Copula Introduction and Its Application in Estimating Portfolio Value at Risk

Zeyu Xiong*

Abstract: In this paper, we mainly discuss copula theory and its applications in estimating value-at-risk(VaR). We present a ARMA-GARCH-copula model to capture dependence between assets returns and to estimate VaR of the portfolio composed of S&P/TSX Composite index and S&P 500 index. The empirical result shows that this approach, compared with traditional approaches, performs good in estimating VaR. Besides, the fact that student t copula fits best to the dependence means that the dependence, contrary to many results, at least in relatively short period, is not asymmetric, and extremes are more likely to happen together in both market.

Keywords: Copula theory; Value at risk; GARCH model; Dependence

1 Introduction

Using joint distribution to model multiple asset returns is always the first part to do analysis in which we need to know the information about the distribution of each asset return and the dependence among them. The traditional approach usually involves estimating both parts(center and covariance matrix) together by assuming that the joint distribution is known such as the most commonly used multivariate normal distribution. In reality, the normality is rarely an adequate assumption. Many empirical works show that such multivariate normal distribution are not able to model asset returns, Ang and Chen [2002] and Longin and Solnik [2001] found that asset returns are more highly correlated in volatile market and in market downturn than in market upturn, and Embrechts [2002] showed that the invariance and symmetry of linear correlation makes it a bad parameter to characterise the asymmetric dependence in downturn and upturn.

Another approach to model multivariate distribution concerns estimating univariate distribution and dependence separately. The theory of copula is a very useful tool to model the dependence because it does not require the assumption of joint normality and it provide a flexible way to decompose any n-dimensional joint distribution into n marginal distribution and a copula function. Copula is a function that glue multivariate distribution function with their marginal distribution functions, we can construct flexible multivariate distributions with different margins and dependent structures, and we will give a formal definition later.

The word copula was first employed in a mathematical sense by Sklar [1959] in the theorem 2.1 which laid the foundation for later researches. But the study of copula and its applications in Statistics and Finance is a rather modern phenomenon in the past several decades due to the efficiency of computation. The book Nelsen [2006] presented a formal and complete introduction to the copula theory. Schweizer and Wolff [1981] showed the superiority of copula by linking 3 famous concordance measures—the Kendall's tau, the Spearman's rho and the Gini indice—with copula. Lindskog and RiskLab [2000] offered us the copula of elliptical distribution which

*Corresponding author, Wuhan Univserity, China

shares lots of tractable properties with multivariate normal distribution. Genest and Mackay [1986] and Frees and Valdez [1998] offered another simple class of copula—Archimedean copula, the simplicity with which they can be formed and other specials properties enable them to frequently appear in copula applications. Bouyé et al. [2007] provided general examples of applications of copula in finance.

In this paper, we mainly discuss copula application in risk management, to be exact, in estimating value-at-risk(VaR). Although VaR is a simple measure, it is not easily to be estimated accurately. We present an ARMA-GARCH-copula approach to estimate VaR of two stock market indices—S&P/TSX Composite index and S&P 500, for each market we recover marginal distribution after eliminating autocorrelation and heteroskedasticity, then we fit different copulas to innovation pairs and generate new pairs according to the dependence, finally we can calculate VaR at different confidence level through the simulated return series. Copula does have a lot of uses, Jondeau and Rockinger [2006] used similar method to show that the time-varying conditional dependency increases between European markets subsequent to movements in the same direction. Cherubini and Luciano [2001] and Bouyé et al. [2007] focus on applications in credit default and derivatives pricing. Palaro and Hotta [2005] and Huang et al. [2009] also discussed the application of conditional copula in estimating VaR of a portfolio. However, instead of eliminating autocorrelation with fixed ARMA model, we choose suitable ARMA model in each stage according to BIC criterion, and by assuming the innovations follow the skewed generalized error distribution(SGED), we are able to take leptokurtosis of return series into consideration. We applied various copulas with different marginal distributions to estimate one-day ahead VaR of the portfolio. In addition, we use 2 approaches to evaluate the copula model, one shows that the dependence is not asymmetric in a relatively short period(i.e. 250 trading days), and the other shows the superiority of copula in estimating VaR by comparing its accuracy with traditional methods(historical simulation, variance-covariance). Generally, Clayton copula and rotated Gumbel copula give suitable VaR of the portfolio at all levels, while student t copula fits the innovation pairs best.

This paper is organized as follows. Section 2 discusses some copula theory. Section 3 presents copula's correlation with commonly used dependence measures. Section 4 presents Elliptical copula family and Archimedean copula family. Section 5 presents the procedure to model different parts and empirical results, followed by conclusion in section 6.

2 Copula Theory

2.1 Some definitions

Definition 2.1 (Nelsen [2006]) *A N-dimensional copula is a function C with the following properties:*

1. *Dom $C = I^N = [0,1]^N$, Ran $C = I = [0,1]$;*
2. *C is grounded*[1] *and N-dimensional increasing*[2]*;*

[1] Suppose function H is defined on $S^N = S_1 * S_2 * \dots S_n$, if there exists $a_i, a_j \in S^N$ so that $H(\boldsymbol{x_1}, a_i, \boldsymbol{y_1}) = 0 = H(\boldsymbol{x_2}, a_j, \boldsymbol{y_2})$ for all $(\boldsymbol{x_1}, \boldsymbol{y_1}), (\boldsymbol{x_2}, \boldsymbol{y_2})$ in S^{N-1} , then we say H is N-dimensional grounded.

[2] C is N-increasing if the C-volume of all N-boxes whose vertices lie in I^N are positive, or equivalently if we have

$$\sum_{i_1=1}^{2} \cdots \sum_{i_N=1}^{2} (-1)^{(i_1+i_2+\cdots+i_N)} C(u_{1,i_1}, \dots, u_{N,i_N}) \geq 0$$

for all u in I^N with $u_{n,1} \leq u_{n,2}$.

3. *C have margins C_n which satisfy $C_n(u) = C(1, \dots u, 1, \dots, 1) = u$ for all u in I.*

Theorem 2.1 (Sklar [1959]) *Let F be an N-dimensional distribution function with continuous margins $F_1, \dots, F_N$. Then there exists a copula C such that for all $x_1, \dots, x_N$:*

$$F(x_1, \dots, x_N) = C(F_1(x_1), \dots, F_N(x_N)) = C(u_1, \dots, u_n) \tag{2.1}$$

where $u_i = F_i(x_i)$ follows uniform distribution[3].
If $F_1, \dots, F_N$ are continuous, then C is unique. Conversely, if C is a copula and $F_1, \dots, F_N$ are distribution function, then the function F defined by (2.1) *is a joint distribution function with margins $F_1, \dots, F_N$.*

Usually, we can directly observe univariate distribution from data, but we have to struggle a lot to get information about the dependence structure. Simple but powerful, (2.1) implies that we can link any kinds of n univariate distribution with any copula to get any well-defined multivariate distribution. Therefore, by constructing appropriate dependence structure, or in other words, copula, we are able to get more flexible multivariate distribution that fits the data.

Definition 2.2 (Fréchet [1951]) *Consider the function M^n, W^n, Π^n defined on $[0,1]^n$ as follows:*

$$\begin{aligned} M^n(\boldsymbol{u}) &= min(u_1, \dots, u_n) \\ W^n(\boldsymbol{u}) &= max(u_1 + \dots + u_n - n + 1, 0) \\ \Pi^n(\boldsymbol{u}) &= u_1 \dots u_n \end{aligned} \tag{2.2}$$

The functions M^n, Π^n are n-copulas for all $n \geq 2$ whereas the function W^n is not a copula for any $n \geq 3$.

