財金論文叢刊
2016 年 6 月，第二十四期，1-20

Modeling Dependence and Optimal Portfolios Using Vine Copula-GJR Approach: A China Stock Market Application

Weichen Sang
Faculty of Finance, Yunnan University of Finance and Economics, Yunnan, China

Jianxu Liu
Faculty of Economics, Chiang Mai University, Chiang Mai, 50200, Thailand

Jiajie Guo
Department of Economic Management, Oxbridge College, Kumming University of Science and Techenology, Kun Ming, 650106, China

Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University, Chiang Mai, 50200, Thailand

Abstract

This paper constructs a framework using the vine copula-GJR model to capture the struct ure interdependence of assets and combines it with simulation studies to calculate the optimal portfolio, expected shortfall (ES), and component ES. This study aims to investigate the lever age effects of the top five firms in Shanghai Stock Exchange market, measure their interdepen dences, forecast their optimal portfolios, and identify the most risky firm at t + 1 period. The major empirical results show that the C-vines demonstrate a better performance than the D-vi nes, and the biggest contributor to the overall risk is PetroChina firm under optimal weights.

Keywords: vine copulas; risk; component expected shortfall; portfolio; China

1. Introduction

In recent years, multivariate copulas have been commonly used in finance, econometrics, economics, and risk management, among others. Guégan and Maugis [2011] presented an econometric study of pair construction copulas. Nikoloulopoulos et al. [2012] used vine copulas to investigate the asymmetric tail dependence of some European financial markets. Czado et al. [2012] used the combination of C-vine copula and maximum-likelihood sequential estimation methods to estimate the conditional dependence of exchange rates. Evidently, vine copulas have become increasingly mature. Compared with vine copulas, standard multivariate copulas, such as multivariate normal and multivariate-t copulas, have become inflexible in high dimensions for limiting different dependency structures between pairs of variables. By contrast, vine approach is more flexible because bivariate copulas can be selected from a wide range of (parametric) families.

Linear correlation is not invariant under nonlinear strictly increasing transformation. Moreover, financial data are more highly correlated during volatile markets and market downturns. Therefore, using common multivariate distributions to measure risk presents some disadvantages. These disadvantages are attributed to copula, which possesses the properties of capturing tail dependence and nonlinear correlation. Some studies have been conducted on risk values or expected shortfall (ES) based on copula approach. Songsak et al. [2014] showed a framework that uses the Monte Carlo simulation and presented the results of vine copula to estimate the ES of an equally weighted portfolio. Huang et al. [2009] estimated the risk value of a portfolio using conditional bivariate copula-GARCH method. Deng et al. [2011] combined the vine copula-GARCH model with the Monte Carlo simulation and the Mean-CVaR model to optimize a portfolio. Mendes and Marques [2012] used robust pair-copula-based estimates on expected return and covariance in the mean–variance analysis of Markowitz to obtain an efficient diversification of investments. Wu and Chen [2006] applied the copula-GARCH model to analyze the risk portfolio of the Chinese stock market. Guégan and Maugis [2011] reported an application of vine copulas in estimating the VaR of a portfolio. Apart from the analysis of portfolio optimization, systemic identification of important institutions and measurement of their contribution to the total loss are also indispensable components of risk management. Banulescu and Dumitrescu [2014] proposed the component ES (CES) to identify systemically important institutions and access their contributions to systemic risk at a specific date. They defined CES as a combination of volatility, correlation, conditional expectations on standardized innovative distributions, and firm size. Although Banulescu and Dumitrescu [2014] invented the CES equation to analyze systemic risks, they used GJR-GARCH and dynamic conditional correlation models to obtain linear correlations and conditional expectations.

Some previous studies have used the vine copula-GARCH model to analyze risks and optimization portfolios in stock markets. However, an analysis of China's stock market portfolio based on vine copula-GJR with CES model has not been reported yet. To more comprehensively analyze portfolios, this paper used the vine copula-GJR model, which allows the construction of a flexible multivariate distribution with different margins and dependent structures by decomposing multivariate density into a series of bivariate copula and marginal densities.

The contribution of this research involves five aspects. First, this study is the first to apply vine copula-GJR with CES model to examine optimal portfolios and illustrate the approach to analyze the Chinese stock market. Second, this study applies many pair copulas to construct C- and D-vine copula structures, thereby improving the flexibility of vine copula. Third, time-varying copulas are estimated in all unconditional pair copulas among vine structures because unconditional dependence may vary. Lastly, this approach is used to analyze the performance of different Chinese stock market portfolios.

In this study, new and necessary insights on dependence structure, portfolio strategies, and risk prevention for investors and financial institutions are obtained using the stock prices of the top five firms in Shanghai stock market, namely, PetroChina, Industrial and Commercial Bank of China (ICBC), Agricultural Bank of China (ABC), and Bank of China (BC) from September 1, 2009 to May 28, 2014. The main findings of this study are as follows. First, based on vine copula-GJR model, long-term unexpected shocks cause large impacts on the volatilities of the top five firms, whereas the influence of short-term unexpected factors is very minimal. Sinopec exhibits leverage effect on the volatility of the structures, but an asymmetric effect is not observed in PetroChina, ICBC, ABC and BC. Nevertheless, PetroChina and Sinopec show the highest linear and rank correlations, whereas the upper and lower dependences of BC and ABC are the highest tail dependences. High degree of persistence is also obtained on the structure dependences of BC and ICBC, as well as of BC and ABC. Second, the model suggests that investors should invest few proportions in ICBC and Sinopec to avoid extreme losses. In addition, the ranks of the top five firms are accurately identified according to their risks. The five firms are ranked as PetroChina, BC, ABC, Sinopec, and ICBC based on their absolute contribution under optimal weights at 95% and 99% confidence levels. In summary, the vine copula-GARCH approach enables the identification of structure dependence and potential nonlinear dependences among the top five firms in China's stock market. The CES based on vine copula-GJR method with the Monte Carlo simulation enables investors to identify high-risk firms and to become perceptive of their risk strengths in China's stock market.

The paper is organized as follows. Section 2 presents the data, and Section 3 introduces the methodologies, including the GJR model, copulas, and C- and D-vine constructions. Section 4 discusses the portfolio optimization model, and Section 5 analyzes the empirical results. Lastly, Section 6 concludes the paper.

2. Data

The data of this study consist of the top five firms in Shanghai Stock Exchange, namely, PetroChina, ICBC, ABC, BC, and Sinopec, which account for 22.8% of market capitalization. These data are obtained from September 1, 2009 to May 28, 2014, which yield a total of 877 observations for each series. Figure 1 shows the close price trends of each firm, which shows high rank correlations. Datasets are transformed logarithmically, and their differences from the initial datasets are obtained; thus, the data obtained correspond to asset returns.

