-regime copulas in reflecting the dynamicity of the data dependency [26]. In their research, Berkhman, Sezado, and Patrelini (2014) used the data of 33 banks and 188 financial institutions in Italy during 2003 – 2011 (extracted from the database italiano perdite operative or DIPO) considering a family of pair copula functions and vine-copula structures. Accordingly, they succeeded to model the dependence between 7 operative events and 8 operative lines while estimating the required capital for covering simultaneous occurrence of these events. Based on the findings of this research, the required capital for covering the operative risks faced by considered financial institutions was on average 38% lower than the baseline set by the Basel committee according to the Basel II. Nicolopolus, Joe and Lee (2010) showed that the vine-copulas constructed based on bivariate t-student copula functions offer a good fit to the data referring to multivariate return on financial assets. Nevertheless, it seems that the return dependencies are stronger for the left tail of the joint distribution of the assets, rather than its right tail. In this study, the presence of such an asymmetry was investigated by using the copulas reflecting the asymmetry of the left and right sides of the distribution. Lu, Alkok, and Brilsford (2013) used asymmetric and elliptic multivariate copula models to predict returns of portfolios of 3 to 12 assets [18]. In this research, assuming no short-selling constraint and a utility function that was characterized by minimizing the conditional VaR, the authors investigated the efficiency threshold obtained by different models and performed a pair-wise comparison of different methods to incorporate scalable asymmetric dependency into the return on asset using the Archimedean copula of Clayton in in-sample and out-sample modes. Based on their findings, the asymmetry modeling of the marginal distributions and dependency modeling using the Archimedean copula of Clayton and C-Vine structure in a well-coordinated way tend to bring about the best possible results in terms of economic and statistic measures, as compared to the models where symmetric and elliptical dependency structures are mixed.

A summary of the relevant literature in Iran is presented in the following.

Keshavarz Hadad and Heirani (2014) published a study entitled "Estimation of Value at Risk in the Presence of Dependence Structure in Financial Returns: A Copula Based Approach" where they investigated the dependence structure between two price indices, namely those referring to chemical products and pharmaceutical companies listed on Tehran Stock Exchange (TSE) during 2004 – 2012. The results indicated an asymmetric dependence structure among the considered variables, with the findings further showing the accuracy and adequacy of the copula-based approach compared to conventional models for predicting the VaR [27].

In a study, Musavi *et al.* (2013) estimated the VaR of selected stocks using conditional GARCH-copula. Results of this research showed that the Gaussian copula model with normal marginal distribution and the Gaussian copula with t-student marginal distribution outperformed the historical simulation as well as variance-covariance methods for estimating the VaR [28]. In 2015, Barghi Osguei *et al.* investigated nonlinear effects of variations of actual exchange rate and crude prices on the price index of TSE following an approach based on the Markov – switching regimes. Results of this research showed that variations of the actual exchange rate and crude oil prices, as exogenous factors, impose a one-lag delayed yet significant positive effect on the stock prices. Considering two lags, however, the effects were found to be negative but still significant [18]. In a study, Pishbahar and Abedi (2017) calculated the VaR of a portfolio of food products using a copula-based approach. The results showed that the copula (*i.e.*, copula functions) – based approach tended to produce more reliable results than alternative methods (historical simulation, multivariate normal distribution, and multivariate t-distribution). On this basis, the maximum weekly expected loss on a portfolio of diaries was estimated at 2.10%. In another piece of research, Bordbar and Heydari (2017) considered the relationship between fluctuations in the oil price and return on stocks of Base Metals, petroleum products, and chemicals using vector autoregression (VAR) and multivariate generalized autoregressive conditional heteroskedasticity (MGARCH) models. They concluded that despite average dependence effects between the markets of oil, Base Metals, and petroleum products, such effects are absent when it comes to the market of chemical industries. Indeed, fluctuations in the world oil market were found to impose no effect on those in the markets of the chemical industries and Base Metals. However, a significant yet negative association was found between the oil market variations and return on the stocks of petroleum products [29].

Research methodology

Data and research scope

This research has been developed in the paradigm of quantitative research works, where decision-making is based on qualitative data. Since the nature of the research problem calls for no specific hypothesis, the present study is focused on a question regarding the structure and characteristics of the dependence among different elements of the financial system of TSE. The considered statistical population includes the indices measuring the market and active industries on the TSE, which are supposed to represent the behavior and variations of different elements of the financial system of the TSE. In this work, daily data on 31 active industries on the TSE and 10 market indices was retrieved from the "website of the TSE" for the period 2001 – 2020. The following tables present a list of the considered indices in this study.

Table 4. Contents of Market indices

Market indices		
Free-Float Index	Total Equal Weight Price	Overall Index (TEDPIX)
Main Board Index	Top 50 Index	Total Equal Weighed
Secondary Index	Total Price Index (TEPIX)	TSE-50 Index (top 50 most active companies)
Industry index		

Table 5. Contents of Industrial sector indices

Industrial sector indices		
Agriculture And Crop Activities	Electrical Equipment	Investment
Coal Mining	Vehicle	Banking
Metallic Mineral Products	Sugar	Other Finance (leasing)
Textile	Multi-discipline	Land Transportation & Storage
Wood Products	Foodstuff (excl. sugar)	Telecommunication
Paper Products	Manufacture of Pharmaceuticals	Building Construction & Mass Housing
Oil Products	Chemical Products	Computer
Rubber & Plastic Product	Tiles and Ceramics	Technical-Engineering
Base Metals	Cement	Petroleum Production (excl. exploration)
Metal Products	Non-metal Mining	Insurance and Pension
Machineries		

* Noteworthily, some industrial sector indices were not considered in this study due to their short history of data reporting and negligible market caps.

Considering the large number and wide diversity of the active industries on the TSE and in an attempt to come up with a more accurate investigation of the dependence structures, we began by clustering all industrial indices based on three scenarios, with the dependence modeling performed separately under different scenarios. The scenarios considered in this research included clustering based on the field of activity in the macro economy (resulting in the following clusters: oil and gas industries, mineral and metallurgical industries, industrial machinery and Land Transportation & Storage, food and pharmaceutical industries, technological industries, construction, and finance/investment-related industries), clustering based on foreign policies and JCPOA (in the following clusters: JCPOA-affected and import-based industries, export-oriented industries that are affected by the world rates, domestic-oriented inflation-affected industries, and finance/investment industries), and clustering based on the market caps of different industries. All calculations and modeling efforts in this research were made using the R software package.

Research methodology

As a first step, logarithmic or continuous return of the indices was evaluated using the following equation:

$$R_t = ln\frac{P_i(t)}{P_i(t-1)} \tag{30}$$

Then, in order to evaluate the considered data statistically, descriptive statistics (*i.e.*, mean, median, standard deviation (SD), trimmed mean (TM), median absolute deviation (MAD), minimum, maximum, range, skewness, kurtosis, and standard error (SE)) were investigated.

Given the requirement of U-form input data for using the copula functions, the next step was to apply the GARCH (1.1) model (calculated from the returns) to obtain residual values as input to the final model.

Subsequently, in order to explore the dependence structures, market indices were separately investigated in a single cluster while the industrial sector indices were evaluated based on the results of clustering under different scenarios. In this stage, a graph of the dependence structure was drawn independently for each cluster, ending up with a dependence matrix.

Following the study, a representative (i.e., central) index was selected in each cluster using the C-Vine functions. Then, drawing the C-Vine structure and fitting different families of copulas to the data, the best copula family for explaining the dependency of different elements of each cluster was identified, with the relevant parameters determined. For each cluster, all possible C-Vine structures up to the second level were drawn to come up with the results.