Theorem 2.2 (Fréchet [1957]) *For any n-dimension copula C and $\boldsymbol{u} = (u_1, \dots, u_n) \in [0,1]^n$, the following inequality always holds:*

$$W^n(\boldsymbol{u}) \leq C(\boldsymbol{u}) \leq M^n(\boldsymbol{u}) \tag{2.3}$$

proof.

$$\begin{aligned} & \forall \boldsymbol{u} = (u_1, \dots, u_n) \in \mathrm{Dom}C \\ & C(\boldsymbol{u}) \leq C(u_1, 1, \dots, 1) \leq u_1 \\ & \vdots \\ & C(\boldsymbol{u}) \leq C(1, \dots, 1, u_n) \leq u_n \\ \Leftrightarrow & C(\boldsymbol{u}) \leq min(u_1, \dots, u_n) = M^n(\boldsymbol{u}) \\ & \forall \boldsymbol{u} \in [u_1, 1] \times \dots \times [u_n, 1] \\ & |C(1, \dots, 1) - C(\boldsymbol{u})| \leq \sum_{i=1}^{n} |1 - u_i| \\ \Leftrightarrow & 1 - C(\boldsymbol{u}) \leq n - \sum_{i=1}^{n} u_i \\ \Leftrightarrow & C(\boldsymbol{u}) \geq \sum_{i=1}^{n} u_i - n + 1 \quad \text{and} \quad C(\boldsymbol{u}) \geq 0 \\ \Leftrightarrow & C(\boldsymbol{u}) \geq max(u_1 + \dots + u_n - n + 1, 0) = W^n(\boldsymbol{u}) \end{aligned}$$

[3] $P(u_i < u) = P(F_i(x_i) < u) = P(x_i < F_i^{-1}(u)) = F_i^{-1}(F_i(u)) = u$

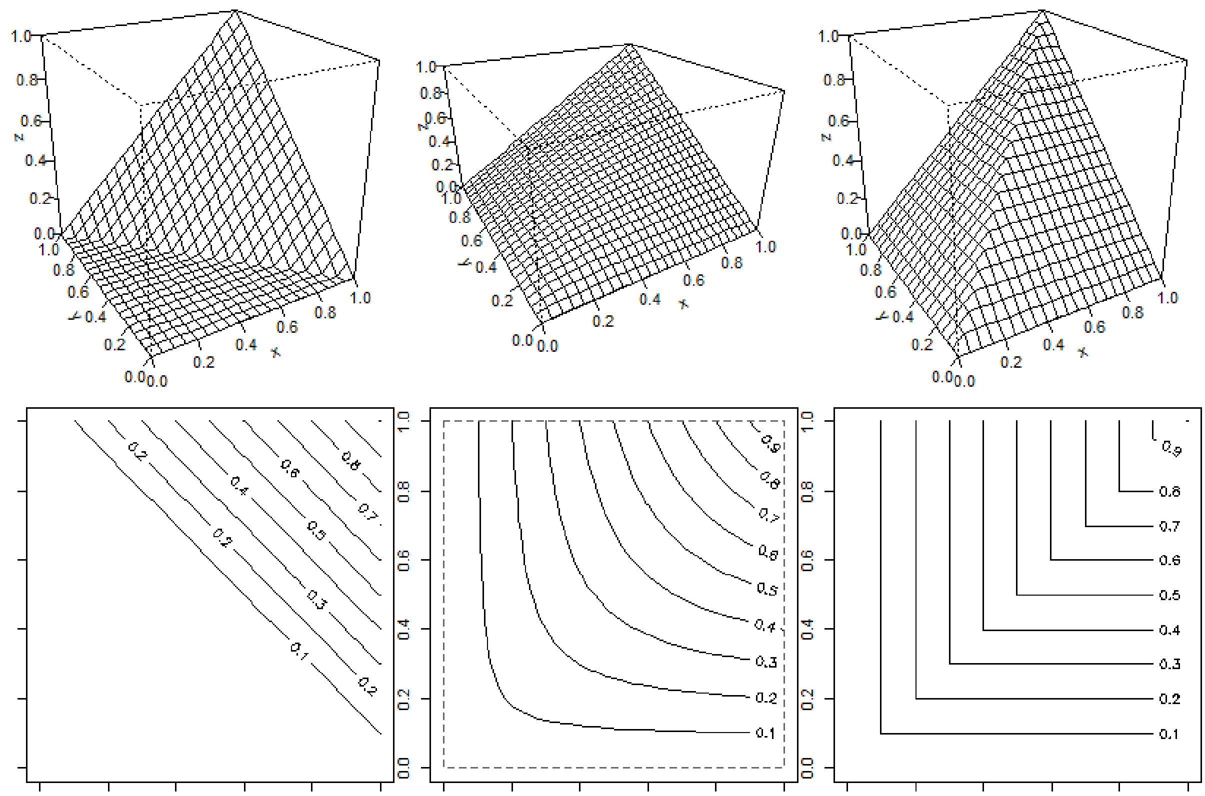

Figure 1: Lower Fréchet, Clayton, upper Fréchet copulas

Here we show a bivariate copula:$C(x, y) = (x^{-\theta} + y^{-\theta})^{-1/\theta}$ distribution[4], where $\theta = 0.9$ together with Fréchet bounds, we can find that copula shares many features with multivariate distribution function.

2.2 Estimation—maximum likelihood method

Let $\boldsymbol{\alpha} = (\boldsymbol{\alpha_1}, \ldots, \boldsymbol{\alpha_N}, \boldsymbol{\alpha_C})$ denote the vector of all parameters, where $\boldsymbol{\alpha_1}, \ldots, \boldsymbol{\alpha_N}$ is the vector of parameters of corresponding univariate distribution $F_1, \ldots, F_2$, $\boldsymbol{\alpha_C}$ is the vector of parameters of N-dimensional copula. Suppose that $\{(x_{1,t}, \ldots, x_{N,t})\}_{t=1}^{T}$ is a sample of size T, then by Sklar's theorem we have:

$$F(x_1, \ldots, x_N; \boldsymbol{\alpha}) = C(F_1(x_1; \boldsymbol{\alpha_1}), \ldots, F_N(x_N; \boldsymbol{\alpha_N}); \boldsymbol{\alpha_C}) \tag{2.4}$$

Then the density function f of F is:

$$\begin{aligned} f(x_1, \ldots, x_N; \boldsymbol{\alpha}) &= c(F(x_1; \boldsymbol{\alpha_1}), \ldots, F(x_N; \boldsymbol{\alpha_N}); \boldsymbol{\alpha_C}) \prod_{n=1}^{N} f_n(x_n; \boldsymbol{\alpha_n}) \\ &= c(u_1, \ldots, u_N; \boldsymbol{\alpha_C}) \prod_{n=1}^{N} f_n(x_n; \boldsymbol{\alpha_n}) \end{aligned}$$

Where $u_n = F_n(x_n; \boldsymbol{\alpha_n})$ for $n = 1, \ldots, N$;

[4]It belongs to Clayton copula which we will discuss later.

$c(u_1,\ldots,u_N;\boldsymbol{\alpha_C}) = \frac{\partial C(u_1,\ldots,u_N;\boldsymbol{\alpha_C})}{\partial u_1\ldots\partial u_N}$;
Therefore, we can get the log likelihood function of sample $(x_{1t},\ldots,x_{Nt}), t = 1.2,\ldots,T$:

$$\ln L(x_1,\ldots,x_N;\boldsymbol{\alpha}) = \sum_{t=1}^{T}(\sum_{n=1}^{N}\ln f_n(x_{nt};\boldsymbol{\alpha_n}) + \ln c(F_1(x_{1t};\boldsymbol{\alpha_1}),\ldots,F_N(x_{Nt};\boldsymbol{\alpha_N});\boldsymbol{\alpha_C}))$$

Thus the maximum likelihood estimate $\boldsymbol{\alpha}$ is given by:

$$\boldsymbol{\alpha} = \underset{\boldsymbol{\alpha}}{\operatorname{argmax}} \ln L(x_1,\ldots,x_N;\boldsymbol{\alpha}) \tag{2.5}$$

Although we can estimate all parameters in (2.5) at the same time, it's inconvenient and computationally inefficient for us to find the optima. The characteristic of copula makes itself suitable for multi-stage estimation. Therefore, Patton [2006] presented that we may use 2-stage estimation:

$$\hat{\boldsymbol{\alpha}}_{\boldsymbol{1}} = \underset{\boldsymbol{\alpha_1}}{\arg\max}\sum_{t=1}^{T}\ln f_1(x_1;\boldsymbol{\alpha_1})$$

$$\vdots$$

$$\hat{\boldsymbol{\alpha}}_{\boldsymbol{N}} = \underset{\boldsymbol{\alpha_N}}{\arg\max}\sum_{t=1}^{T}\ln f_N(x_N;\boldsymbol{\alpha_N})$$

Second step:

$$\hat{\boldsymbol{\alpha}}_{\boldsymbol{C}} = \underset{\boldsymbol{\alpha_C}}{\arg\max}\sum_{t=1}^{T}c(F(x_1;\hat{\boldsymbol{\alpha}}_{\boldsymbol{1}}),\ldots,F(x_N;\hat{\boldsymbol{\alpha}}_{\boldsymbol{N}});\boldsymbol{\alpha_C})$$

That's to say, we firstly estimate parameters $\boldsymbol{\alpha_n}, n = 1,2,\ldots,N$ for marginal distributions, then we take them as already known parameters in copula function to estimate $\boldsymbol{\alpha_C}$. This method could greatly simplify the numerical procedure, and Patton [2006] showed 2-stage estimator is consistent and normal.

3 Measure of association

Copula is the dependence structure between random variables, naturally, copula can be used in the study of dependence or association between random variables. Though there are various ways to measure dependence like linear correlation, Kendall's tau, Spearman's rho, they all have their own pros and cons and usually they do not contain adequate information about the dependence. Moreover, it is impressive that these traditional dependence measures can all be expressed with copula Schweizer and Wolff [1981].