Table 1 presents the descriptive statistics for the asset returns of the stock prices of the top five firms. All the series show approximately zero means, negative skewness except that in ABC, and excessive kurtosis. Moreover, the data for each variable are not normally distributed because the Jarque–Bera test rejected the null hypothesis for each index return. Results show that PetroChina, ICBC, BC, and Sinopec are skewed to the left, whereas ABC is skewed to right; all variables are peak and fat-tailed. Table 1 also shows that the empirical Kendall's τ of the components of each pair are closely related; the biggest dependence is observed between PetroChina and Sinopec.

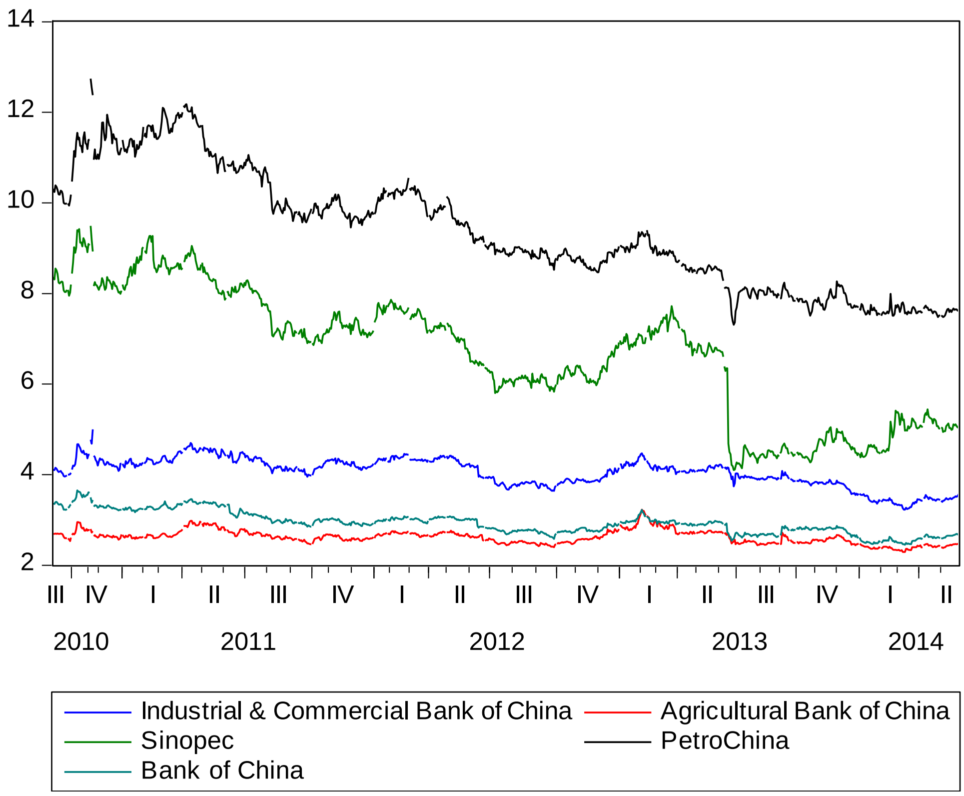

Figure 1 Top five stock prices of China

Table 1. Data Description and Statistics

	PetroChina	ICBC	ABC	BC	Sinopec
Mean	-0.0003	-0.0002	-0.0001	-0.0003	-0.0006
Median	0.0000	0.0000	0.0000	0.0000	-0.0012
Maximum	0.1111	0.0717	0.0964	0.0615	0.0953
Minimum	-0.1227	-0.1233	-0.0459	-0.0673	-0.3052
Std. Dev.	0.0115	0.0116	0.0110	0.0097	0.0180
Skewness	-0.0990	-0.7637	0.8145	-0.2914	-5.3763
Kurtosis	28.5616	22.6663	11.6464	10.2380	96.8675
Jarque-Bera	23877.69	14218.31	2828.880	1926.801	326197.9
Probability	0.0000	0.0000	0.0000	0.0000	0.0000
Kendall’s tau					
PetroChina	1.0000	0.3364	0.3180	0.3495	0.5198
ICBC	0.3363	1.0000	0.5182	0.5169	0.3206
ABC	0.3180	0.5181	1.0000	0.5171	0.2951
BC	0.3495	0.5169	0.5171	1.0000	0.3532
Sinopec	0.5198	0.3206	0.2951	0.3532	1.0000

3. Methodology

According to studies in recent years, vine copulas have fixed application mode. As a general rule, the ARMA-GJR model is used first to fit the time series, which is assumed to have skewed student-t or skewed generalized error distributions. The corresponding standardized residuals of each margin are then considered as i.i.d. sample over time. Second, using cumulative distribution, the corresponding standardized residuals are transformed to obtain a uniform distribution between 0 and 1. Third, the preferred sequences of the vine copulas in different variable orders are selected according to their influence. In addition, the best family for each bivariate copula model is selected and estimated among the available copula families. Normally, the sequential maximum-likelihood method is used to obtain the starting values, then the maximum-likelihood method is used to estimate the parameters of the vine copulas, and the Akaike information criterion (AIC) values are evaluated to determine the best copulas. Remarkably, the empirical Kendall’s τ, Spearman’s ρ, and degree of freedom of the T copula can be used to determine the extent of the influence, which is similar to the studies of Aas et al. [2009] and Czado et al. [2012].

3.1 Specification for marginal distribution

Glosten, Jagannathan and Runkle [1993] proposed GJR-GARCH model to judge whether there exist leverage effect in stock market. The form of the ARMA (P, Q)-GJR (K, L) model can be expressed as

$$r_t = c + \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \psi_i \varepsilon_{t-i} + \varepsilon_t \tag{1}$$

$$\varepsilon_t = h_t \cdot \eta_t \tag{2}$$

$$h_t^2 = \omega + \sum_{j=1}^{K} \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^{K} \lambda_j \mathrm{I}[\varepsilon_{t-j} < 0]\varepsilon_{t-j}^2 + \sum_{i=1}^{L} \beta_i h_{t-i}^2 \tag{3}$$

where $\sum_{i=1}^{p} \phi_i < 1$, $\omega > 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$, $\alpha_j + \lambda_j \geq 0$ and $\sum_{j=1}^{K} \alpha_i + \sum_{i=1}^{L} \beta_i + \frac{1}{2}\sum_{j=1}^{K} \lambda_i < 1$. The leverage coefficient λ_j is applied to negative standardized residuals, giving negative changes additional weight. η_t is called standardized residuals, which is assumed to be the skewed student-t or skewed generalized error distribution in this study. They can capture the characteristics of heavy tail and asymmetry. Fernandez and Steel [1998] proposed skewed student-t distribution, the formula is shown as follow,

