Journal of Business Theory and Practice
ISSN 2372-9759 (Print) ISSN 2329-2644 (Online)
Vol. 6, No. 4, 2018
www.scholink.org/ojs/index.php/jbtp

Original Paper

Copula Model Selection of Stock Return Time Series Using Information Complexity

Jana Salim[1*]

[1] School of Economics, Shandong University, Jinan, China

[*] Jana Salim, School of Economics, Shandong University, Jinan, China

Received: September 20, 2018 Accepted: September 29, 2018 Online Published: October 15, 2018
doi:10.22158/jbtp.v6n4p294 URL: http://dx.doi.org/10.22158/jbtp.v6n4p294

Abstract

In this paper we estimate the correlation between four different stock return prices. To accomplish this, we use the copula models to study the dependency structure between the variables. The original variables of interest are mapped into more manageable variables by considering joint and marginal distributions of these variables. Then a correlational structure between these variables are obtained. We fit several well-known copula models to the portfolio of the stock return price dataset using consistent information complexity (CICOMP) criterion along with other AIC-type criteria to choose the best copula functional model. CICOMP predominated the AIC-type criterion, both in the case when the fitted models are correctly specified. We expect to get more realistic results using other copula distributions contrary to the Gaussian copula used by Li (2000) that fails to capture the dependence between extreme events.

Keywords

bicopula modeling, choosing the best copula model, information complexity, kendall correlation

1. Introduction

Market and Credit risk have been traditionally the most severe risk that many financial institutions are exposed to and for which the most regulatory capital is required. Market risk is caused by several periods of distress where stock markets can experience. Credit portfolio risk is manifested by the default of the structured securities and that was obvious during and after the 2008 housing crisis. The dependence structure is also important since the implications of dependence are applied in the models for pricing those derivatives and calculating the quantitative risk measures like the Value at Risk (VaR). Investigating and forecasting stock prices require too much attention on the statistical weight like studying the interactions and dependencies between those multivariate variables.

Hence the main question is centered about the possibility to capture the dependence between the return stock prices without any assumptions made on their marginal distributions or the joint distribution. But instead, by finding the best fit joint distribution for the multivariate random variables using information complexity criterion ICOMP. There are many options to solve this problem, such as Akaike's Information Criterion (AIC), Schwartz Bayesian Criterion (SBC), Rissanen's Minimum Discription Length (MDL), for some examples. Based on the results, for the first time, in this paper we introduce Bozdogan's (2010, 2016) Consistent Information Complexity (CICOMP) criterion to fit the best bivariate vine copulas for capturing the correlation between the stock return prices.

As is well known, the co-variance matrix only captures the linear dependence in the data for special distributions such as normal (or Gaussian) distribution. Hoeffding (1940, 1941) studied non-parametric measures of association such as Spearman's rho in multivariate distribution.

So copula is the appropriate modeling technique needed to estimate the joint distribution hence the dependence between stock return prices. The word copula was mentioned for the first time in Sklar's (1959) work in his famous Sklar's theorem.

A copula decouples the risk associated with the portfolio dependence structure from the individual risks of each obligor. There are many copula functions to fit that measures the portfolio dependence such as the normal copula, which assumes that the latent variables follow a multivariate normal distribution. Normal copula (Figure 1) has been incorporated by Li (2000), where he used the one factor Gaussian model and this what actually lead to the crisis in 2008 and to the credit derivative obligations to meltdown.

The copula model has featured attractive models for measuring different aspects of dependence in finance. For example the dependence between the probability of default of the securities CDOs using copula (Frey et al., 2001). Embrechts et al. (2002) introduced the financial applications of Copula models in risk management (Embrechts et al., 2002), Frey and McNeil (2003). Pricing of derivatives has been studied by Cherubini et al. (2004), Rosenberg and Schuermann (2006). Estimating the dependence between stock markets (Jondeau & Rockinger, 2006); between exchange rates (Patton, 2006a; Bartram et al., 2007; Necula, 2010). And finally Contagion dependence among financial markets (Durante & Jarowski, 2010; Boero et al., 2011).