We usually consider association between 2 objectives. Let $(x_i,y_i),(x_j,y_j)$ denote 2 observations from a random vector (X,Y), we say that (x_i,y_i) and (x_j,y_j) are concordant if $(x_i-x_j)(y_i-y_j) > 0$ and discordant if $(x_i-x_j)(y_i-y_j) < 0$. It is a natural definition, if we find that x_i is smaller(larger) than x_j and, at the same time, y_i is smaller(larger) than y_j, we are more confident that random variable X and Y tend to move in the same direction and vice versa.

3.1 Kendall's tau

Kendall's tau is defined as the probability of concordance minus probability of discordance.

Definition 3.1 (Kruskal [1958]) *Let $(X_1, Y_1), (X_2, Y_2)$ be i.i.d. random vectors, each with joint distribution function H. Then the population version of Kendall's tau is:*

$$\tau = \tau_{X,Y} = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0] \tag{3.1}$$

Let $\{(x_i, y_i)\}_{i=1}^n$ be n distinct observations from the vector (X, Y), then there are $\binom{n}{2}$ pairs of observations need to be signed as concordant or discordant. Let c denote the number of concordant pairs and d the number of discordant pairs, then the sample version of Kendall's tau is:

$$\tau = \frac{c - d}{c + d} = \frac{c - d}{\binom{n}{2}} \tag{3.2}$$

Theorem 3.1 (Nelsen [2006]) *Let (X_1, Y_1) and (X_2, Y_2) be independent vectors of random variables with joint distribution H_1 and H_2 respectively, with common margins F and G. Let C_1, C_2 be the copulas of (X_1, Y_1) and (X_2, Y_2) respectively, so $H_i(x, y) = C_i(F(x), G(y))$ $i = 1, 2$. Then the Kendall's tau of (X, Y) is:*

$$\tau = \tau(C_1, C_2) = \tau(C_1, C_2) = 4 \iint_{I^2} C_1(u, v) \mathrm{d}C_2(u, v) - 1 \tag{3.3}$$

where $u = F(x), v = G(y)$

The proof can be seen on $p159$ of Nelsen [2006]. Furthermore, τ is easily evaluated for pairs of the basic copulas M, W, Π. Because the support of M and W is the diagnal $v = u$ and $v = 1 - u$ respectively, it follows that if g is an integrable function whose domain is I^2, then

$$\begin{aligned} \tau(g, M) &= \iint_{I^2} g(u, v) \mathrm{d}M(u, v) = \int_0^1 g(u, u) \mathrm{d}u \\ \tau(g, W) &= \iint_{I^2} g(u, v) \mathrm{d}W(u, v) = \int_0^1 g(u, 1 - u) \mathrm{d}u \end{aligned} \tag{3.4}$$

3.2 Spearman's rho

Definition 3.2 (Kruskal [1958]) *Let $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3)$ be independent random vectors with a common joint distribution function H and copula C. Then the population Spearman's rho is defined to be proportional to the probability of concordance minus discordance for two pairs—(X_1, Y_1) and (X_2, Y_3):*

$$\rho_{X,Y} = 3(P[(X_1 - X_2)(Y_1 - Y_3) > 0] - P[(X_1 - X_2)(Y_1 - Y_3) < 0]) \tag{3.5}$$

Let $\{(x_{1,i}, y_{1,i})\}_{i=1}^m, \{(x_{2,i}, y_{2,i})\}_{i=1}^n, \{(x_{3,i}, y_{3,i})\}_{i=1}^n$ be distinct observations from vector (X, Y), then there are $m \times n$ pairs of combination. Let c denote the number of concordant pairs and d the number of discordant pairs, then the sample version of Kendall's tau is:

$$\rho = 3\frac{c - d}{mn}$$

Note that while the joint distribution function of (X_1, Y_1) is $H(x, y)$, the joint distribution function of (X_2, Y_2) is $F(x)G(y)$ since X_2 and Y_3 is independent. Therefore, the copula between X_1 and Y_1 is C while the copula between X_2 and Y_3 is Π. So Spearman's rho is:

$$\rho_{X,Y} = 3\tau_{C,\Pi} = 12 \iint_{I^2} C(u, v) \mathrm{d}u\mathrm{d}v - 3$$

Although both Kendall's tau and Spearman's rho measures the difference between probability of concordance and discordance with given copula, they are quite different in many cases. In fact, they hold some inequalities like $-1 \leq 3\tau - 2\rho \leq 1$, $\frac{1+\rho}{2} \geq (\frac{1+\tau}{2})^2$ and $\frac{1-\rho}{2} \geq (\frac{1-\tau}{2})^2$. We won't talk too much about their association and other concordance measures like Gini's g and Blomqvist's β, please see Nelsen [2006] for details. And we shall see later that Archimedean copulas have a deep correlation with Kendall's tau and Spearman's rho, calculating these measures could greatly simplify the estimation of their parameters.

4 Copula Families

4.1 Empirical copula

Definition 4.1 (Empirical copula, Deheuvels [1979]) *Let $\mathbb{X} = \{(x_1^t, x_2^t, \ldots, x_N^t)\}_{t=1}^T$ denote a sample set, then the empirical copula distribution is given by:*

$$\boldsymbol{C}_e(t_1/T, \ldots, t_N/T) = \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}_{x_1^t \leq x_1^{(t_1)}, \ldots, x_N^t \leq x_N^{(t_N)}} \tag{4.1}$$

where $x_n^{(t)}$ are the order statistics and $1 \leq t_1, \ldots, t_N \leq T$.

Figure 2 shows an empirical copula distribution of a part of the data set we will consider in section 5. As we can see, the empirical copula become less smooth on the district near the diagonal district. However, by comparing the middle in figure 1 with figure 2, we are still confident that the former can be a good estimation of the latter through a suitable parameter.

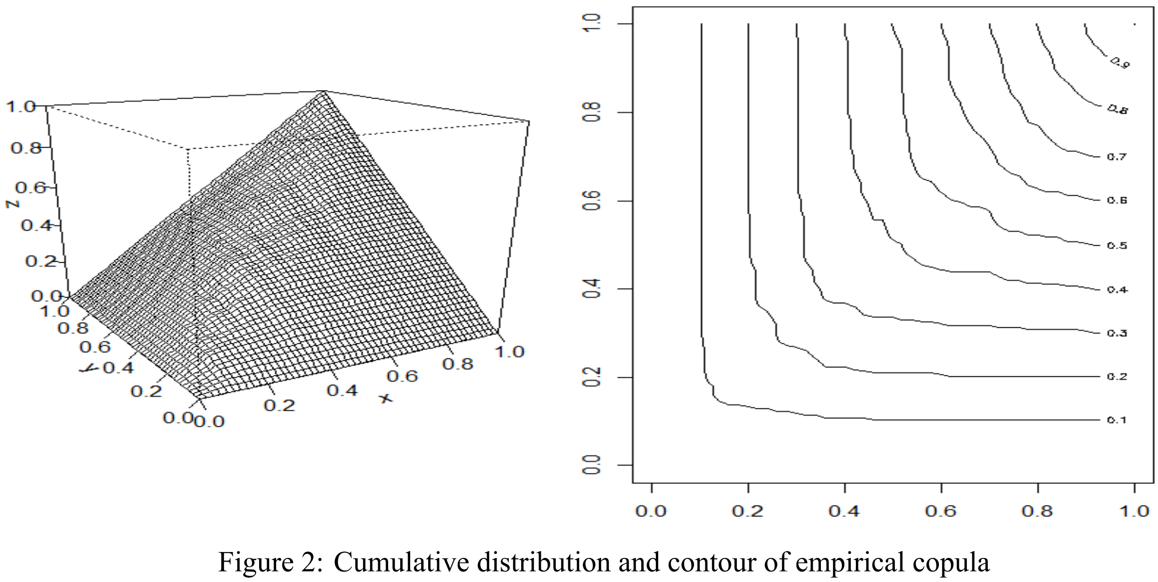

Figure 2: Cumulative distribution and contour of empirical copula

4.2 Elliptical copula

The class of elliptical distributions offers us a variety of multivariate distributions like multivariate gaussian distribution and multivariate t distribution which has lots of tractable properties. Elliptical copula is the copula of elliptical distribution.