At a next step, a new matrix containing representatives of different clusters was formed and the dependency of the representatives was calculated by the Kendall rank correlation coefficient (*i.e.*, Kendall’s tau).

Subsequently, R-Vine functions in combination with all families of pair copulas were used to draw the dependence structure for the new matrix, ending up with a model of association of different elements based on the best-fit copula function for each cluster. Noteworthily, in this stage, the R-Vine model was built based on the strongest-correlation associations between the elements. This iterative process was continued until the last stage where all elements in both groups of indices were included (Model 1).

At the next step, dependence structure was investigated with the R-Vine model coupled with the Gaussian and t-student copulas following a procedure that was similar to the previous step (Models 2 and 3).

In order to investigate the dependence structure in more detail, for all cluster-representing matrixes, dependence fitting was conducted based on all families of copula functions, including Gaussian and t-student copulas, but this time with C-Vine copulas in full mode, with the rest of the procedure resembling that of the previous steps (Models 4 to 6).

Finally, an ultimate matrix was constructed to compare the 6 models. At a final step, Wang test was used to perform pairwise comparisons and identify the best model for explaining the dependence structure between the market and industry indices under the three considered scenarios separately.

A summary of different steps of this research is demonstrated in the following figure:

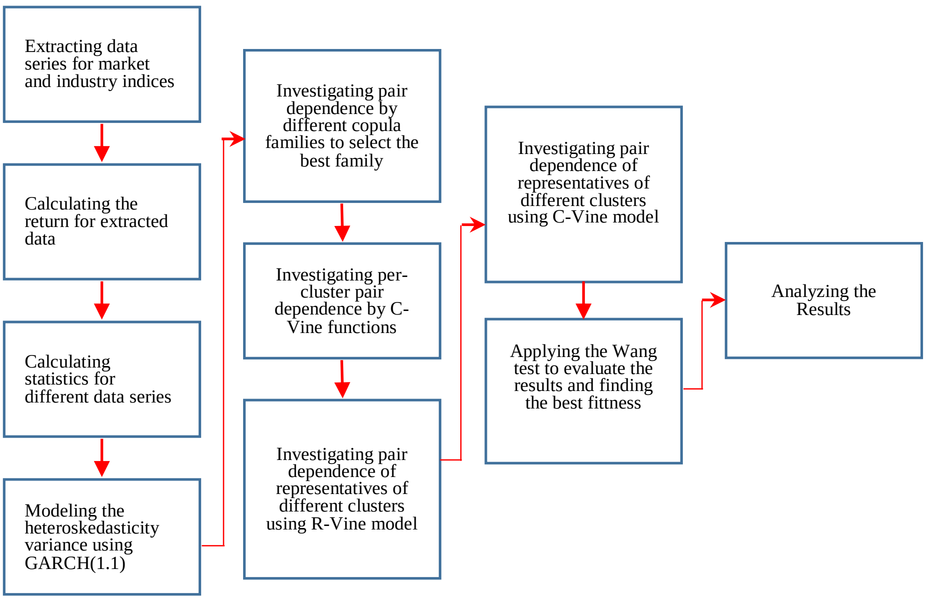


Fig. 1. Different steps of the analysis performed in this study.

Models used in this research

GARCH(1.1)

The GARCH(1.1) is the simplest generalized autoregressive conditional heteroskedasticity model. According to this model, the best predictor of the variance in the proceeding period has contributions from the weighted average of the long-run variance, current-period predicted variance, and new information of the current period, which is obtained as the square of the latest residual, as follows [29]:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \alpha_{t-1}^2 + \beta \sigma_{t-1}^2 \qquad \text{u\~N}(0,\sigma^2) \tag{31}$$

$$\alpha_\circ \geq 0, \alpha_1, \beta \geq 0, \alpha_1 + \beta < 1 \tag{32}$$

where σ_t^2 is predicted variance for the period t, u_{t-1}^2 is the square of residual (error term) for the period $t-1$, σ_{t-1}^2 is the predicted variance for the period $t-1$, and α_0, α_1, and β are the model parameters and used to predict future periods (*i.e.,* indicating average values). Information on the fluctuations during the previous period (measured as square of the latest residual) is known as ARCH, with the latest predicted variance called GARCH. The general model of GARCH is expressed as GARCH(q, p) where p and q refer to the ranks of the ARCH and GARCH, respectively. The following equation shows how these variables are related to one another:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_1 u_{t-1}^2 + \sum_{i=1}^{g} \beta_j \sigma_{t-1}^2 \tag{33}$$

Vine copula functions

In this research, we use two particular types of regular copulas called C-Vine and R-Vine [12]. In the C-Vine model, there is a key variable for the dependence structure, which is positioned at the center of the structure. This implies that all other models are directly associated with this variable. In general, three choices must be determined before an R-Vine copula can be implemented: (1) vine structure selection, (2) pair copula selection, and (3) parameterization of the selected pair copulas.

Regular vine joints

In this subsection, we elaborate on the theory of regular vine (R-Vine) copulas, how is an R-Vine matrix developed, and how imply can such a matrix represent the density. A review of the studies published by Bedford and Cooke (2001, 2002), and Kurowicka and Cooke (2006), a vine is a graph with two nodes connected to one another through a unique link.
The vine $v = (T_1, \dots, T_{n-1})$ has *n* components if:

1. T_1 is a tree with the nodes $N_1 = \{1, \dots\dots\dots, n\}$ and edge E_1.
2. For $i = 2, \dots, n-1$, T_1 is a tree with the nodes $N_i = E_{i-1}$ and edge E_i.
3. For $i = 2, \dots, n-1$ and $\{a, b\} \in E_i$ where $a = \{a_1, a_2\}$ and $b = \{b_1, b_2\}$, the condition $(a \cap b) = 1$ must be satisfied.

In other words, an n-component R-Vine copula is a nested set of $n-1$ trees in such a way that the edges of the j^{th} tree are nodes to the $(j+1)^{\text{th}}$ tree. The initial condition for

connectedness of two nodes on the $(j+1)^{\text{th}}$ tree is that these two nodes share the same node on the j^{th} tree. It is worth noting that the set of nodes on the first tree includes all members (*i.e.*, $1, \dots, n$), while the set of edges is a set of $n-1$ pair of these members. On the second tree, the set of nodes includes the member pairs while the set of edges is formed by pairing the members. The complete union of an edge refers to a set of edges corresponding to all members. Letting the nodes a and b be connected to one another through an edge, the conditioned and conditioning sets for this edge exhibit symmetric and cross differences to the complete union of a and b [30].
The following diagram shows a sample R-Vine for seven variables.