In estimating the dependence structure between four stock indices PX, SP500, BUX and DAX, Necula (2010) has found that the t-copula and the Gumbel-Clayton mixture copulas are the best fit copula functions to capture the correlation of two financial return series.

Several attempts had been made to develop parsimonious model, for example Bozdogan's (1987) work resolves issues related to the second term of AIC such as the consistency similar to SBC. Bozdogan's (1987) extended AIC further and obtained another dimension consistent criterion ICOMP .

Vine Copulas have been recently proposed as one of the most powerful alternative tool to the multivariate copulas (Joe, 1996).

GLM-based models emphasizes too how the dependence in each pair of conditioned variables relies on the conditioning variables. Others proposed functions LASSO, Tibshirani (1996) and SCAD by Fan and Li (2001), Bedford and Cooke (2001, 2002). Spiegelhalter et al. (2002) used Deviance Information Criterion (DIC) and other criteria to select the best copula fit model.

Czado and Min (2011) studied vine copulas in a Bayesian framework. Ingrid Hobæk Haff (2013) studied the asymptotic characteristics of the sequential estimators for vine copulas.

Vine copulas captures the asymmetry as well as the tail dependency of the underlying portfolio through decomposing the multivariate copula densities into bivariate ones (Schepsmeier et al., 2013). Many papers have used also the Akaike Information Criterion (AIC) and BIC for model selection (Gronnberg, Steffen, & Hjort, 2014).

Therefore, for ICOMP, in addition to the lack of fit, the lack of parsimony and the profusion of complexity are data-adaptively adjusted by the entropic complexity of the estimated IFIM across the competing alternative models as the parameter spaces of these models are constrained in the model fitting process data-adaptively.

In this paper we attempt to fit and select the best unbiased vine copula using the Bozdogan's (2010, 2016) consistent Information Complexity (CICOMP) criterion. We then estimate the correlation matrix of the given return stock prices using the Kendall's Tau based on the fitted chosen copula model.

The rest of the paper is organized as follows, section 2 reviews the literature review of the concept of copula and ICOMP selection criteria, section 3 analyses the data used, section 4 refers to the methodology, section 5 discusses the empirical results, section 6 concludes.

2. Method

A copula, following Mcneil et al. (2005b, pp. 184-228) can be defined in d dimensions as follows:

A d-dimensional copula $C:[0,1]^d \to [0,1]$ is a cumulative function with uniform marginal distributions.

The notation $C(u) = C(u_1, u_2, \ldots, u_d)$ is subsequently reserved for the multivariate distribution functions, which represent copulas. Following McNeil et al. (2005b, p. 185) three properties characterize a copula such that every function satisfying them is a copula (Sklar, 1959):

1) $C(u) = C(u_1, u_2, \ldots, u_d)$ is increasing in each component u_i.

2) By setting $u_i = 1$ for all $j \neq i$ the marginal component i is attained and since it must be uniformly distributed, $C(1, \ldots 1, u_i, 1 \ldots 1) = u_i$.

3) For $a_i \leq b_i$ the probability $P(U_1 \in [a_1, b_1], \ldots, U_d \in [a_d, b_d])$ has to be nonnegative.

4) $C(U_1, U_2, \ldots, U_d) \leq P(U_1 \leq u_1, \ldots\ldots, U_d \leq u_d)$.

Consider a d-dimensional distribution function with marginal distributions $F_1, \ldots, F_d$. Then there exists a copula $C:[0,1]^d \to [0,1]$, such that :

$$F(x_1, .., x_d) = C\big(F_1(x_1), \ldots, F_d(x_d)\big) for all x_i, \ldots, x_d in [-\infty, +\infty] \quad (1)$$

Examining the implications of equation (1) for the copula itself and making use of the property $F_i(F_i^{-1}(y)) \geq y$ one obtains:

$$C(u) = F(F_1^{-1}(u_1), \ldots.., F_d^{-1}(u_d)) \quad (2)$$

The relation described in equation (1) typically represents the starting point for simulations that are based on a given copula and given marginal while equation (2) is more a theoretical instrument to get the copula from a multivariate distribution function.