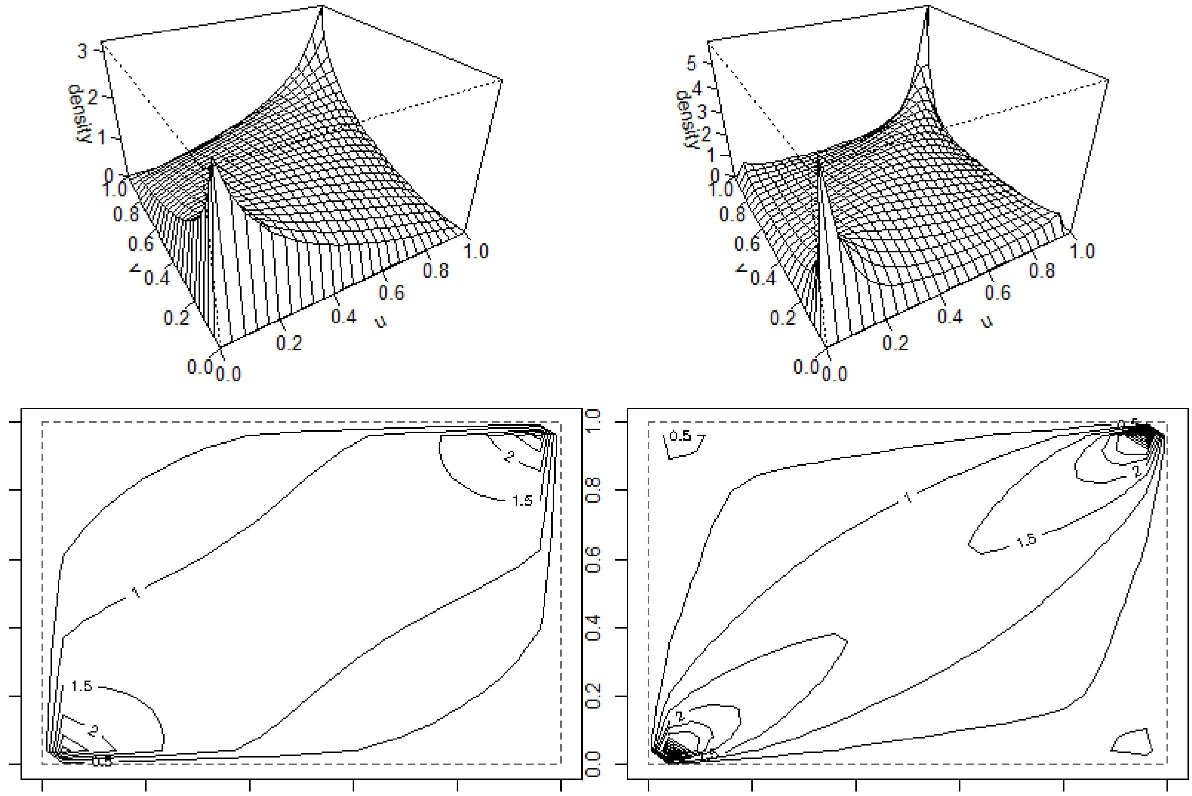

Figure 3: Density distribution and contour of bivariate gaussian copula and t copula($\tau = 1/3$)

Definition 4.2 (Multivariate gaussian copula) *Let ρ be a symmetric, positive definite matrix with diagonal elements equal to 1, and Φ_ρ the standardized multivariate normal distribution with correlation matrix ρ. The multivariate gaussian copula is then defined as follows:*

$$C(u_1, \ldots, u_N; \rho) = \Phi_\rho(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_N))$$

In the bivariate case, the copula expression can be written as:

$$C(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\{-\frac{s^2 - 2\rho st + t^2}{2(1-\rho^2)}\} \mathrm{d}s\mathrm{d}t$$

Definition 4.3 (Multivariate student t copula) *Let ρ be a symmetric, positive definite matrix with diagonal elements equal to 1, $\Phi_{\rho,\upsilon}$ the standardized multivariate Student t distribution with υ degrees of freedom and correlation matrix ρ. The multivariate student t copula is then defined as follows:*

$$C(u_1, \ldots, u_N; \rho, \upsilon) = t_{\rho,\upsilon}(t_\upsilon^{-1}(u_1), \ldots, t_\upsilon^{-1}(u_N))$$

In the bivariate case, the copula expression can be written as:

$$C(u, v) = \int_{-\infty}^{t_\nu^{-1}(u)} \int_{-\infty}^{t_\nu^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \{1 + \frac{s^2 - 2\rho st + t^2}{\nu(1-\rho^2)}\}^{-(\nu+2)/2} \mathrm{d}s\mathrm{d}t$$

From figure 3, we are able to find that both copulas are symmetric, so the dependence structure modelled by them is symmetric which means that extremes in both directions will happen with the same probability. However, being similar to t distribution, the density of t copula in both tails is heavier than gaussian copula, so variables with t copula are easier to fall into extremes than gaussian copula even though the concordance measures τ are the same. This is why copula is more superior in characterising dependence, measures like Kendall's tau only contain partial information about the dependence, while copula preserves all the information.

4.3 Archimedean copula

Elliptical copulas can be derived from elliptical distributions using Sklar's theorem. Though they have many tractable properties, there are however, drawbacks: elliptical copulas do not have closed form expressions and are restricted to have radial symmetry[5]. While in many cases, we hope to model asymmetric dependence where there is a stronger dependence in big losses than big gains. Such asymmetry can not be modelled by elliptical copulas, here comes the Archimedean copulas.

There are a number of reasons to study Archimedean copulas—Archimedean copulas allow for various dependence structure including asymmetric and symmetric; all frequently used Archimedean copulas have closed form expressions; and they are not derived from multivariate distribution functions using Sklar's theorem which means that they need further conditions to ensure that they are proper n-copulas.

Here we only discuss 3 typical Archimedean copulas—Gumbel, Clayton and Frank copula, there are still other Archimedean copulas, we recommend Genest and Mackay [1986] Genest and Rivest [1993] and Joe [1997] for further references.

Definition 4.4 (Pseudo-inverse function) *Let ϕ be a continuous, strictly decreasing function from I to $[0, \infty]$ such that $\phi(1) = 0$. The pseudo-inverse of ϕ is ϕ^{-1} given by:*

$$\phi^{[-1]}(t) = \begin{cases} \phi^{-1}(t), & \text{if } 0 \leq t \leq \phi(0) \\ 0, & \phi(0) \leq t \leq \infty. \end{cases} \tag{4.2}$$

so if $\phi(0) = \infty$, then $\phi^{[-1]} = \phi^{-1}$

Definition 4.5 (Archimedean copula) *Let ϕ be a continuous, strictly decreasing function from I to $[0, \infty]$ such that $\phi(1) = 0$, and $\phi^{[-1]}$ be the pseudo-inverse function defined in* (4.2)*. Let C be the function from I^N to I given by:*

$$C(u_1, \ldots, u_N) = \phi^{[-1]}(\phi(u_1) + \cdots + \phi(u_N)) \tag{4.3}$$

Then C satisfies the conditions in definition 2.1 is Archimedean copula, and ϕ is called the generator of the copula.

Archimedean copulas preserve several desired properties(C is symmetric[6], associative[7], etc.) and they allow to extend the modelling of bivariate series to n-variate series, though at the same time, suffering from lack of free parameter choice in high dimension. Moreover, Archimedean copulas simplify calculus, (3.3) is given by

$$\tau = 1 + 4 \int_0^1 \frac{\varphi(u)}{\varphi'(u)} \mathrm{d}u$$

[5] $C = \hat{C}$, where $\hat{C}$ is the survival copula, details can be seen in Nelsen [2006], p32

[6] $C(u, v) = C(v, u)$

[7] $C(u_1, C(u_2, u_3)) = C(C(u_1, u_2), u_3)$

Table 1
Parameter of copulas and their correlation

Copula	$\varphi(u)$	$\boldsymbol{C}()$	Parameter range(α)	Kendall τ
$\Pi(\boldsymbol{u})$	$-\ln u$	$u_1 u_2 \ldots u_N$		0
Gumbel	$(-\ln u)^\alpha$	$\exp\left(-[\sum_{n=1}^{N}(-\ln u_n)^\alpha]^{\frac{1}{\alpha}}\right)$	$[1,\infty)$	$1-\alpha^{-1}$
Clayton	$\frac{1}{\alpha}(u^{-\alpha}-1)$	$(\sum_{n=1}^{N} u_n^{-\alpha} - N + 1)^{-\frac{1}{\alpha}}$	$(-1,0)\cup(0,\infty)$	$\frac{\alpha}{\alpha+2}$
Frank	$-\ln(\frac{e^{-\theta\alpha}-1}{e^{-\theta}-1})$	$-\frac{1}{\alpha}\ln(1+\frac{\prod_{n=1}^{N}(e^{-\alpha u_n}-1)}{(e^{-\alpha}-1)^{N-1}})$	$(-\infty,+\infty)\setminus 0$	$\frac{4}{\alpha}(D_1(-\alpha)-1)-1$

$D_k(x)$ is the Debye function given by $D_k(x) = \frac{k}{x^k}\int_0^x \frac{t^k}{e^t-1}\mathrm{d}t$ for any positive integer k

And it can be simplified further given the generator $\phi(u)$. The following are some classical Archimedean copulas' generators, parameter constraints and their relationship with Kendall's tau:

Note that the 5th column gives the relation between Archimedean copula's parameter and Kendall's tau, so it offers us another way to estimate the parameter α in Archimedean copula. To be exact, if we observe k pairs of sample from the variable X and Y, then we can calculate sample Kendall's tau by (3.2), finally we can inverse it according to the 5th column in table 1 to get parameter α. It is a faster way to estimate the parameter, while it is easier to be polluted by outliers and we will not use it in later part.