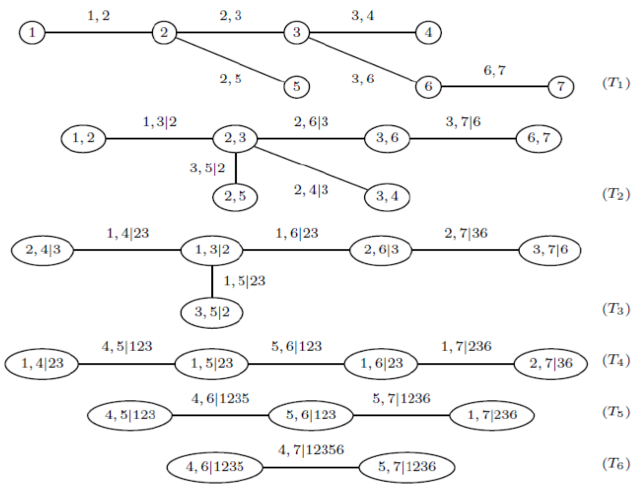


Fig. 2. sample R-Vine for seven variables

Canonical vine joints

As mentioned earlier, a C-Vine structure is a special case of an R-Vine structure. On the first tree of a C-Vine structure, the dependency is modeled for each pair based on a particular variable called the *root node* using bivariate copulas. Next, upon conditioning, the root node of the first tree, the dependency of copula pairs is modeled based on the second root node. In general, a root node is selected on each tree and all copula pairs are modeled based on this root node and conditioning of the previous root node. A C-Vine tree exhibits a star-like structure, as shown in the figure, with its density expressed as follows:

$$f(x) = \prod_{k=1}^{d} f_k(x_k) \times \prod_{i=1}^{d-1} \prod_{i=1}^{d-i} c_{i,i+j|l:(i-1)}\left(F(x_i|x_1, \dots\dots x_{i-1}), F\left(x_{i+j}|x_1, \dots\dots x_{i-1}\right)\middle|\theta_{i,i+j|l:(i-1)}\right) \tag{34}$$

in which f_k $(k = 1, \dots, d)$ indicates the marginal densities and $c_{i,i+j|l:(i-1)}$ denotes the density of bivariate copulas with the parameters $\theta_{i,i+j|l:(i-1)}$ [4].

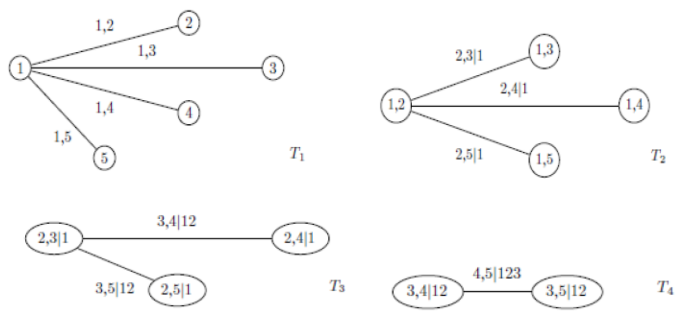


Fig. 3. An example of a C-Vine corresponding to five variables.

Kendall's tau

The Kendall rank correlation coefficient (*i.e.*, Kendall's tau) is a good alternative to linear correlation coefficient. The Kendall's tau for the two variables X and Y is defined as follows:

$$\rho_\tau(X,Y) = 4\int_0^1\int_0^1 C(u,v)dC\,(u,v) - 1 \tag{35}$$

where $C(u,v)$ is the joint distribution of the two variables X and Y. For the Gaussian, t, and other implicit copulas, the relationship between the linear correlation coefficient and Kendall's tau can be expressed as follows:

$$cor(X,Y) = sin\left(\frac{\pi}{2}\rho_\tau\right) \tag{36}$$

Vuong's closeness test (1989)

This test is useful for comparing two non-interfering models. Similar to Clarke's test, this test is based on the likelihood ratio in relation to the "Kullback–Leibler index" and measures the divergence between two statistical models. Letting c_1 and c_2 be two density functions corresponding to competing bivariate copulas with estimated parameters of B_1 and B_2, one can perform the Vuong's closeness test by calculating the standardized sum of the likelihood logarithm difference over all data points. Likelihood logarithm difference between the observations $i = 1,\ldots,N$ and $j = 1, 2, \ldots, u_{i,j}$ is written as follows [14]:

$$m_i = log\left(c_1\left(u_{i,1},u_{i,2}\middle|\hat{B}_1\right)\right) - log\left(c_2\left(u_{i,1},u_{i,2}\middle|\hat{B}_2\right)\right) \tag{37}$$

The point-by-point standardized sum of the likelihood logarithm difference, v, can be expressed as follows:

$$v = \frac{\frac{1}{n}\sum_{i=1}^{N} m_i}{\sqrt{\sum_{i=1}^{N}(m_i - \bar{m})}^2} \tag{38}$$

Vuong showed that, asymptotically, v exhibits a normal distribution. If $v > -\Phi^{-1}\left(1-\frac{\alpha}{2}\right)$, then the copula model 1 is preferred over the copula model 2 at an error level of α. Conversely, if $v < -\Phi^{-1}\left(1-\frac{\alpha}{2}\right)$, then the copula model 2 is preferred over the copula model 1. Finally, if $|v| \le -\Phi^{-1}\left(1-\frac{\alpha}{2}\right)$, then no definite decision can be made between the two models and the null hypothesis (*i.e.*, statistical equivalence of the models) cannot be rejected.

Research results

The following table shows the results in terms of descriptive statistics of the considered data for explaining the dependence structure between the market indices and active industry sector indices on the TSE.

Table 6. Descriptive statistics of the considered data.

Index	Mean	SD	Median	Trimmed mean	MAD	Minimum	Maximum	Range	Skewness	Kurtosis	SE
Free-Float Index	-0.002	0.137	0.003	0.005	0.018	-3.104	0.075	3.179	-21.982	494.466	0.006
Main Board Index	-0.001	0.120	0.003	0.004	0.018	-2.702	0.075	2.777	-21.722	486.663	0.005
Secondary Index	-0.003	0.161	0.003	0.004	0.017	-3.649	0.075	3.725	-22.212	501.397	0.007