$$p(\eta_i \left| v, \lambda \right.) = \frac{2}{\lambda + \frac{1}{\lambda}} \left\{ f_v(\frac{\eta_i}{\lambda}) I_{[0,\infty)}(\eta_i) + f_v(\lambda \eta_i) I_{(-\infty,0)}(\eta_i) \right\}, \tag{4}$$

and the skewed generalized error distribution can be expressed as

$$f_{sged}(\eta_t \left| v, \gamma \right.) = v(2\theta \cdot \Gamma(1/v))^{-1} \cdot \exp(-\frac{\left|\eta_t - \delta\right|^v}{(1 - sign(x_t - \delta)\lambda)^v \theta^v}) \tag{5}$$

where

$$A = \Gamma(2/v)\Gamma(1/v)^{-0.5}\Gamma(3/v)^{-0.5}$$

$$S(\lambda) = \sqrt{1 + 3\lambda^2 - 4A^2\lambda^2}$$

$$\delta = 2\lambda \cdot AS(\lambda)^{-1}$$

$$\theta = \Gamma(3/v)^{-0.5}\sqrt{\Gamma(1/v)}S(\lambda)^{-1}$$

where $f_v(.)$ is unimodal and symmetric around zero, and λ is the skewness parameter that is defined from 0 to ∞; I denotes the indicator function and v is degree of freedom.

3.2 Copulas

Traditionally, we would use classical families of bivariate distributions such as bivariate normal, student-t, lognormal, and extreme value distributions to fit the sample data and to estimate the pairwise dependence. However, these classical families restrict the individual behavior of each variable because the individual behaviors have to be described by the same parametric family of univariate distribution.

Copula models are able to completely circumvent this restriction. A copula function can decompose the joint distribution into two parts: marginal distribution and dependence structure. Sklar [1959] was the first person to give the definition of a copula: Let $x = (x_1, x_2, \cdots x_n)$ be a random vector with the joint distribution function H and the marginal distribution $F_1, F_2, \cdots F_n$; then, there exists a function C which is called a copula function.

$$F(x_1, x_2, \cdots, x_n) = C(F_1(x_1), F_2(x_2), \cdots, F_n(x_n)) \tag{6}$$

In the light of formula (6), the copula function can be expressed as follows:

$$C(u_1, u_2, \cdots, u_n) = F(F_1^{-1}(u_1), F_2^{-1}(u_2), \cdots, F_n^{-1}(u_n)) \tag{7}$$

If F_i is an absolutely continuous function that is strictly increasing, we have the density function as follows:

$$\begin{aligned} f(x_1, \cdots\cdots x_n) &= \frac{\partial F(x_1, \cdots\cdots, x_n)}{\partial x_1 \cdots \partial x_n} \\ &= \frac{\partial C(u_1 \cdots, u_n)}{\partial u_1 \cdots \partial u_n} \times \prod \frac{\partial F(x_i)}{\partial x_i} \\ &= c(u_1 \cdots, u_n) \times \prod f_i(x_i) \end{aligned} \tag{8}$$

Copulas have many merits that help tremendously in studying dependence structures. It is worth mentioning that copulas can measure both the rank dependence and the tail dependence. Generally, we use Kendall's tau to reflect the size of a nonlinear correlation. Rank correlation reflects the monotonic dependence between variables. The empirical Kendall's tau coefficient is defined as follows:

$$\tau = \frac{c-d}{c+d} = \frac{2(c-d)}{n(n-1)} \tag{9}$$

where c is the number of concordant pairs, and d is the number of discordant pairs. In addition, the Kendall's tau coefficient can be calculated using the copula function (see Nelsen [2006] for details) as

$$\tau_{X,Y} = \tau_K = 4\iint_{[0,1]^2} C(u_1, u_2) dC(u_1, u_2) - 1 \tag{10}$$

where X and Y are continuous random variables and C is the copula function, and u_1 and u_2 are the values of the marginal probability distribution functions. Moreover, there is the possibility that the tail dependence may divide into upper tail and lower tail dependences, the definitions of which are written as

$$\lambda_L := \lim_{u \to 0^+} P[Y \le G^{-1}(u) \mid X \le F^{-1}(u)] = \lim_{u \to 0^+} \frac{C(u,u)}{u} \tag{11}$$

and

$$\lambda_{up} := \lim_{u \to 1^-} P[Y > G^{-1}(u) \mid X > F^{-1}(u)] = \lim_{u \to 1^-} \frac{1-2u+C(u,u)}{1-u}, \tag{12}$$

where F and G are the marginal cumulative distribution functions of X and Y, respectively. Some copulas, such as the Gaussian and Frank copulas, possess the characteristic $\lambda_{up} = \lambda_L = 0$, while most copulas can capture upper tail and/or lower tail dependence. For instance, Clayton copula can measure lower tail dependence, and Gumbel copula can measure upper tail dependence, while student-t copula reflects symmetric tail dependence.

Patton [2006] proposed the time-varying copulas that include the Gaussian and symmetric Joe copulas. Manner and Reznikova [2012], Wu et al. [2012] further researched the time-varying copulas. In our study, we follow the formula of the research of Wu et al. [2012]. The formula of time-varying T copula can be expressed as

$$\rho_t^* = c + \beta \cdot \rho_{t-1}^* + \alpha \cdot (u_{t-1} - 0.5)(v_{t-1} - 0.5) \tag{13}$$

where $\rho_t^* = -\ln[(1-\rho_t)/(1+\rho_t)]$, which is used to ensure the correlation fall within (-1, 1). We fix the parameter degree of freedom of T copula. So the time-varying correlation is reflected by the linear correlation ρ.

3.3 C-vine and D-vine constructions

Vine copulas are a kind of multivariate copulas which use bivariate copulas to construct multivariate distributions, and specify the dependence and conditional dependence of selected pairs of random variables and all marginal distribution functions. A d-dimensional vine copula is decomposed into d(d-1)/2 pair-copulas, and the densities of multivariate vine copulas can be factorized in terms of pair-copulas and margins. Bedford and Cooke [2001, 2012] proposed two subclasses of PCC, which we call as C-vine and D-vine copulas. Note that C-vine and D-vine copulas have been widely used in finance asset returns and other data, for example, in studies conducted by Aas et al. [2009], Min and Czado [2010], and Czado [2010].