Figure 4 shows the density distribution and contour of Gumbel, Clayton, Frank copula respectively. We can see that Gumbel copula has heavier upper tail, while Clayton copula has heavier lower tail, so they can be used in modelling asymmetric dependence structure, thus they and their extensions are extremely useful in characterising tail dependence, and Joe [1997] presented using tail coefficient to study dependence between extreme values. Unlike Gumbel and Clayton copulas which can only capture co-movement in the same direction, or non-negaitve correlation, Frank copula can be used in modelling variables with negative correlation. However, this flexibility doesn't make Frank copula a good option in modelling dependence in financial market since its density in both tails is even lighter than the gaussian copula, naturally it is hardly used in the market full of leptokurtosis, and we won't show its result in empirical part.

Figure 5 gives the contour of joint distribution with Kendall's tau equalling to $1/3$ and margins following standard normal distribution. The joint distributions constructed by different copulas have much difference even though they have the same concordance measure(τ) and the same marginal distribution. Thus, simple concordance measures together with univariate distribution is not able to describe multivariate distribution, with the help of copula, we are more flexible in modelling multivariate distribution.

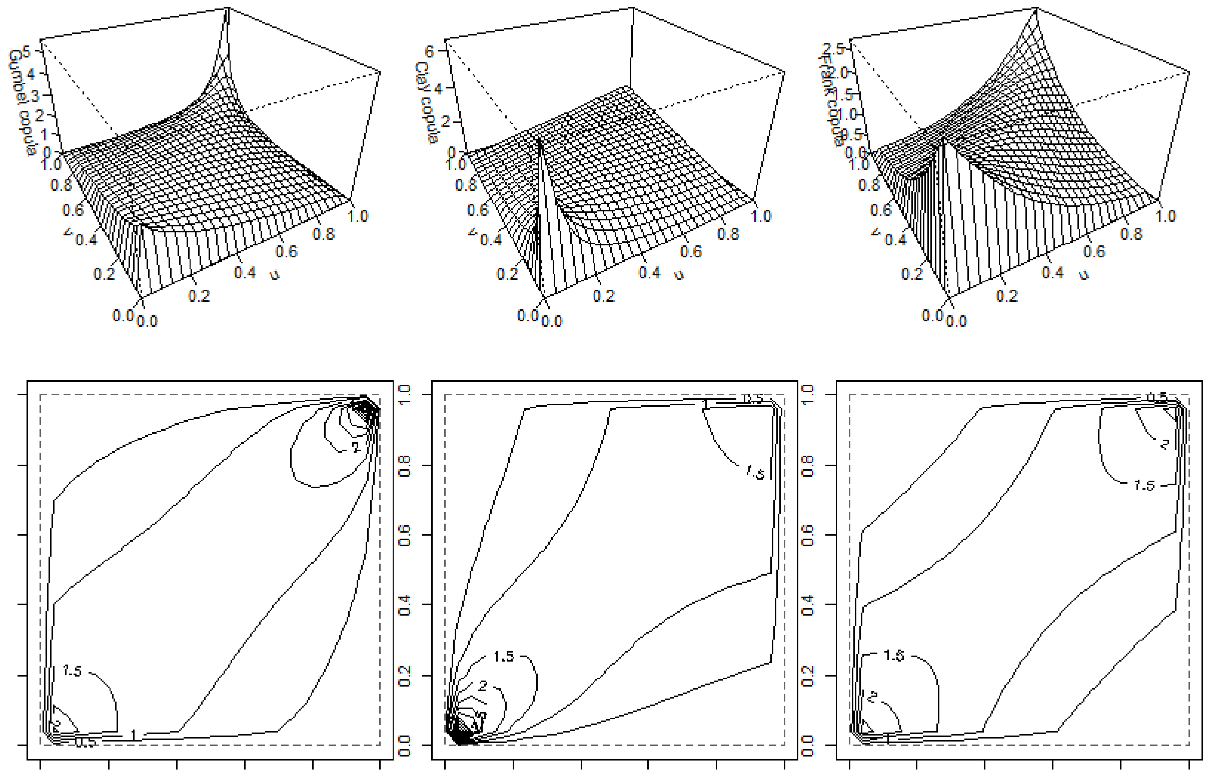

Figure 4: Density distribution and contour of bivariate Gumbel, Cayton, Frank copula(τ=1/3)

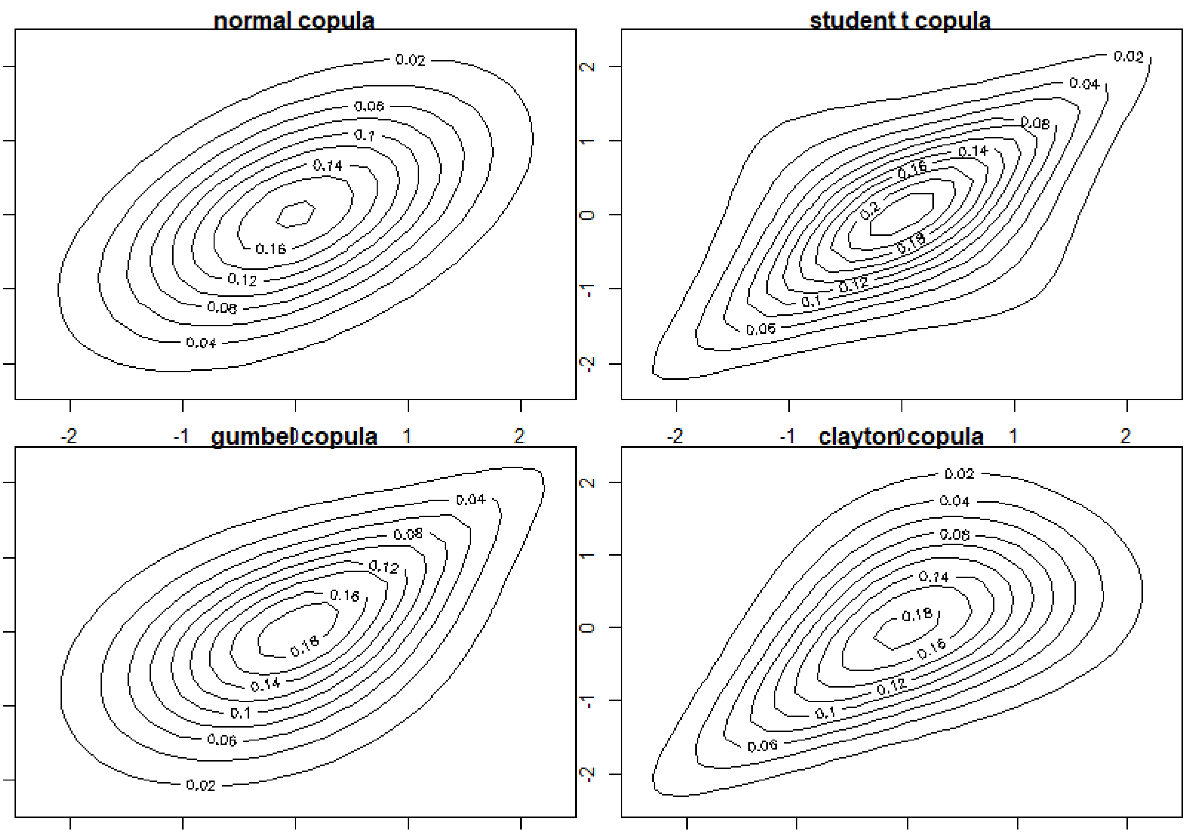

Figure 5: Density contour of variables with the same margins and different copulas

5 Model and Application

We aim at estimating VaR using GARCH-copula methodology and to check the performance of VaR. VaR is a forecast of a given percentile, usually in the lower tail, of the distribution of returns on a portfolio over some period. The VaR of a portfolio at time t, with confidence level $1-\alpha$ is defined as:

$$VaR_t(\alpha) = \text{inf}\{s : F_{p,t}(s) \geq \alpha\} \quad \alpha \in (0,1)$$

where $F_{p,t}$ is the distribution function of the portfolio return $R_{p,t}$ at time t. Therefore, we can rewrite it as $P(R_{p,t} \leq VaR_t(\alpha)) = \alpha$, which means that we are $100(1-\alpha)\%$ confident that the loss at the time t will not exceed $VaR_t(\alpha)$. Not only does it help us understand the concept of VaR, but it offers us a good way to check the accuracy of VaR. By checking the number of observation exceeding VaR in the out-of-sample test, we are able to judge the performance of the methodology. And Kupiec [1995] defines a statistic to help us do statistical inference which we will use in the out-of-sample test.