Industry index	-0.001	0.128	0.003	0.004	0.018	-2.887	0.075	2.962	-21.855	490.651	0.006
Total Equal Weight Price	0.002	0.046	0.004	0.005	0.018	-0.941	0.075	1.017	-15.913	320.699	0.002
Top 50 Index	-0.003	0.167	0.003	0.005	0.018	-3.806	0.075	3.881	-22.258	502.787	0.007
Total Price Index (TEPIX)	0.001	0.079	0.003	0.004	0.018	-1.747	0.075	1.823	-20.337	445.612	0.003
Overall Index (TEDPIX)	-0.002	0.132	0.003	0.004	0.018	-2.991	0.075	3.066	-21.920	492.605	0.006
Total Equal Weighed	0.002	0.061	0.004	0.005	0.017	-1.301	0.075	1.376	-18.686	397.793	0.003
TSE-50 index (top 50 most active companies)	0.004	0.023	0.003	0.005	0.018	-0.060	0.158	0.218	0.438	3.721	0.001
Agriculture And Crop Activities	0.003	0.034	0.004	0.005	0.020	-0.574	0.075	0.650	-9.565	161.607	0.001
Coal Mining	0.006	0.041	0.003	0.005	0.021	-0.060	0.767	0.827	12.521	233.471	0.002
Metallic Mineral Products	0.001	0.084	0.003	0.004	0.018	-1.856	0.075	1.932	-20.557	452.064	0.004
Textile	0.007	0.077	0.003	0.004	0.018	-0.060	1.712	1.772	20.196	441.644	0.003
Wood Products	-0.001	0.126	0.003	0.004	0.021	-2.833	0.225	3.058	-21.433	478.347	0.005
Paper Products	0.002	0.047	0.003	0.004	0.021	-0.934	0.081	1.015	-14.737	289.143	0.002
Oil Products	-0.004	0.194	0.003	0.005	0.020	-4.426	0.075	4.501	-22.380	506.473	0.008
Rubber & Plastic Product	0.002	0.062	0.003	0.005	0.019	-1.327	0.075	1.402	-18.712	398.569	0.003
Base Metals	-0.001	0.119	0.004	0.005	0.019	-2.675	0.075	2.750	-21.678	485.337	0.005
Metal Products	0.002	0.059	0.004	0.005	0.019	-1.247	0.075	1.322	-18.235	384.993	0.003
Machineries	0.002	0.057	0.003	0.005	0.018	-1.195	0.075	1.270	-18.018	378.869	0.002
Electrical Equipment	-0.004	0.186	0.003	0.004	0.018	-4.240	0.075	4.316	-22.355	505.719	0.008
Vehicle	0.002	0.040	0.003	0.004	0.019	-0.766	0.075	0.841	-13.259	250.964	0.002
Sugar	0.003	0.044	0.003	0.005	0.018	-0.880	0.075	0.955	-15.017	296.684	0.002
Multi-discipline	0.003	0.038	0.003	0.004	0.018	-0.709	0.075	0.784	-12.613	234.752	0.002
Foodstuff (excl. sugar)	0.005	0.029	0.003	0.005	0.018	-0.060	0.445	0.505	6.601	100.222	0.001
Manufacture of Pharmaceuticals	0.003	0.029	0.004	0.005	0.018	-0.442	0.075	0.517	-6.868	103.162	0.001
Chemical Products	0.004	0.024	0.003	0.004	0.018	-0.221	0.075	0.297	-1.662	14.436	0.001
Tiles and Ceramics	0.006	0.046	0.004	0.005	0.019	-0.060	0.930	0.990	15.504	310.610	0.002
Cement	0.008	0.099	0.004	0.005	0.019	-0.060	2.213	2.273	21.171	470.316	0.004
Non-metal Mining	0.005	0.033	0.003	0.005	0.019	-0.060	0.573	0.633	9.601	164.370	0.001
Investment	0.006	0.051	0.003	0.004	0.018	-0.060	1.076	1.136	17.057	352.604	0.002
Banking	0.008	0.101	0.004	0.004	0.020	-0.060	2.277	2.337	21.242	472.417	0.004
Other Finance (leasing)	0.006	0.061	0.003	0.004	0.019	-0.060	1.316	1.376	18.579	395.085	0.003
Land Transportation & Storage	0.005	0.034	0.003	0.005	0.018	-0.060	0.606	0.666	10.242	178.973	0.001
Telecommunication	0.007	0.083	0.002	0.004	0.017	-0.060	1.854	1.914	20.594	453.263	0.004
Building Construction & Mass Housing	0.009	0.104	0.003	0.005	0.019	-0.060	2.348	2.408	21.325	474.874	0.005
Computer	0.003	0.037	0.002	0.004	0.017	-0.691	0.075	0.766	-12.379	228.987	0.002
Technical-Engineering	0.008	0.095	0.003	0.004	0.020	-0.060	2.131	2.191	20.983	464.734	0.004
Petroleum Production (excl. exploration)	0.008	0.103	0.003	0.004	0.020	-0.060	2.321	2.382	21.241	472.367	0.005
Insurance and Pension	0.005	0.029	0.004	0.005	0.019	-0.060	0.433	0.493	6.228	92.793	0.001

As is evident in the above table, the distribution of the returns corresponding to most of the indices is skewed left with higher than normal kurtosis values. According to previous studies, it has been figured out that the distribution of financial data may not follow a normal distribution function. In this research, thanks to the copula functions, no assumption was made regarding the data distribution.

Once finished with standardizing the residuals according to the GARCH(1.1) model, the standardized values were used as a foundation for estimating the vine-copula structures. Pair copula functions were developed between each pair of variables based on the AIC-to-BIC ratio, and the vine structure describing the joint distribution of time-series was obtained using the Kendall rank correlation coefficient. Based on this, one can identify the best-fit C-Vine structure.

At the next step, a representative index was obtained for each cluster of the market and industrial sector indices under various scenarios using the C-Vine copula by fitting different families of the copula functions. A summary of the results of fitting for different clusters and different scenarios is reported in the following table:

Table 7. Results of applying C-Vine function for determining per-cluster representative index.

Results of applying C-Vine function for determining per-cluster representative index	**Representative index**	**Loglik**	**AIC**	**BIC**
Market cluster	TEDPIX	23,976.08	-47,802.16	-47,482.12
Scenario 1 – industrial clusters				

Results of applying C-Vine function for determining per-cluster representative index	Representative index	Loglik	AIC	BIC
Cluster 1: Oil and gas	Chemical products	1,430.49	-2,848.98	-2,823.38
Cluster 2: Mining and metallurgy	Mining	4,281.17	-8,540.34	-8,493.40
Cluster 3: Industrial and transportation	Land Transportation & Storage	7,145.21	-14,232.41	-14,108.66
Cluster 4: Food products and pharmaceuticals	Food products (excl. sugar)	12,231.77	-24,355.54	-24,125.11
Cluster 5: Technology	Computer	1,001.34	-1,998.68	-1,990.14
Cluster 6: Construction	Cement	2,914.78	-5,819.56	-5,798.23
Cluster 7: Finance and investment	Investment	4738.43	-9,456.86	-9,414.19
Scenario 2 - industrial clusters				
Cluster 1: Import-oriented and JCPOA-affected	Technical and engineering	9,235.85	-18,369.70	-18,152.07
Cluster 2: Export-oriented and world prices-affected	Multi-discipline	5,638.97	-11,243.94	-11,171.40
Cluster 3: IRR-based and domestic inflation-driven	Land Transportation & Storage	32,139.41	-63,824.82	-62,856.16
Cluster 4: Finance and investment	Investment	1,634.09	-3,266.17	-3,261.91
Scenario 3 - industrial clusters				
Cluster 1: market cap above 400 TB IRR	Multi-discipline	5,015.26	-10,012.51	-9,974.11
Cluster 2: market cap below 400 TB IRR	Investments	4,832.38	-9,626.75	-9,545.68
Cluster 3: market cap below 100 TB IRR	Food products (excl. sugar)	1,595.90	3,179.80	-3,154.20
Cluster 4: market cap below 60 TB IRR	Land Transportation & Storage	3,157.08	-6,292.15	-6,245.21
Cluster 5: market cap below 30TB IRR	Rubber & Plastic Product	5,882.62	11,731.23	-11,658.69
Cluster 6: market cap below 15 TB IRR	Metal products	5,840.82	-11,647.64	-11,575.10
Cluster 7: market cap below 5 TB IRR	Oil production (excl. exploration)	6,269.60	-12,501.19	-12,420.11

Considering the above tables, it is evident that, in the cluster containing the market indices, the TEDPIX was expectedly identified as the representative of the index. That is, this index exhibits the strongest dependency with other market indices.

Representative industrial sector indices were selected based on the values reported in the table. For example, under Scenario 1, where the clustering was conducted based on the field of activity of different industries, the chemical products index was identified as representative of the oil and gas sector. In the finance and investment sector, the investment industry was identified as representative with the strongest dependency on other indices in the same sector. The same procedure was applied to other sectors to identify the representatives. Considering the results reported in the above table and the identified indices representing the market and industrial sectors under the three scenarios, three final clusters were selected for evaluating the dependence structure between the industry and market indices. Different components of these three clusters are listed in the following table:

Table 8. Representatives of industrial-sector indices.