For C-vines and D-vines, the densities are, respectively [2009],

$$f(x_1, x_2, \cdots, x_d) = \prod_{k=1}^{d} f(x_k) \times \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{j,i+j|1,\ldots,j-1}(F(x_j | x_{1:j-1}), F(x_{i+j} | x_{i+j|1:j-1})) \tag{14}$$

$$f(x_1, x_2, \cdots, x_d) = \prod_{k=1}^{d} f(x_k) \times \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{i,i+j|i+1,\ldots,i+j-1}(F(x_i | x_{i+1:i+j-1}), F(x_{i+j} | x_{i+1|i+1:i+j-1})) \tag{15}$$

where $F(\cdot|\cdot)$ is the conditional distribution which can be got from the first deviation of copulas, for example,

$$F(x_1 | x_2) = \frac{\partial C_{x_1,x_2}(F(x_1), F(x_2))}{\partial F(x_2)}, \tag{16}$$

is the conditional distribution, given one variable. The marginal conditional distributions of given multivariate variables can be expressed by the form F(X|υ),

$$F(x|\upsilon)=\frac{\partial C_{x,\upsilon_j|\upsilon_{-j}}(F(x|\upsilon_{-j}),F(\upsilon_j|\upsilon_{-j}))}{\partial F(\upsilon_j|\upsilon_{-j})}, \tag{17}$$

where υ stands for all the conditional variables. The copula family used in our work includes Gaussian copula, T copula, Clayton copula, Gumbel copula, Frank copula, BB1, BB7, BB8, and rotate copulas. Different copulas have different characteristics. In order to accurately capture the dependency, we apply copulas as much as possible.

In this study, we turn to the sequential maximum likelihood estimation method for obtaining the initial values, and then use the maximum likelihood estimation method to estimate the parameters of the C-vine and D-vine copulas. Aas et al. [2009] and Czado et al. [2012] introduced the detailed calculation process. A brief process of sequential maximum likelihood estimation can be described as follows. First, the maximum likelihood estimation method is used to estimate the parameters of each non-conditional copula; second, the observations are computed by using the conditional distribution function (Formula (16)) and the known non-conditional copulas in the first step; third, the parameters of the conditional copulas are estimated on one variable; fourth, the observations for the copulas are computed, given two variables, by using Formula (17); and, last, the copulas, given two variables, are estimated using the observations from the fourth step. Through these five steps, we can get the initial values of the 4-dimensional vine copulas. If there are more 4-dimensional variables, the observations may be obtained by using Formula (17) as well.

4. Portfolio Optimization Model

In risk management, VaR is probably the most popular risk measure. But, it does not satisfy the property of subadditivity. There is an alternative method which is called as expected shortfall, which satisfies the property of subadditivity and provides a more conservative measure of losses relative to VaR. Also, the component expected shortfall might be useful as a supplement to expected shortfall in terms of measuring systemic risk. By following the method discussed by Rockafellar and Uryasev [2002], the optimization approach-based minimum expected shortfall can be formulated as follows:

$$\min ES_\beta(W)=\min\{VaR_\beta(W)+\frac{1}{q(1-\beta)}\sum_{k=1}^{q}[-W^T r_k-VaR_\beta(W)]^+\}, \tag{18}$$

where $[t]^+=\max(t, 0)$ and $\sum_{i=1}^{n} w_i=1$. $VaR_\beta(W)$ is the VaR under the β confidence level and the W portfolio allocations, and r_k is the return vector from the Monte Carlo simulation of the vine copula.

Banulescu and Dumitrescu [2014] defined the component expected shortfall as that which measures the absolute contribution of a firm to the risk of the financial system; the systemic risk can be measured by linearly aggregating the component losses, as well. They are given by

$$CES_{it} = w_{it} \frac{\partial ES_{m,t-1}(VaR_{\beta}(W))}{\partial w_{it}} \tag{19}$$

$$CES\%_{it}(VaR_{\beta}(W)) = \frac{CES_{it}(VaR_{\beta}(W))}{SES_{t-1}(VaR_{\beta}(W))} \times 100 \tag{20}$$

and $SES_{t-1}(VaR_{\beta}(W)) = \sum_{i=1}^{n} CES_{it}(VaR_{\beta}(W))$ (21)

With these formulas, the absolute contribution of a firm and the systemic risk can be measured under given optimal weights, based on the minimum ES portfolio. For out-of-sample systemic risk, Banulescu and Dumitrescu [2014] show it may be solved using (22):

$$CES\%_{i,T+1}(\tilde{C}) = \frac{CES_{i,T+1}(\tilde{C})}{\sum_{i=1}^{n} CES_{i,T+1}(\tilde{C})} \tag{22}$$

where $\tilde{C}$ is set to be the out-of-sample VaR-HS for cumulative market returns at a coverage rate of 1%, 5% and 10%.

In this study, the forecasting of the t+1 period optimization portfolios is carried out by combining the vine copula-GJR model with ES and CES, and a collection of the return r_k at period t+1 is generated by the Monte Carlo simulation. The process, to summarize, can be illustrated probably in four steps. First, use the estimation results of the preferred vine copula to generate the random number 10,000. Second, use the inverse function of the corresponding marginal distribution (skewed student-t or skewed generalized error distribution) of each variable to get the standardized residuals. Third, forecast the value of each variable at the t+1 period by using the ARMA-GJR model; thus, 10,000 possible returns are generated at the t+1 period for each variable, which can be expressed as

$$r_{m,t+1}^{n} = \hat{c} + \sum_{i=1}^{p} \hat{\phi}_{m,i} r_{m,t-i+1} + \sum_{j=1}^{q} \hat{\psi}_{m,i} \varepsilon_{m,t-i+1} + \hat{h}_{m,t+1} \eta_{m,t+1}^{n}, \tag{23}$$

where n=1,2, 10,000, m equals to the number of variables, $\eta_{m,t+1}^{n} = F_{m}^{-1}(u_{m,t+1}^{n})$, and u_m is from the simulation of the preferable vine copula, $\hat{h}_{m,t+1}$ is forecasted by variance equation of ARMA-GJR model. Last, by giving an unknown weight to each variable, the optimal portfolio weights of the selected assets are estimated under the minimum ES framework. Repeating 1000 times these four steps, it can able to obtain 1000 weights for each variable. Therefore, the mean value of weights and standard error can be obtained for each variable, respectively. In addition, $CES\%_{i,t+1}$ can be obtained by equation 22.

In addition, a C-vine copula is used to generate random numbers. The specific process of generation—for example, take the generation of the C-vine copula of four variables — can probably be summarized in the following steps:

(1) Independently generate z_1, z_2, z_3, and z_4 from the uniform distribution on [0, 1].

(2) Let $z_1=u_1$ and $z_2=F(u_2|u_1)$; then, u_2 can be calculated using the inverse function of Formula (16), namely, $u_2=F^{-1}(z_2|u_1)$.