In this paper, we model the return series of portfolio composing of S&P/TSX composite index and S&P 500 with equal weight in the following sequence:

- An exploratory data analysis is done.
- The ARMA-GARCH approach are fitted in each period for both series to get innovation pairs.
- Various bivariate copulas are fitted to the innovation pairs.
- Some evaluation criteria and comparisons are done to select suitable models.

5.1 Data Description

We consider the portfolio composing of S&P/TSX composite index and S&P 500 with the same weight though the approach is flexible in other weight assumptions. After eliminating observations when a holiday occurred in one country, the data contains 3060 daily log returns of closing price, ranging from Jan. 1^{st} 2006 to May 31^{th} 2018.

Table 2 shows that both series are asymmetric and have large kurtosis, the Shapiro-Wilk tests confirm the violation of normal distribution. Ljung-Box tests suggest that both series need to be corrected for serial correlation and arch effect, and in figure 6 we can see the volatility cluster in which large absolute returns tend to follow large absolute returns and the same for small returns, these are 2 reasons why we include ARMA-GARCH into the marginal model.

5.2 Modelling the margin

ARMA-GARCH model and its extensions have been popular in modeling univariate return series in many authors, and with no loss of simplicity and practicability, $ARMA(p,q)-GARCH(1,1)$ will be used here. In particular, we select suitable p, q order according to BIC criterion, and we assume that innovations follow skewed generalized error distribution(SGED) to model the skewness and kurtosis.

$$\begin{aligned}
X_t &= \phi_0 + \sum_{i=1}^{p} \phi_i X_{t-i} + a_t - \sum_{j=1}^{q} \theta_j a_{t-j} \\
a_t &= \sigma_t \varepsilon_t, \quad \varepsilon_t \sim SGED(\xi, \nu) \\
\sigma_t^2 &= \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2
\end{aligned} \tag{5.1}$$

Table 2
Descriptive statistic

Statistic	TSE	S&P 500
Annulized mean	0.842%	1.756%
Annulized standard deviation	17.378%	18.751%
Minimum	-9.788%	-9.470%
Median	0.072%	0.060%
Maximum	9.370%	10.246%
Sample skewness	-0.671	-0.624
Sample kurtosis	14.279	12.968
$QW(10)$	65.5^a	80.1^a
$QW^2(10)$	3808^a	2699^a
SK	0.884^a	0.880^a

$QW(K)$ is the Ljung-Box statistics for series with K lags, $QW^2(K)$ is for the squared series.
SK is the Shapiro-wilk normality statistic.
In this and following tables, significance is denoted by the superscripts at the $1\%(^a), 5\%(^b)$ and $10\%(^c)$ levels.

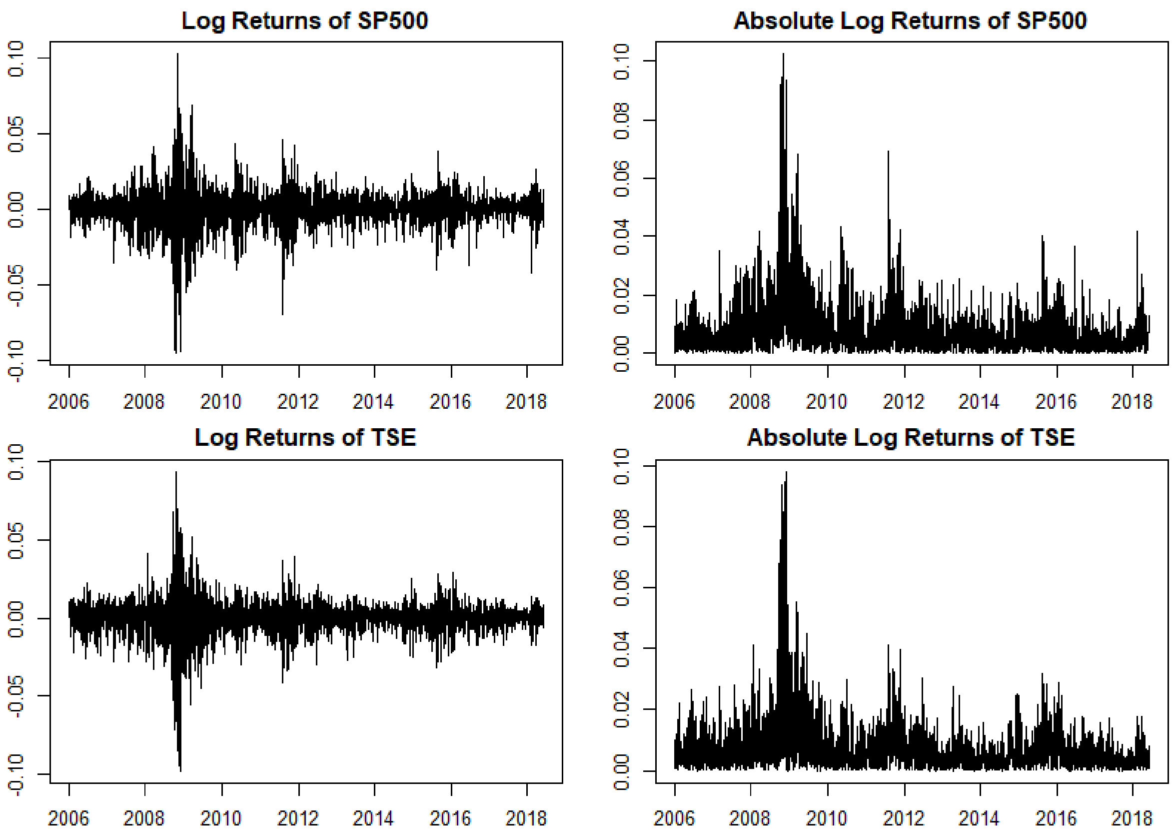

Figure 6: Daily returns and absolute returns of S&P 500 and TSE

Table 3
Marginal model summary

Model	Parameter	TSE	S&P 500
ARMA	p1	0	0
	q1	0	0
GARCH(1,1)	ω	$2.064 * 10^{-5}$	$4.953 * 10^{-6}$
	α	0.138^a	0.116^b
	β	0.834^a	0.881^a
	ξ	0.789^a	0.876^a
	ν	1.894^a	1.374^a
	BIC	-4.504	-4.503
	Log likelihood	576.807	576.659
	$QW(10)$	4.324	13.881
	$QW(20)$	14.41	25.889
	$QW^2(10)$	16.536^c	11.488
	$QW^2(20)$	21.962	15.497
	KS	0.042	0.052

ARMA model selects ARMA(0,0) according to the 'BIC' criterion. All Ljung-Box statistics are not significant at 5% level, so both series and their squared ones are not correlated.
KS is the Kolmogorov-Smirnov statistic, it does not reject that $SGED^{-1}(\varepsilon)$ follows uniform distribution from 0 to 1.

As an example, we select data from 29^{th} July, 2008 to 1^{st} Aug, 2009 to model their margins. The financial crisis took place in this period, so we can see clear volatility cluster and Ljung-Box test confirm the existence of autocorrelation and arch effect. ARMA-GARCH model works as a filter in order to have innovation processes which are serially independent. As table 3 shows, the Ljung-Box test applied to the residuals of the model does not reject the hypothesis of no autocorrelation and arch effect for both series at 5% level, and the estimated distribution of residuals is reasonable according to KS test, so the model is adequate. And model summaries using other data are similar which we will not show the detail.

5.3 Modelling the dependence

Figure 7 shows plots of return series and standardized innovations from 29^{th} July, 2008 to 1^{st} Aug, 2009. The standardized innovations are more discrete after eliminating autocorrelation and arch effect, while the positive dependence between both innovations still exists since points distribute around the diagonal. And the distribution of the innovation vector $(u_t, v_t) = (SGED_{1,t}(\varepsilon_{1,t}), SGED_{2,t}(\varepsilon_{2,t}))$ are assumed to have some specific dependence structure which will be modelled by different copulas.