Description	Representative of market cluster	Representatives of industrial-sector indices

Final cluster under scenario 1	TEDPIX	Chemical products, Mining, Land Transportation & Storage, Food products (excl. sugar), Computer, Cement, Investment
Final cluster under scenario 2	TEDPIX	Technical and engineering, Multi-discipline, Land Transportation & Storage, Investment
Final cluster under scenario 3	TEDPIX	Multi-discipline, Investments, Food products (excl. sugar), Land Transportation & Storage , Rubber & Plastic Product, Metal products, Oil production (excl. exploration)

Following this study, dependence structure was investigated between different components of each cluster separately using the R-Vine and C-Vine functions. Moreover, for the sake of fitting and comparing the results from different copula functions when applying each of the mentioned vine-copula models, different families of copula functions were considered in three schemes, namely considering all copula function families, considering Gaussian copulas, and, finally, considering t-student copula functions. To sum up, the following cases were considered:

Scenario 1

1.1. Scenario 1, modeling based on R-Vine functions and all families of copula functions
1.2. Scenario 1, modeling based on R-Vine functions and Gaussian copula functions
1.3. Scenario 1, modeling based on R-Vine functions and t-student copula functions
1.4. Scenario 1, modeling based on C-Vine functions and all families of copula functions
1.5. Scenario 1, modeling based on C-Vine functions and Gaussian copula functions
1.6. Scenario 1, modeling based on C-Vine functions and t-student copula functions

Scenario 2

2.1. Scenario 2, modeling based on R-Vine functions and all families of copula functions
2.2. Scenario 2, modeling based on R-Vine functions and Gaussian copula functions
2.3. Scenario 2, modeling based on R-Vine functions and t-student copula functions
2.4. Scenario 2, modeling based on C-Vine functions and all families of copula functions
2.5. Scenario 2, modeling based on C-Vine functions and Gaussian copula functions
2.6. Scenario 2, modeling based on C-Vine functions and t-student copula functions

Scenario 3

3.1. Scenario 3, modeling based on R-Vine functions and all families of copula functions
3.2. Scenario 3, modeling based on R-Vine functions and Gaussian copula functions
3.3. Scenario 3, modeling based on R-Vine functions and t-student copula functions
3.4. Scenario 3, modeling based on C-Vine functions and all families of copula functions
3.5. Scenario 3, modeling based on C-Vine functions and Gaussian copula functions
3.6. Scenario 3, modeling based on C-Vine functions and t-student copula functions

The modeling procedure for Case 1.1 is described in the following.

Firstly, the dependence of different components of a cluster is determined as Kendall's tau.

Table 9. Kendall rank correlation coefficient matrix for the Case 1.1

Index	TEDPIX	Metallic Mineral Products	Food products (excl. sugar)	Chemical	Cement	Investment	Land Transportation & Storage	Computer
TEDPIX	1	0.945	0.872	0.924	0.886	0.915	0.892	0.861
Metallic Mineral Products	0.945	1	0.866	0.903	0.877	0.907	0.896	0.86
Food products (excl. sugar)	0.872	0.866	1	0.887	0.859	0.888	0.892	0.855
Chemical	0.924	0.903	0.887	1	0.849	0.898	0.893	0.871
Cement	0.886	0.877	0.859	0.849	1	0.872	0.87	0.808

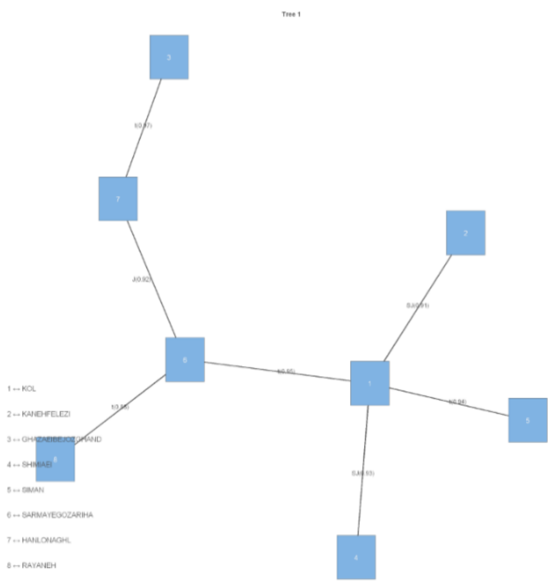
Investment	0.915	0.907	0.888	0.898	0.872	1	0.912	0.878
Land Transportation & Storage	0.892	0.896	0.892	0.893	0.87	0.912	1	0.848
Computer	0.861	0.86	0.855	0.871	0.808	0.878	0.848	1

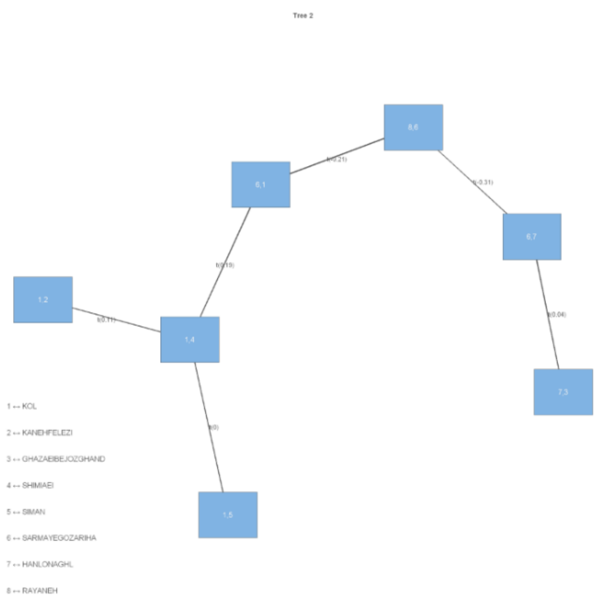
Next, based on the R-Vine structure and considering all copula function families, the best function for expressing different dependencies among the components of the considered cluster was obtained, as reported in the following table.

Table 10. the best function for expressing different dependencies among the components.

tree	edge	family	cop	par	par2	tau	utd	ltd
1	1,5	2	t	1	2	0.94	0.94	0.94
	1,2	16	SJ	21.64	0	0.91	-	0.97
	1,4	16	SJ	27.47	0	0.93	-	0.97
	6,1	2	t	1	2	0.95	0.95	0.95
	7,3	2	t	1	2	0.97	0.97	0.97
	6,7	6	J	24.64	0	0.92	0.97	-
	8,6	2	t	0.98	2	0.88	0.88	0.88
2	4,5;1	2	t	0	16.97	0	0	0
	4,2;1	2	t	0.17	15.09	0.11	0	0
	6,4;1	2	t	0.3	16.31	0.19	0.01	0.01
	8,1;6	2	t	-0.33	3.46	-0.21	0.04	0.04
	6,3;7	2	t	0.07	5.12	0.04	0.06	0.06
	8,7;6	2	t	-0.46	5.1	-0.31	0.01	0.01
3	2,5;4,1	2	t	-0.27	4.51	-0.17	0.02	0.02
	6,2;4,1	2	t	0.54	3.31	0.36	0.31	0.31
	8,4;6,1	1	N	0.02	0	0.01	-	-
	7,1;8,6	2	t	-0.39	4.91	-0.26	0.01	0.01
	8,3;6,7	2	t	0.42	2.57	0.27	0.3	0.3
4	6,5;2,4,1	2	t	-0.09	3	-0.06	0.09	0.09
	8,2;6,4,1	2	t	0.03	7.55	0.02	0.02	0.02
	7,4;8,6,1	2	t	0.01	21.69	0	0	0
	3,1;7,8,6	2	t	-0.63	4.09	-0.43	0.01	0.01
5	8,5;6,2,4,1	2	t	-0.91	2.04	-0.74	0	0
	7,2;8,6,4,1	2	t	0.14	11.76	0.09	0.01	0.01
	3,4;7,8,6,1	2	t	0.03	27.91	0.02	0	0
6	7,5;8,6,2,4,1	2	t	0.28	8.83	0.18	0.04	0.04
	3,2;7,8,6,4,1	2	t	-0.41	6.15	-0.27	0	0
7	3,5;7,8,6,2,4,1	2	t	0.13	9.74	0.08	0.02	0.02