(3) Assume that $z_3 = F(u_3 \mid u_1, u_2) = \dfrac{\partial C(F(u_2 \mid u_1), F(u_3 \mid u_1))}{\partial F(u_2 \mid u_1)}$; then, $u_3= F^{-1}(z_3|u_1,u_2)$.

(4) By the same token, we assume that

$$z_4 = F(u_4 \mid u_1, u_2, u_3) = \frac{\partial C(F(u_3 \mid u_1, u_2), F(u_4 \mid u_1, u_2))}{\partial F(u_3 \mid u_1, u_2)}$$

Then we can get $u_4= F^{-1}(z_4|u_1,u_2,u_3)$. Regardless of the number of variables that are present, it is advisable that we use Formula (17) to generate the recursive random numbers of the C-vine or D-vine copula structure.

5. Application to China Sector Indices

5.1 Results of ARMA-GJR model

ARMA (1, 1)-GJR (1, 1), ARMA (2, 2)-GJR (1, 1), and ARMA (1, 0)-GJR (1, 1) are selected as the financial return data for PetroChina and Sinopec, ABC and BC, and ICBC, respectively. Table 2 shows the results of ARMA-GJR models for all the asset returns. With the assumption that the standardized residuals of PetroChina, ABC, BC, and Sinopec have skewed student-t distributions and that of ICBC have skewed generalized error distribution, the α and β parameters are found significant at 90% confidence level. This result indicates that all returns exhibit clustered and persistent volatilities. Furthermore, the asymmetric parameter γ for Sinopec is significant and exhibits leverage effect on structural volatility, but an asymmetric effect is not observed in PetroChina, ICBC, ABC, and BC markets. The skewness parameter λ and degree of freedom υ are significant at 99.99% confidence level, which imply that the skewed student-t and skewed generalized error distributions fit very well for each asset returns.

Prior to copula analysis, the marginal distribution of each series is tested if it satisfies the assumptions of serial independence and uniform distribution (0, 1). If any of these two assumptions is rejected, then inaccurate specification of the marginal distribution may cause incorrect fit to the copulas. Thus, testing these two assumptions for marginal distribution and model specification is critical in constructing multivariate distribution models using copulas [2006]. The Box–Ljung and Kolmogorov–Smirnov (KS) tests are used to evaluate the serial independence of the marginals and their distribution specifications, respectively. The test results in Table 3 show that none of the probability values obtained in the KS test rejected the null hypothesis at 95% confidence level. Moreover, all the return data passed the Box–Ljung test at 90% confidence level. Therefore, both tests have confirmed the use of the skewed student-t and skewed generalized error distributions to fit the margins. The margins are also found to satisfy the assumptions of uniform i.i.d. on (0, 1).

Table 2 Results of ARMA-GJR Model

	PetroChina	ICBC	ABC	BC	Sinopec
Ar1	0.8132*** (0.0799)	-0.6653*** (0.0358)	-0.1729*** (0.0395)	-0.1579*** (0.0106)	-0.9282*** (0.0935)
Ar2	---	---	0.7841*** (0.0408)	-0.9549*** (0.0020)	---
Ma1	-0.8736*** (0.0655)	0.6148*** (0.0371)	0.1234*** (0.0238)	0.1352*** (0.0040)	0.9431*** (0.0820)
Ma2	---	---	-0.8110*** (0.0706)	0.9531*** (0.0040)	---
ω	0· (0)	0** (0)	0· (0)	0 (0)	0* (0)
α	0.0683* (0.0332)	0.1548** (0.0517)	0.0811* (0.0321)	0.0749· (0.0432)	0.1099*** (0.0092)
β	0.8941*** (0.0428)	0.7279*** (0.0772)	0.8934*** (0.0428)	0.7858*** (0.1118)	0.8771*** (0.0336)
γ	-0.0290 (0.0320)	-0.0631 (0.0634)	-0.0416 (0.0373)	-0.0184 (0.0447)	-0.1015*** (0.0089)
λ	1.0680*** (0.0413)	1.0079*** (0.0249)	1.0196*** (0.0377)	1.0848*** (0.0470)	1.1208*** (0.0406)
υ	3.5434*** (0.4738)	1.0896*** (0.0671)	3.6564*** (0.5122)	3.7450*** (0.4767)	3.3876*** (0.3996)
LogL	2851.619	2826.319	2819.515	2924.174	2924.174
AIC	-6.4849	-6.4272	-6.4071	-6.6458	-6.6458

Note: Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘·’0.1.
The numbers in the parentheses are the standard deviations.

Table 3 KS Test and Box-Ljung Test

	PetroChina	ICBC	ABC	BC	Sinopec
KS test statistics	0.0322	0.0410	0.0420	0.0394	0.0209
P value	0.3240	0.1051	0.0905	0.1308	0.8378
P value of LM test					
First moment	0.7350	0.9885	0.4214	0.9587	0.1382
Second moment	0.1584	0.7897	0.6235	0.5997	0.8737
Third moment	0.6561	0.9899	0.3467	0.7639	0.5636
Fourth moment	0.1098	0.8695	0.8512	0.8107	0.9229

5.2 Results for copulas

In this paper, copula families, including the Gaussian, T, Clayton, Frank, Gumbel, Joe, BB1, BB6, BB7, BB8, and rotate copulas, are used. Further details regarding these copula families, which include their properties and characteristics, are discussed by Liu and Sriboonchitta [2012] and Sriboonchitta et al. [2013].

According to the maximum rank correlation principle of the empirical Kendall's τ method, the orders of C- and D-vine copula structures are 4, 1, 2, 3, 5 and 1, 5, 4, 2, 3, respectively, with 1, 2, 3, 4, and 5 representing PetroChina, ICBC, ABC, BC, and Sinopec, respectively. Tables 4 and 5 present the estimated results for the C- and D-vine copulas. Table 4 shows that the C-vine copula has smaller values than that of the D-vine copula in Table 5, which explains the preference for C-vine copula over D-vine copula in this study.

The following discussions attempt to explain the estimated results in Tables 4 and 5. The first column of each table shows all the best copula families among the candidates based on AIC. As shown, the T, T, T, T, T, Frank, T, Frank, Clayton, and Clayton copulas are more suitable than the other families under the C-vine copula structure, whereas the T, T, T, T, Frank, Rotate Gumbel 180 degrees, T, Gumbel, Frank, and Frank copulas are the suitable copula families under the D-vine copula structure. As noted above, the C-vine copula better accounts for the potential asymmetry of structure dependence. Considering that four of the selected C-vine copulas are non-elliptical, different C-vine copulas are selected, namely, the Gaussian and the student-t copulas. In addition, half of the D-vine copulas are non-elliptical. The choice of flexible vine copula model for the joint distribution of asset returns in this study, which accounts for possible tail dependence between variables, is well justified.