We are going to use 4 different copulas—gaussian, student t, Clayton and Gumbel, to model the dependence. Parameters in each copula will be estimated by 2-step maximum likelihood method talked in section 2.2. Note that, we hope to use Gumbel copula's feature on top-right tail to model the dependence in bottom-left case, thus rotated Gumbel copula will also be included. It is defined as:

$$C_{Rotated-Gumbel}(u_1, \ldots, u_N; \alpha) = \sum_{i=1}^{N} u_i - (N-1) + C_{Gumbel}(1-u_1, \ldots, 1-u_N; \alpha)$$

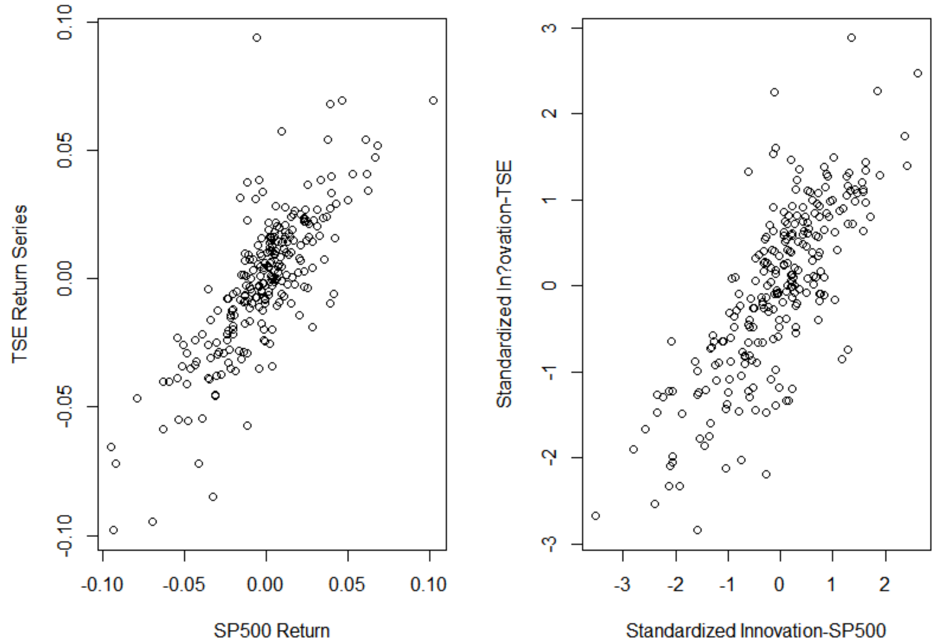

Figure 7: Scatter plot of return series and innovation pairs

where C_{Gumbel} is the Gumbel copula. In fact, any copula has reflect, and this kind of copula is known as survival copula. Given the same parameter, the density distribution of rotated copula is merely the reflect of its original copula, so it is a good alternative for us to model the bottom-left dependence using Gumbel copula.

Figure 8,9 shows the cumulative density distribution of copula on the diagonal. Generally, Student t and Gaussian copula fit relatively well regardless in the whole part or in the lower tail, because the empirical copula always distribute around them. However, in the lower tail, the Clayton copula and Gumbel copula are almost the upper bound and the lower bound of empirical copula since Clayton copula puts the most weight in lower tail while Gumbel copula puts the least. Rotated Gumbel is much similar to Clayton copula with just a little lower than Clayton copula but higher than empirical copula most of time in the lower tail.

In order to evaluate the goodness of fit directly, like Palaro and Hotta [2005], we use the quadratic distance between empirical copula and the estimated copula as a criterion.

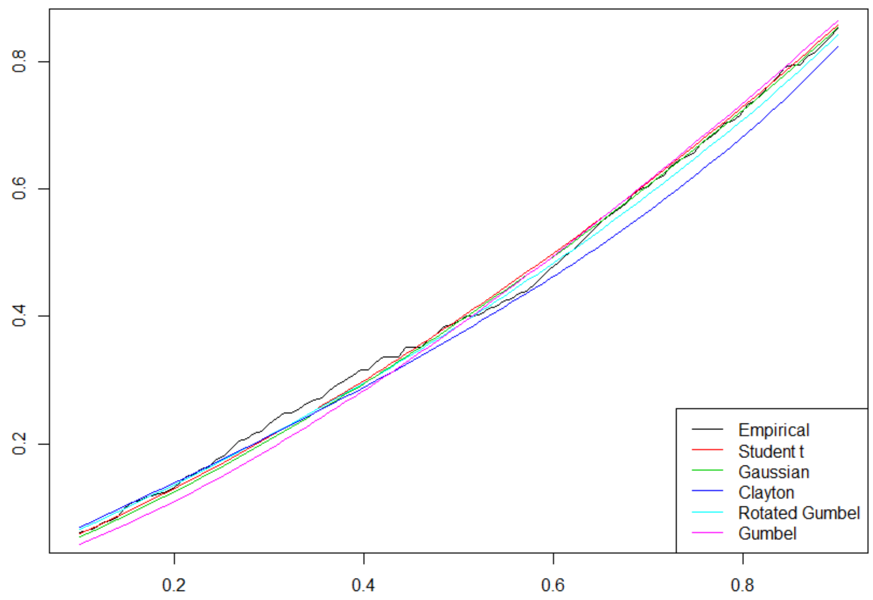

Figure 8: Cumulative distribution from 0 to 1

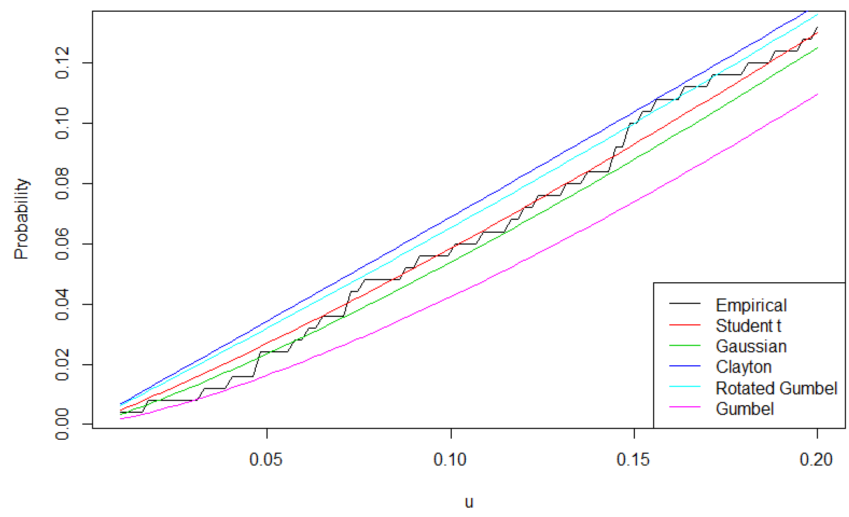

Figure 9: Cumulative distribution from 0 to 0.2

Definition 5.1 (Quadratic distance) *For any region $\boldsymbol{D} \subset I^N$, the distance function $d(\cdot, \cdot)$ between copula C_1 and C_2 is:*

$$d(C_1, C_2) = \sqrt{\sum_{i}^{m} (C_1(\boldsymbol{a_i}) - C_2(\boldsymbol{a_i}))^2}, \quad \boldsymbol{a_i} \in \Lambda \tag{5.2}$$

where $\Lambda = \{\boldsymbol{a_1}, \ldots, \boldsymbol{a_m}\}$ is a finite set of points with dimension N in $\boldsymbol{D}$.

Table 4 reports the mean of quadratic distance in evenly spaced grid in all periods. Similar to the figures, Student t copula has the least distance between empirical copula, then is the Gaussian copula. Rotated Gumbel fits better than Clayton copula and Gumbel copula in the lower tail but not in the global part. More importantly, the result indicates that the dependence of both series

Table 4
Quadratic distance between the estimated and empirical copula in different areas

$\boldsymbol{D}$	Gaussian	Student t	Clayton	Rotated Gumbel	Gumbel
$[0,1]^2$	0.721%	0.665%	1.544%	1.745%	1.028%
$[0,0.2]^2$	0.685%	0.633%	0.771%	0.774%	1.510%
$[0,0.1]^2$	0.480%	0.453%	0.630%	0.456%	0.993%
$[0,0.05]^2$	0.324%	0.328%	0.503%	0.338%	0.535%

is symmetric after eliminating the autocorrelation and arch effect since Student t copula is closer than Clayton copula and rotated Gumbel copula. This confirms the conclusion in Longin and Solnik [2001] that "high volatility per se(i.e. large absolute returns) doesn't seem to lead to an increase in conditional correlation. Correlation is mainly affected by the market trend", because correlations in both tails do not distinguish each other given no arch effect and autocorrelation, the larger losses in the bear market are more likely to be caused by larger volatility. What's more, although the dependence seems to be symmetric, it still has heavy tails indicating that extreme events are more likely to happen together than in the multivariate normal case because student t copula fits better than gaussian copula.

5.4 Estimating VaR

We take following steps to estimate 1-day ahead VaR:

1. We take 250 days as time window which is almost the trading days of 1 year.
2. We build $ARMA(p,q)-GARCH(1,1)$ model for each series in the interval and impose estimated SGED distribution to residuals to get $(u_t, v_t) = (SGED_{1,t}(\varepsilon_{1,t}), SGED_{2,t}(\varepsilon_{2,t}))$ which follows uniform distribution.
3. We fit different types of copula to these pairs to estimate parameters, maximum likelihood method in section 2.2 is used.
4. We generate 400 pairs (u'_t, v'_t) based on the estimated copula.
5. According to the marginal model, we impose inverse SGED distribution function to (u'_t, v'_t) to get $(SGED_{1,t}^{-1}(u'), SGED_{2,t}^{-1}(v'))$ and multiply them with predicted 1-day ahead volatility to get the simulated residuals.
6. Combining the ARMA model with residuals, we finally get 400 pairs of simulated returns through which we can calculate the portfolio return, and thus VaR at different quantiles.
7. We repeat step 2-6 1000 times and take the mean of them as 1-day ahead VaR of the last day in the interval.
8. Then we move the window ahead 5 trading days and calculate VaR till the end of the data.