Moreover, R-Vine structure of the trees described in the above table is as follows:





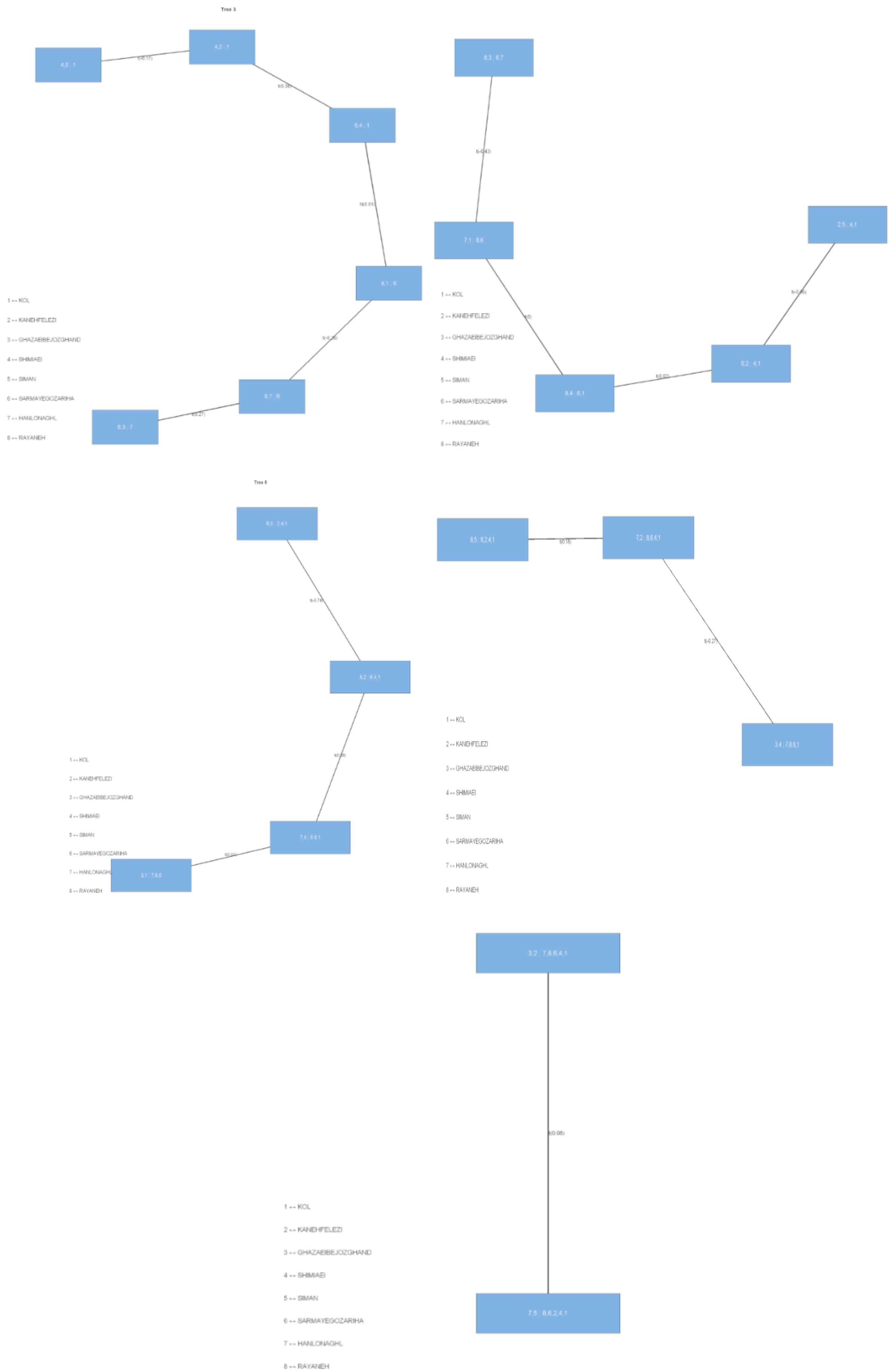


Fig. 4. R-Vine structure of the trees

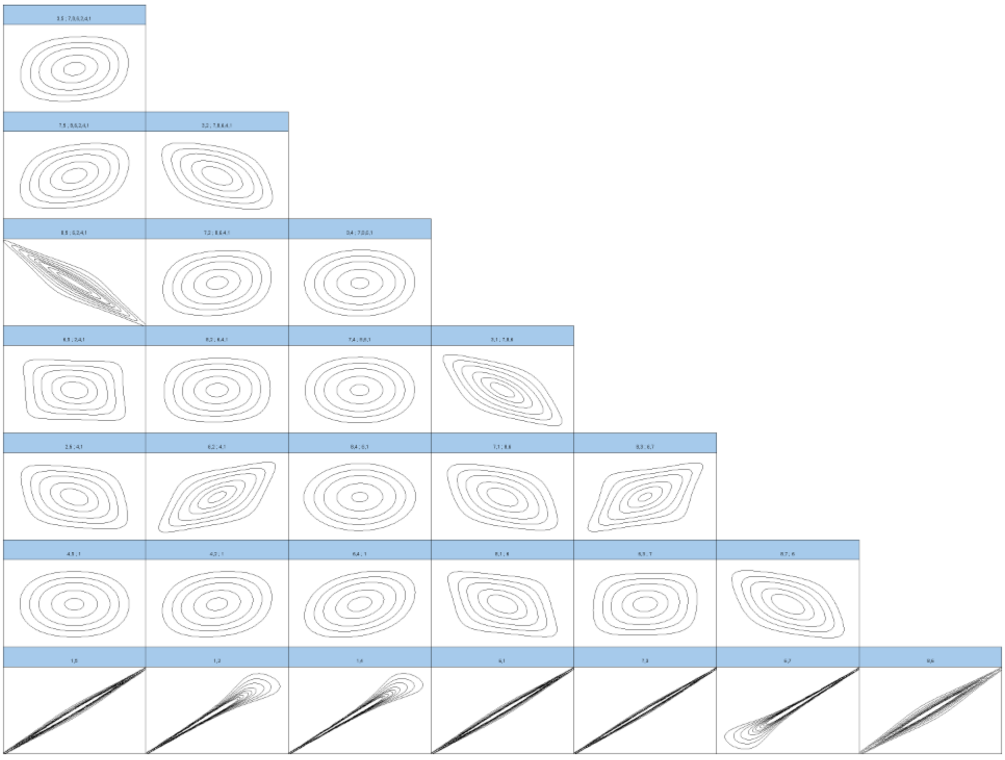


Fig. 5. the dependence structure

The best approach for identifying the dependence structure in an extreme sequence is offered by the copula functions, which can well reflect nonlinear and tail dependencies. Accordingly, we selected the best copula corresponding to the maximum likelihood for each pair of variables. Once finished with investigating the type of dependence between each pair of the considered series, the results were analyzed. A summary of the results of dependence structure modeling between the indices in Case 1.1 (Scenario 1) is reported in the below table.