The results show the highest linear and rank correlations between 1 (PetroChina) and 5 (Sinopec). The linear correlation and Kendall's τ are equal to 0.7036 and 0.5059, respectively. The upper and lower dependences between 4 (BC) and 3 (ABC) are the highest tail dependences. However, unconditional and conditional bivariate copulas exist in vine copulas. The implications of unconditional copulas can be directly interpreted compared with that of conditional dependences, which are more difficult to explain. Nonetheless, the conditional copulas are explained by comparing them with unconditional copulas. First, the dependence parameter between 2 (ICBC) and 5 conditional on 1 and 4 is not significant, which implies that ICBC and Sinopec are independent from PetroChina and BC. Conversely, the estimated parameter between 2 and 5 conditional 4 is significant, which indicates that PetroChina affects the dependence between ICBC and Sinopec. Second, the C_{41} and $C_{41|5}$ are compared. Sinopec converts the T copula structure of 1 and 4 to Frank copula and modifies their Kendall's τ values from 0.3601 to 0.1757. Third, both C_{15} and $C_{15|4}$ are T copula, and the estimated parameter changes are not large. Thus, the effect of BC on the dependence between 1 and 5 is assumed to be relatively low. Lastly, when PetroChina, ICBC, and BC are set as condition variables, the dependence between ABC and Sinopec becomes independent.

Therefore, as more information is provided, the correlations become lower.

Table 4. Estimated Results of C-vine Copula

Copulas	parameters	Lower and upper tail	Kendall's tau	AIC
T (C_{41})	0.5359*** (0.0237)	0.1824	0.3601	-291.6867
	6.2036*** (1.3194)	0.1824		
T (C_{42})	0.6992*** (0.0165)	0.3024	0.4929	-616.7345
	6.0000*** (1.1377)	0.3024		
T (C_{43})	0.7050*** (0.0185)	0.4455	0.4981	-669.0335
	3.1063*** (0.4451)	0.4455		
T (C_{45})	0.5351*** (0.0241)	0.1702	0.3595	-300.1966
	6.5815*** (1.6589)	0.1702		
T ($C_{12\|4}$)	0.2092*** (0.0328)	0.0112	0.1341	-44.7292
	12.2468** (4.7366)	0.0112		
Frank ($C_{13\|4}$)	1.2472*** (0.2012)	0	0.1364	-32.5288
T ($C_{15\|4}$)	0.6065*** (0.0223)	0.2115	0.4148	-404.7854
	6.5987*** (1.5949)	0.2115		
Frank ($C_{23\|14}$)	2.4140*** (0.2193)	0	0.2539	-127.5699
Clayton ($C_{25\|14}$)	0.0470 (0.0338)	0	0.0229	-0.1826
Clayton ($C_{35\|124}$)	0.0261 (0.0289)	0	0.0129	1.0987
Sum			2.5036	-2486.349

Note: The significant codes: 0 "***" 0.001 "**" 0.01 "*" 0.05 "·"0.1.
The numbers in the parentheses are the standard deviations.

Table 5. Estimated Results of D-vine Copula

Copulas	parameters	Lower and upper tail	Kendall's tau	AIC
T (C_{15})	0.7136*** (0.0169) 4.5386*** (0.8920)	0.3761 0.3761	0.5059	-643.8708
T (C_{54})	0.5327*** (0.0246) 5.5783*** (1.2841)	0.2022 0.2022	0.3576	-301.0207
T (C_{42})	0.6991*** (0.0165) 5.3542*** (0.8414)	0.3274 0.3274	0.4928	-617.2735
T (C_{23})	0.6988*** (0.0168) 5.2930*** (0.9795)	0.3296 0.3296	0.4925	-611.6604
Frank ($C_{14\|5}$)	1.6228*** (0.2106)	0	0.1757	-58.3364
R-Gumbel ($C_{25\|4}$)	1.0921*** (0.0238)	0.1135	0.0843	-21.8005
T ($C_{34\|2}$)	0.4566*** (0.0295) 5.6065*** (1.2888)	0.1629 0.1629	0.3018	-194.5773
Gumbel ($C_{12\|45}$)	1.1031*** (0.0257)	0.1254	0.0934	-23.8564
Frank ($C_{35\|24}$)	0.5245** (0.2097)	0	0.0581	-4.5265
Frank ($C_{13\|245}$)	0.5452** (0.2052)	0	0.0604	-5.0871
Sum			2.5036	-2482.01

Note: The significant codes: 0 "***" 0.001 "**" 0.01 "*" 0.05 "·"0.1.
The numbers in the parentheses are the standard deviations.

Despite the outstanding performance of vine copula in multivariate distributions, the unconditional dependence may still vary. Thus, time-varying copulas are estimated for all unconditional bivariate copulas of the vine structures. Considering that all unconditional copulas are T copula, the time-varying T copula model is used. The estimated results of time-varying copulas are shown in Table 6. In terms of AIC, the time-varying C_{42} and C_{43} demonstrate better explanations than the static T copula, whereas other copulas exhibit the opposite. The autoregressive parameters α in the time-varying T copulas C_{42} and C_{43} are equal to 0.9257 and 0.9202, respectively, which imply a high degree of persistence on the structure dependences of BC and ICBC, BC, and ABC. The dynamic Kendall's τ from the T copulas C_{42} and C_{43} are illustrated in Figure 2. The smallest and largest Kendall's τ values on BC and ICBC are approximately 0.38, and 0.62, respectively, which indicate a constant shift of nonlinear correlations with time. For BC and ABC, the smallest and largest Kendall's τ values are 0.42 and 0.63, respectively. In addition, the dependences between BC and ICBC as well as BC and ABC show similar volatility structures.

Table 6. Estimated Results of Time-Varying Copulas

Copulas	ω	α	β	υ	AIC
T (C_{41})	2.3235 ***	-0.9598 ***	0.5479	5.7218 ***	-290.5702
	(0.1491)	(0.0081)	(0.3709)	(1.5020)	
T (C_{42})	0.0891 ***	0.9257 ***	0.7879 **	5.4658 ***	-621.0715
	(0.0237)	(0.0115)	(0.2927)	(1.0152)	
T (C_{43})	0.1051**	0.9202***	0.7754*	3.0363***	-672.6256
	(0.0325)	(0.0177)	(0.3334)	(0.4425)	
T (C_{45})	1.0367	0.1462	0.1198	5.2537***	-296.4920
	(2.0488)	(1.6588)	(0.9831)	(1.1603)	
T (C_{15})	2.5425·	-0.3837	0.4988	4.8271***	-639.4169
	(1.5229)	(0.8111)	(0.8214)	(1.0272)	
T (C_{23})	3.5097***	-0.9543***	-0.3686	5.4145***	-609.1286
	(0.1429)	(0.0315)	(0.2485)	(0.5441)	

Note: The significant codes: 0 "***" 0.001 "**" 0.01 "*" 0.05 "·"0.1.
The numbers in the parentheses are the standard deviations.