Through above steps, we can get 1-day ahead VaR calculated by different models from 9^{th} Jan, 2007 to 31^{th} May, 2018 in every week, the number of points is 563. In order to asses the accuracy of VaR estimates, we backtested the method at 90%, 95%, 99% and 99.5% confidence level by the following procedure. At time t, we use the data from $t-250$ to t to build the model and to estimate 1-day ahead $VaR_{t,\alpha}$ at α level, then we test whether the return of portfolio at

Table 5
Number of observations where the loss exceed the estimated VaR

	Gaussian	Student t	Clayton	Rotated Gumbel	Gumbel	VaR hist	VaR delta	VaR sum
				$\alpha = 10\%(56)$				
number	58	58	58	58	59	59	57	54
P-value	80.111%	80.111%	80.111%	80.111%	69.59%	69.59%	91.062%	75.568%
				$\alpha = 5\%(28)$				
number	35	35	31	33	35	38	39	31
P-value	19.760%	19.760%	58.060%	35.547%	19.76%	$6.833\%^c$	$4.572\%^b$	58.060%
				$\alpha = 1\%(6)$				
number	10	10	8	9	12	9	16	8
P-value	$9.429\%^c$	$9.429\%^c$	34.279%	18.756%	$1.881\%^b$	18.756%	$0.033\%^a$	34.279%
				$\alpha = 0.5\%(3)$				
number	6	6	4	4	10	5	11	4
P-value	$9.780\%^c$	$9.780\%^c$	50.357%	50.357%	$0.086\%^a$	23.822%	$0.021\%^a$	50.357%

”VaR hist” is the historical simulation method calculated from portfolio return.
”VaR delta” is the variance-covariance method where we assume portfolio return follows normal distribution and we calculate standard deviation and mean from both series and their correlation.
”VaR sum” is the weighted sum of VaR calculated from historical simulation of each series.

time $t + 1$ R_{t+1} is below $VaR_{t,\alpha}$, then we forward to time $t + 5$ to re-estimate the model and repeat the whole process. Therefore, we have a total of 562 tests for VaR at each level.

Table 5 reports the number of observation exceeding the estimated VaR at α level. Besides, it includes a statistic defined as:

$$K = -2\ln[(1-p)^{n-m}p^m] + 2\ln[(1-m/n)^{n-m}(m/n)^m]$$

where p denotes the assumed probability of exception happening, m denotes the real number of exceptions, and n denotes the number of total tests. Kupiec [1995] shows that K follows $\chi(1)$ distribution, so K is a two-sided test which could help us to do statistical inference easily.

Basically, if we select 5% as reject level, then only Gumbel model and ”VaR delta” will be removed, the rest are adequate at all levels. If we raise the reject level to 10%, then Clayton model, rotated Gumbel model and ”VaR sum” are still adequate. If we compare the result horizontally, ”VaR delta” is the best at 10% level, Clayton model and ”VaR sum” are the best at 5% and 1% levels, Clayton model, rotated Gumbel model and ”VaR sum” are the best at 1% level. But note that, VaR is not a coherent measure which means $VaR(R_p) < \sum_{i=1}^{N} VaR(R_i)$, though ”VaR sum” is among the best in our simple bivariate case, it suffers when the dimension increase. Therefore, in the back-test, the Clayton model and rotated Gumbel model perform the best, while we still can not reject gaussian model and student t model are good ones using frequently used 5% level. That is to say, ARMA-GARCH-copula performs good theoretically and empirically.

6 Conclusion

This paper introduces the use of copula in modelling the dependence between series and in estimating portfolio value-at-risk. We first provide copula's power and flexibility in modelling multivariate distribution, then we introduce some frequently used copula classes and their superiority over other correlation or concordance measures. Finally, we resort to using copula to model the dependence between S&P/TSX composite index and S&P 500 and estimating the VaR of portfolio composing of both series. In particular, we build ARMA-GARCH model with flexible ARMA order for each series to eliminate autocorrelation and arch effect, then we fit different copulas to innovation pairs and use a criterion to determine the goodness of fit of these copulas, finally, we get estimated 1-day ahead VaR according to the ARMA-GARCH-copula model built in each interval, and we test the accuracy of VaR using out-of-sample test and compare them with traditional approaches for estimating VaR such as variance-covariance and historical simulation.

Basically, we use 2 methods to do evaluation. According to the quadratic distance between estimated copula and empirical copula, student t copula fits the best for the dependence. It is interesting because the asymmetric dependence where losses in the bear market are larger than gains in the bull market disappear after eliminating autocorrelation and arch effect, at least for a relatively short period(i.e. 250 trading days). Furthermore, though the dependence is not asymmetric, that the tail of student t copula is heavier than that of gaussian copula indicates that extreme events are more likely to happen together in the real world. In the VaR test, in addition to Gumbel copula, VaR calculated by ARMA-GARCH-copula is generally better than traditional approaches. Although student t copula and gaussian copula fit good with empirical copula, Clayton copula and rotated Gumbel copula which have heavier left tail provide more accurate VaR limits at all levels. Note that, the weighted sum VaR calculated by each component in the portfolio, though performs good in our result too, it suffers from many disadvantages including dimension ,coherence and exaggeration. By using copula to model dependence, we can get a more robust VaR which can help us allocate assets more reasonable.

References

Andrew Ang and Joseph Chen. Asymmetric correlations of equity portfolios. *Journal of financial Economics*, 63(3):443–494, 2002.

Eric Bouyé, Valdo Durrleman, Ashkan Nikeghbali, Gaël Riboulet, and Thierry Roncalli. Copulas for finance - a reading guide and some applications. *Social Science Electronic Publishing*, 2007.

Umberto Cherubini and Elisa Luciano. Value-at-risk trade-off and capital allocation with copulas. *Economic Notes*, 30(2):235–256, 2001.

P Deheuvels. La fonction de dépendance empirique et ses propriétés. académie royale de belgique. *Bulletin de la Classe des Sciences*, 65(5):274–292, 1979.

P Embrechts. Correlation and dependency in risk management : properties and pitfalls. *Risk Management Value at Risk & Beyond*, pages 176–223, 2002.

Maurice Fréchet. Sur les tableaux de corrélation dont les marges sont données. *Ann. Univ. Lyon, 3^ e serie, Sciences, Sect. A*, 14:53–77, 1951.

Maurice Fréchet. Les tableaux de corrélation et les programmes linéaires. *Revue de l'Institut international de statistique*, pages 23–40, 1957.

Edward W Frees and Emiliano A Valdez. Understanding relationships using copulas. *North American Actuarial Journal*, 2(1):1–25, 1998.

Christian Genest and Jock Mackay. The joy of copulas: Bivariate distributions with uniform marginals. *American Statistician*, 40(4):280–283, 1986.

Christian Genest and Louis Paul Rivest. Statistical inference procedures for bivariate archimedean copulas. *Publications of the American Statistical Association*, 88(423):1034–1043, 1993.

Jen Jsung Huang, Kuo Jung Lee, Hueimei Liang, and Wei Fu Lin. Estimating value at risk of portfolio by conditional copula-garch method. *Insurance Mathematics & Economics*, 45(3): 315–324, 2009.

Harry Joe. *Multivariate models and multivariate dependence concepts*. CRC Press, 1997.

Eric Jondeau and Michael Rockinger. The copula-garch model of conditional dependencies: An international stock market application. *Journal of International Money Finance*, 25(5): 827–853, 2006.

William H. Kruskal. Ordinal measures of association. *Publications of the American Statistical Association*, 53(284):814–861, 1958.

Paul Kupiec. Techniques for verifying the accuracy of risk measurement models. *Social Science Electronic Publishing*, 3(2):73–84, 1995.

Filip Lindskog and ETH RiskLab. *Modelling dependence with copulas and applications to risk management*. PhD thesis, Master Thesis, ETH Zürich, 2000.

Francois Longin and Bruno Solnik. Extreme correlation of international equity markets. *The journal of finance*, 56(2):649–676, 2001.

Roger B Nelsen. *An Introduction to Copulas*. Springer,, 2006.

Helder P Palaro and Luiz Koodi Hotta. Using conditional copula to estimate value at risk. *Journal of Data Science*, 4(1):93–115, 2005.

Andrew J. Patton. Estimation of multivariate models for time series of possibly different lengths. *Journal of Applied Econometrics*, 21(2):147–173, 2006.

Berthold Schweizer and Edward F Wolff. On nonparametric measures of dependence for random variables. *The annals of statistics*, pages 879–885, 1981.

M Sklar. Fonctions de repartition an dimensions et leurs marges. *Publ. inst. statist. univ. Paris*, 8:229–231, 1959.