Table 11. Results of the six models used for modeling the structure dependence under Scenario 1.

Model	**loglik**	**par**	**AIC**	BIC
R-vine-all	10,520.0	52.0	-20,935.9	-20,714.0
R-vine-normal	4,605.0	28.0	-9,153.9	-9,034.4
R-vine-tStudent	11,319.4	56.0	-22,526.8	-22,287.8
C-vine-all	10,824.0	52.0	-21,542.0	-21,315.8
C-vine-normal	4,483.0	28.0	-8,910.0	-8,790.5
C-vine-tStudent	11,523.1	56.0	-22,934.2	-22,695.3

Based on these results, it is evident that the best likelihood was obtained when C-Vine functions were coupled with t-student copulas. Accordingly, on the first tree (*i.e.*, first level) the Land Transportation & Storage industry showed the highest correlation to the TEDPIX, with the corresponding index producing the best performance by fitting by the t-student copula functions. At the second level, the highest correlation was seen between the Land Transportation & Storage and computer industries, provided the TEDPIX is varying. This implies that both industries are affected by domestic variations in prices. At the second level, the highest correlation was seen for the food and computer industries, provided the TEDPIX and Land Transportation & Storage indexes are varying. At the fourth level, if the TEDPIX, Land Transportation & Storage, and computer indices are varying, the highest changes were seen in the food and computer industries. At the fifth level, the results show that a change in the TEDPIX, Land Transportation & Storage, computer, and food industries imposes the largest

impact on the Metallic Mineral Products and investment industries. At the sixth level, it was observed that any change in the TEDPIX, Land Transportation & Storage, computer, food products, and Metallic Mineral Products indices imposes the largest effects on the indices of chemical and investment industries. At the seventh level, finally, the strongest dependencies were observed between the cement and investment industries, should any change occur to other indices. Based on the above-presented results, it is clear that the dependence structure modeling by the vine-copula functions could well extract and express conditional and pair relationships of different industries to one another in the considered cluster.
Final results of the six cases for modeling the dependence structure under Scenarios 2 and 3 are given in the following tables.

Table 12. Results of the six cases for modeling the dependence structure under Scenarios 2.

	loglik	par	AIC	BIC
R-vine-all	6,028	18	-12,021	-11,944
R-vine-normal	2,611	10	-5,202	-5,159
R-vine-tStudent	6,405	20	-12,771	-12,685
C-vine-all	6,008	19	-11,978	-11,896
C-vine-normal	2,438	10	-4,856	-4,814
C-vine-tStudent	6,406	20	-12,771	-12,686

Table 13. Results of the six cases for modeling the dependence structure under Scenarios 3.

	loglik	par	AIC	BIC
R-vine-all	11,629	52	-23,153	-22,931
R-vine-normal	4,666	28	-9,275	-9,156
R-vine-tStudent	12,078	56	-24,044	-23,805
C-vine-all	11,844	51	-23,585	-23,367
C-vine-normal	4,499	28	-8,941	-8,822
C-vine-tStudent	12,377	56	-24,642	-24,403

As shown in these tables, the application of the C-Vine model coupled with t-student copula functions led to better fitness values to the dependence structure of the market and industry indices in the relevant clusters. Under Scenario 2, where the clustering was based on foreign policies and the JCPOA, the Land Transportation & Storage industry exhibited the highest correlation to the TEDPIX. At the next level and provided a change in the TEDPIX and Land Transportation & Storage index occurs, the largest variations occur in the multi-discipline industries index. Focusing on the third tree, changes in the TEDPIX, Land Transportation & Storage, and multi-disciplinary indices were found to be most correlated to the technical and engineering index.

Under Scenario 3, where the clustering was performed on the basis of market caps for different industries, the highest correlation to the TEDPIX was exhibited by the Land Transportation & Storage index followed by the multi-disciplinary, technical and engineering, and investment industries, respectively, should the indices of the previous tree have varied.

Knowing that different cases were investigated under each scenario, the results of different models were subjected to the Vuong's test (1989) to compare the outcomes. As mentioned earlier, results of modeling under each scenario were finally compared by means of Vuong's tests, as reported in the following.

Table 14. Pairwise comparison of the six modeling cases under Scenario 1.

Model	stat	p.value	stat-aic	p-value	stat-bic	p-value
rv all vs normal	2.23	0.03	2.22	0.03	2.20	0.03
rv all vs t	-0.37	0.71	-0.37	0.71	-0.37	0.71
rv normal vs t	-12.52	0.00	-12.46	0.00	-12.35	0.00
cv all vs normal	2.26	0.02	2.25	0.02	2.23	0.03
cv all vs t	-0.33	0.74	-0.33	0.74	-0.32	0.75
cv normal vs t	-10.24	0.00	-10.20	0.00	-10.11	0.00
cv all vs rv all	1.96	0.05	1.96	0.05	1.94	0.05

Table 15. Pairwise comparison of the six modeling cases under Scenario 2.

Model	stat	p.value	stat-aic	p-value	stat-bic	p-value
rv all vs normal	1.93	0.05	1.92	0.05	1.91	0.06
rv all vs t	-0.26	0.79	-0.26	0.79	-0.26	0.80
rv normal vs t	-10.48	0.00	-10.46	0.00	-10.40	0.00
cv all vs normal	2.93	0.00	2.92	0.00	2.90	0.00
cv all vs t	-0.56	0.57	-0.56	0.57	-0.56	0.58
cv normal vs t	-7.59	0.00	-7.57	0.00	-7.53	0.00
cv all vs rv all	-0.03	0.98	-0.03	0.98	-0.03	0.97

Table 16. Pairwise comparison of the six modeling cases under Scenario 3.

Model	stat	p.value	stat-aic	p-value	stat-bic	p-value
rv all vs normal	2.20	0.03	2.19	0.03	2.18	0.03
rv all vs t	-0.16	0.87	-0.16	0.87	-0.16	0.88
rv normal vs t	-19.77	0.00	-19.70	0.00	-19.54	0.00
cv all vs normal	2.21	0.03	2.20	0.03	2.19	0.03
cv all vs t	-0.19	0.85	-0.19	0.85	-0.18	0.85
cv normal vs t	-14.74	0.00	-14.69	0.00	-14.58	0.00
cv all vs rv all	1.49	0.14	1.50	0.13	1.51	0.13

For example, the above tables indicate that, under Scenario 1, modeling the dependence structure using the R-Vine coupled with all families of copula functions led to no significant difference to the results when the R-Vine was coupled with the Gaussian copulas. This was while a significant difference was observed when the t-student copulas were used rather than all families of copulas, so that the null hypothesis could not be rejected. The same finding was seen with the C-Vine models as well.

Discussion and conclusion

With the growth and development of financial markets, the number of factors affecting the market trends has increased abruptly, so that variations in any segment of the financial system can potentially affect other parts of the system. On the other hand, the portfolio managers are always trying to understand various fluctuations in the market and draw a clearer image of the changes in the value of their portfolio considering the mentioned factors. For this purpose, they tend to consider nonlinear dependence structures to better comprehend the evolution of assets in the capital market. As far as dependence structure modeling is concerned, the vine-copula functions are superior to many other families of the copula functions including the Gaussian copula, Archimedean copula functions, and hierarchical Archimedean copula function (HAC). In a vine-copula function, the dependence structure is modeled by performing pairwise comparisons with a wide family of bivariate copulas. In this research, considering the large nature of the statistical population (including 10 market indices and 31 indices measuring active industries in the TSE) and structural differences for each index, we began by clustering different indices based on the most common criteria under three different scenarios. Then, by evaluating the structure of dependence between the market indices and those industries that represented the relevant clusters, an attempt was made to couple all families of copula functions with the C-Vine and R-Vine models to identify direct and indirect dependencies of these indices and graph them properly.