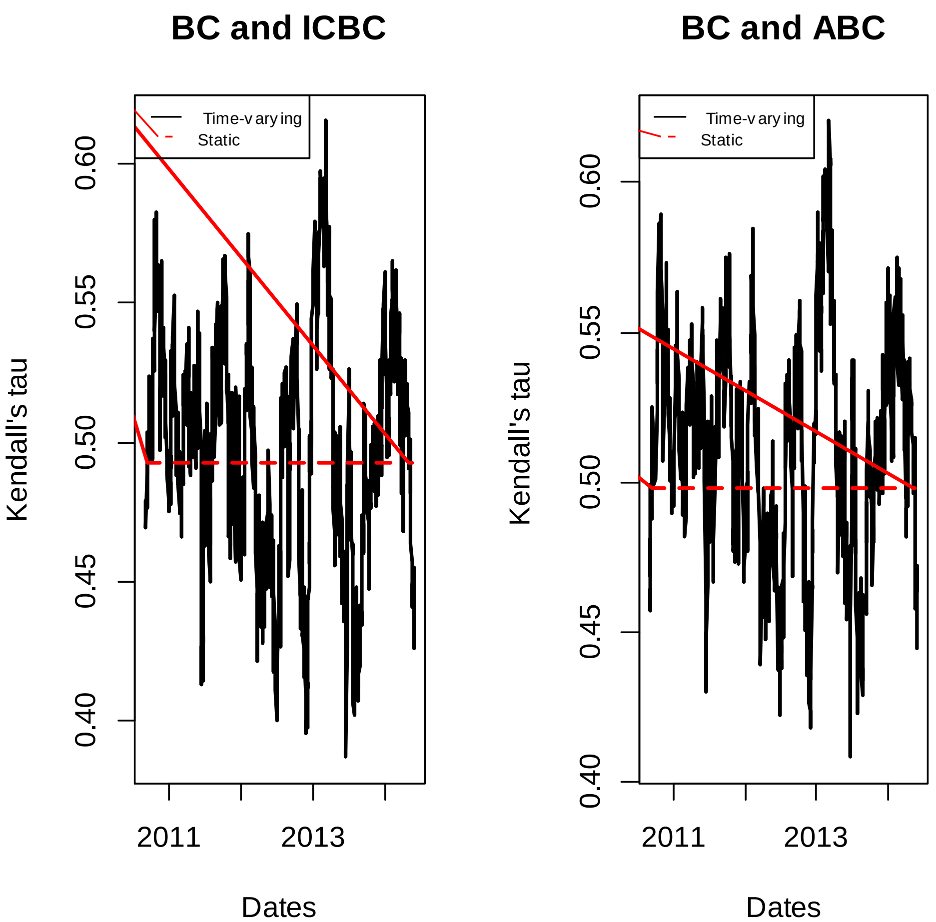

Figure 2. Dependence estimates (Kendall's tau) from T copula

5.3 Portfolio optimization

As stated in the section entitled "Portfolio Optimization Model," the minimum ES (MES) and C-vine copula with Monte Carlo simulation framework is calculated at 90%, 95%, and 99% confidence levels. Table 7 shows the optimal portfolio-based MES values at 90%, 95%, and 99% confidence levels. W1, W2, W3, W4, and W5 represent the weight of PetroChina, ICBC, ABC, BC, and Sinopec, respectively. MES is the minimum expected shortfall under optimal portfolios, and EWES is the equally weighted expected shortfall. As shown in Table 7, all mean weight values are significant at 1% confidence level, and the ES under optimal portfolio is less than the EWES on each confidence level. These results suggest that the optimal portfolios are beneficial in reducing investment risk. The weights of ICBC show the lowest investment allocation at all confidence levels, whereas the weights of PetroChina show the highest investment allocation. The weights of ICBC, BC, and Sinopec gradually increase as the risk increases, which indicates that the three assets largely contribute to the risk. The weights of ABC decrease from 21.14% to 17.81% because investors become more risk tolerant to achieve higher expected returns at higher confidence levels. Table 8 shows the absolute contribution of each share based on CES. Two types of absolute contributions are calculated. One is based on optimal portfolios, and the other is based on equal weight. First, all CES under optimal weights are smaller than equally weighted CES. This result verifies the importance of investment allocation in reducing speculative risk. Second, the absolute contributions of PetroChina and Sinopec are found negative at 90% confidence level, which indicates a portfolio risk diversification. In addition, the absolute contributions of each share have large differences under equally weighted CES. Therefore, more contributions are considered to have more risk. Lastly, the institutions are accurately ranked according to their risks. Based on their absolute contribution under optimal weights at 95% and 99% confidence levels, the ranking of the top five firms is PetroChina, BC, ABC, Sinopec, and ICBC.

Table 7. Optimal Portfolio Based on Minimum Expected Shortfall

	W1	W2	W3	W4	W5	MES	EWES
90%	0.4346***	0.0001**	0.2114***	0.3491***	0.0048***	0.0082***	0.0092***
	(0.0013)	(0)	(0.0015)	(0.0016)	(0.0004)	(0)	(0)
95%	0.4362***	0.0012***	0.2032***	0.3542***	0.0052***	0.0109***	0.0121***
	(0.0016)	(0.0002)	(0.0018)	(0.0021)	(0.0004)	(0)	(0)
99%	0.4231***	0.0106***	0.1781***	0.3806***	0.0075***	0.0184***	0.0203***
	(0.0027)	(0.0007)	(0.0031)	(0.0037)	(0.0005)	(0)	(0)

Note: The significant codes: 0 "***" 0.001 "**" 0.01 "*" 0.05 "·"0.1.
The numbers in the parentheses are the standard deviations.

Table 8. The Absolute Contribution of Each Share Based on Component Expected Shortfall

	PetroChina	ICBC	ABC	BC	Sinopec	SES
EW90	-0.0128	0.7696	0.3371	0.0364	-0.1303	0.0121
Weight90	0.0835	0	0.3537	0.5558	0.0069	0.0113
EW95	0.1836	0.2306	0.2364	0.1196	0.2296	0.0132
Weight95	0.3916	0.0016	0.2375	0.3667	0.0025	0.0131
EW99	0.1682	0.1958	0.1665	0.1697	0.2995	0.0477
Weight99	0.4225	0.0102	0.1699	0.3865	0.0107	0.0390

Note: EW90, EW95, and EW99 represent the equal-weight CES at confidence levels 90%, 95%, and 99%, respectively. Weight90, weight95, and weight99 represent the optimal portfolio obtained from Table 7. SES represents the systemic risk.