For this purpose, once finished with calculating daily returns of the considered indices, the GARCH(1.1) model was applied to obtain standardized values of residuals before using the C-Vine function to identify the representative index of each cluster. Under Scenario 1, the representative indices included those measuring the chemical products, metal minerals, Land Transportation & Storage, food products, computer, cement, and investment industries. Under Scenario 2, the representative indices were those related to the technical and engineering, multi-disciplinary, Land Transportation & Storage, and investment industries. And under Scenario 3, the representative indices were those corresponding to multi-disciplinary, investment, food products, Land Transportation & Storage, Rubber & Plastic Product, metal products, and oil production (excl. exploration). Subsequently, for all of the representative indices based on the R-Vine and C-Vine models, dependence structures were modeled using the entire family of copula functions, Gaussian copulas, and t-student copulas. The results indicated that, under either of the three scenarios, more accurate results were obtained when the C-Vine model rather than the R-Vine model was used. Focusing on the families of copula functions, the results showed that the best fitness values were obtained by the t-student functions. Finally, results of all cases of dependence structure modeling under the three scenarios were subjected to pairwise comparisons using Vuong’s test. The results showed significant similarity between the cases where all copula families were considered and those where t-student copulas were engaged, while the results were significantly different to those of Gaussian models. This proves that the distribution of financial data may not follow a normal distribution.

Finally, it is worth mentioning that the present research showed that, considering the formula for calculating the TEDPIX on TSE, one can properly extract pair dependence structures between different variables and indices of different industries by clustering them based on the investor’s requirements, thereby improving the composition of the investor’s portfolio.

References

[1] Khalili S, Tehrani R. Dependence structure between Iranian financial system's sub sectors: a vine copula approach. Journal of Investment Knowledge. 2019 Jun 22;8(30):51-72.
[2] Rodríguez-Moreno M, Peña JI. Systemic risk measures: The simpler the betterr. Journal of Banking & Finance. 2013 Jun 1;37(6):1817-31.
[3] Mansoori Gargari H, Khodavaisi H. Comparing Logistic Growth Model and Competing Models for Predicting the Overall Index in the Tehran Stock Exchange. Financial Research Journal. 2019 Aug 23;21(2):265-92.
[4] Esmaeil Pishbahar, Sahar Abedi, Measuring portfolio Value at Risk: The application of copula approach, Financial Engineering and Protfolio Management, 2017; 8(30): 55-73. magiran.com/p1683977
[5] Bokusheva R, 2010. Measuring the dependence structure between yield and weather variables. ETH Zurich, Institute for Environmental
[6] Brechmann E, Czado C, Paterlini S. Flexible dependence modeling of operational risk losses and its impact on total capital requirements. Journal of Banking & Finance. 2014 Mar 1;40:271-85.
[7] Czado C, Brechmann EC, Gruber L. Selection of vine copulas. InCopulae in mathematical and quantitative finance 2013 (pp. 17-37). Springer, Berlin, Heidelberg.
[8] Fischer M, Köck C, Schlüter S, Weigert F. An empirical analysis of multivariate copula models. Quantitative Finance. 2009 Oct 1;9(7):839-54.
[9] Joe H, Kurowicka D, editors. Dependence modeling: vine copula handbook. World Scientific; 2011.
[10] Nelsen RB. An introduction to copulas. Springer Science & Business Media; 2007 Jun 10.
[11] Cherubini U, Luciano E, Vecchiato W. Copula methods in finance. John Wiley & Sons; 2004 Oct 22.
[12] Aas K, Czado C, Frigessi A, Bakken H. Pair-copula constructions of multiple dependence. Insurance: Mathematics and economics. 2009 Apr 1;44(2):182-98.
[13] Joe H. Dependence modeling with copulas. CRC press; 2014 Jun 26.
[14] Schepsmeier U, Brechmann EC. Modeling dependence with C-and D-vine copulas: The R package CD vine. J. Stat. Software. 2013;52(3):1-27.
[15] Yu W, Yang K, Wei Y, Lei L. Measuring Value-at-Risk and Expected Shortfall of crude oil portfolio using extreme value theory and vine copula. Physica A: Statistical Mechanics and its Applications. 2018 Jan 15;490:1423-33.
[16] Bedford T, Cooke RM. Probability density decomposition for conditionally dependent random variables modeled by vines. Annals of Mathematics and Artificial intelligence. 2001 Aug;32(1):245-68.
[17] Dissmann J, Brechmann EC, Czado C, Kurowicka D. Selecting and estimating regular vine copulae and application to financial returns. Computational Statistics & Data Analysis. 2013 Mar 1;59:52-69.
[18] Sayadi M, Karimi N. Modeling the Dependency Structure between Stocks of Chemical Products Return, Oil Price and Exchange Rate Growth in Iran; an Application of Vine Copula. Journal of Economic Modeling Research. 2020 Feb 10;10(38):45-94.
[19] Wang K, Chen YH, Huang SW. The dynamic dependence between the Chinese market and other international stock markets: A time-varying copula approach. International Review of Economics & Finance. 2011 Oct 1;20(4):654-64.
[20] Skoglund J, Erdman D, Chen W. A mixed approach to risk aggregation using hierarchical copulas. Journal of risk management in financial institutions. 2013 Mar 1;6(2):188-205.
[21] Brechmann EC. Hierarchical Kendall copulas and the modeling of systemic and operational risk (Doctoral dissertation, Technische Universität München).
[22] Jalaei Esfandabadi ,S A, Salehi Asfiji, Elham Shivaei, Modeling the relationship between price index in financial markets Noorullah , the relationship of exchange in the Iranian economy (Merton price jump pattern and the approach of conditional capillary functions). Journal of Financial Economics, 1397; 12 (42): 1-24.

[23] Sukcharoen K, Leatham DJ. Hedging downside risk of oil refineries: A vine copula approach. Energy Economics. 2017 Aug 1;66:493-507.
[24] Reboredo JC, Ugolini A. Downside/upside price spillovers between precious metals: A vine copula approach. The North American Journal of Economics and Finance. 2015 Nov 1;34:84-102.
[25] BenSaïda A. The contagion effect in European sovereign debt markets: A regime-switching vine copula approach. International Review of Financial Analysis. 2018 Jul 1;58:153-65.
[26] Brechmann E, Czado C, Paterlini S. Flexible dependence modeling of operational risk losses and its impact on total capital requirements. Journal of Banking & Finance. 2014 Mar 1;40:271-85.
[27] Keshavarz Haddad G, Heyrani M. Estimation of Value at Risk in the Presence of Dependence Structure in Financial Returns: A Copula Based Approach. Journal of Economic Research (Tahghighat-E-Eghtesadi). 2014 Dec 22;49(4):869-902.
[28] MOUSAVI M, RAGHFAR H, MOHSENI M. Estimation of the Value of Risky Stocks (Using Conditional Copila-Garch Metho
[29] Pishbahar, Esmaeil, and Sahar Abedi. "Measuring portfolio Value at Risk: The application of copula approach." (2017): 55-73.
[30] Dissmann J, Brechmann EC, Czado C, Kurowicka D. Selecting and estimating regular vine copulae and application to financial returns. Computational Statistics & Data Analysis. 2013 Mar 1;59:52-69.