6. Conclusions

In this paper, the combination of vine copula-GJR model with ES and CES is used to model the structure dependence of multivariate China stock exchange market and fitted models are used to simulate data and assess the variability of the risk measures.

First, the daily data are analyzed. The marginals are modeled using the skewed student-t and generalized error distributions. Sinopec exhibits leverage effect on the volatility structures, but an asymmetric effect is not observed in PetroChina, ICBC, ABC, and BC markets. Second, the C-vine copula is observed to perform better in terms of AIC. The T, T, T, T, T, Frank, T, Frank, Clayton, and Clayton are the best copulas among the candidates for 5D C-vine structures. The Kendall τ rank correlation and tail dependence are also reported. Third, a high degree of persistence on structure dependence between BC and ICBC as well as BC and ABC is found. Fourth, the risk measures of equally weighted and optimal portfolios are assessed using simulations. The ES and component shortfall are calculated using the simulation from C-vine copula structure instead of that from multivariate distribution because the C-vine copula structure is possibly more credible considering that vine copulas allow different structure dependences between pairs of variables and can also identify tail dependence. Finally, the findings revealed that investors become more risk tolerant to achieve higher expected returns.

Acknowledgment

We would like thank the referees for their constructive comments and suggestions. We would like also to express our gratitude to our colleagues with whom we have had the pleasure of discussing copulas and their applications through the years.

References

Aas, K., Czado, C., Frigessi, A, and Bakken, H. (2009), "Pair-copula construction of multiple dependence. Insurance": Mathematics and Economics 44, pp. 182–198.

Banulescu, G. D., and Dumitrescu, E. I. (2014), "Which are the SIFIs? A Component Expected Shortfall (CES) Approach to Systemic Risk". Journal of banking & finance. http://dx.doi.org/10.1016/j.jbankfin.2014.01.037.

Bedford, T., and Cook, R. M. (2001), "Monte Carlo simulation of vine dependent random variables for applications in uncertainty analysis". In 2001 Proceedings of ESREL2001, Turin, Italy.

Bedford, T., and Cooke, R. M. (2002), "Vines-a new graphical model for dependent random variables". Annals of Statistics 30 (4), pp. 1031–1068.

Czado, C. (2010), "Pair-Copula Constructions of Multivariate Copulas. In P. e. a. Jaworski (Ed.), Copula Theory and Its Applications". Lecture Notes in Statistics, Volume 198, Berlin Heidelberg, 93–109. Springer-Verlag.

Czado, C., Schepsmeier, U., and Min, A. (2012)," Maximum likelihood estimation of mixed C-vines with application to exchange rates". Statistical Modelling 12, pp. 229–255.

Deng, L., Ma, C., and Yang. W. (2011), "Portfolio Optimization via Pair Copula-GARCH-EVT-CVaR Model". Systems Engineering Procedia 2, pp. 171–181.

Fernandez, C., and Steel, M. F. J. (1998), "On bayesian modeling of fat tails and skewness". Journal of the American Statistical Association 93, pp. 359–371.

Glosten, L. R., Jaganathan, R., and Runkle, D. E. (1993), "Relationship between the expected value and the volatility of the nominal excess return on stocks". The Journal of Finance 48, pp. 1779-1801.

Guégan, D. and Maugis, P. A. (2011), "An Econometric Study of Vine Copulas". International Journal of Economics and Finance 2(5), pp. 2–14.

Huang, J. J., Lee, K. J., Liang, H., and Lin, W. F. (2009), "Estimating value at risk of portfolio by conditional copula-GARCH method". Insurance: Mathematics and Economics 45, pp. 315–324.

Liu, J., and Sriboonchitta, S. (2012), "Analysis of Volatility and Dependence between the Tourist Arrivals from China to Thailand and Singapore: A Copula-based GARCH Approach". Uncertainty Analysis in Econometrics with Applications Advances in Intelligent Systems and Computing, 200, pp. 283–294.

Manner, H., and Reznikova, O. (2012), "A survey on time-varying copulas: Specification, simulations and application". Econometric Reviews 31(6), pp. 654–687.

Mendes, B. V. M., and Marques, D. S. (2012), "Choosing an optimal investment strategy: The role of robust pair-copulas based portfolios". Emerging markets reviews 13, pp. 449–464.

Min, A., and Czado, C. (2010), "Bayesian inference for multivariate copulas using pair-copula constructions". Journal of Financial Econometrics 8 (4), pp. 511–546.

Nelsen, R. (2006), “An introduction to copulas”. Springer, 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, pp. 3659–3673.

Patton, A. J. (2006), “Modelling asymmetric exchange rate dependence,” International Economics Review 47 (2), pp. 527–556.

Rockafellar, R. T., and Uryasev, S. (2002), “Conditional Value-at-risk for general loss distributions”. Journal of Banking & Finance. 26, pp. 1443–1471.

Sklar, M. (1959),” Fonctions de r´epartition `an dimensions et leurs marges” .Publ. Inst. Statist. Univ. Paris 8, pp. 229–231.

Sriboonchitta, S., Liu, J., Kreinovich, V., and Nguyen, H. T. (2014), “A Vine Copula Approach to Analyzing Financial Risk and Co-movement of the Indonesian, Philippine and Thailand Stock Markets”. Modeling Dependence in Econometrics, Springer Verlag, Berlin, Heidelberg, pp. 241–254.

Sriboonchitta, S., Liu, J. and Wiboonpongse, A. (2014), “Vine Copula–Cross Entropy Evaluation of Dependence Structure and Financial Risk in Agricultural Commodity Index Returns”. Modeling Dependence in Econometrics, Springer Verlag, Berlin, Heidelberg, pp. 275–287.

Sriboonchitta, S., Nguyen, H. T., Wiboonpongse, A., and Liu, J. (2013), “Modeling volatility and dependency of agricultural price and production indices of Thailand: Static versus time-varying copulas”. International Journal of Approximate Reasoning. 54, pp. 793–808.

Wu, C. C., Chung, H., and Chang, Y. H. (2012),” Economic value of co-movement between oil price and exchange rate using copula-based GARCH models”. Energy Economics 34 (1), pp. 270–282.

Wu, Z., Chen, M., and Ye, W. (2006), “Risk Analysis of Portfolio by Copula-GARCH”. Systems Engineering Theory and Practice 26, pp. 45–52.