2.1 Elliptical Copulas

Elliptical copulas are easily obtained. The Gaussian copulas are elliptical.

• **Gaussian copula**

The Gaussian copula is defined by:

$$C_p^{Ga}(u, v) = \Phi_2(\Phi^{-1}(u), \Phi^{-1}(v)) \quad (3)$$

Where $\Phi_2(.,.;p)$ is the joint distribution function of two standard normal distributed random variables with a correlation coefficient $\rho \in (-1, -1)$, Φ is the $N(0,1)$ cumulative distribution function (cdf) and Φ^{-1} (the quantile function) is its functional inverse.

The density of the bivariate Gaussian copula is given by

$$c(u_1, u_2, \rho) = \frac{1}{\sqrt{1-\rho^2}} \exp\{\frac{x_1^2 + x_2^2 - 2\rho x_1 x_2 + (1-\rho^2)(x_1^2 + x_2 2)}{2(1-\rho^2)}\} \quad (4)$$

Where $X_1 = \Phi^{-1}(u)$and $X_2 = \Phi^{-1}(v)$

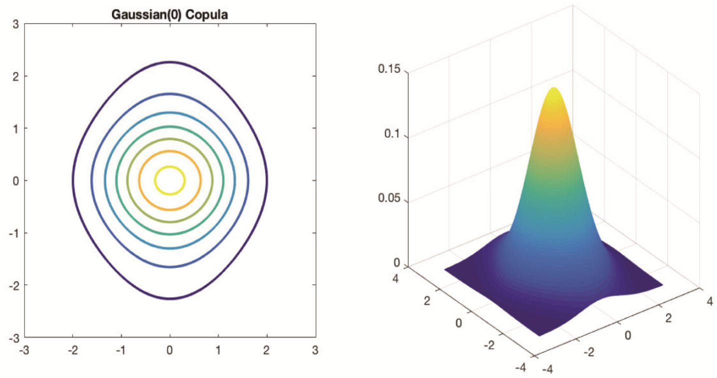

Figure 1. Contour and 2D Plots of Gaussian Copula with Correlation ρ=0.5

• **Multivariate Gaussian copula**

The multivariate Gaussian copula function is applied to a joint distribution with correlation matrix R, it is defined by:

$$C_R(u_1,\ldots,u_d) = \Phi_R(\Phi^{-1}(u_1),\ldots,\Phi^{-1}(u_d)), \quad (5)$$

Where Φ_R is the distribution function of the joint random variables. The variables are normal and standardized with a correlation matrix R.

• **Student's t copula**

The student's t copula is of the form:

$$C_T(u_1,\ldots,u_n) = T_{\upsilon,R}(T_\upsilon^{-1}(u_1),\ldots,T_\upsilon^{-1}(u_n)) \quad (6)$$

Where $T_{\upsilon,R}$ is the multivariate student t-distribution function with υ is the degree of freedom. Mean vector is 0 and correlation matrix R. The student's t factor model can be interpreted as a student's t copula. The student's t copula has tail dependence in both tails.

2.2 Kendall's Tau and Spearman's Rho

Instead of concentrating on the data itself, it is a popular approach in non-parametric statistics to focus on the ranks of data. This concept has given rise to Kendall's tau and Spearman's rho, which are the two important estimators of correlation. By focusing on ranks one obtains a scale-invariant correlation estimate which is advantageous when working with copulas. Rank correlations will offer a potential way to fit copulas to data.

Although the rank correlations are better suited to the analysis of a joint distribution of a financial data than linear correlations but there is a relation between the linear correlation coefficient ρ, Kendall's tau τ and Spearman's rho ρ_s:

Kendall's tau for the Gaussian copula is given by:

$$\tau_K = \frac{2}{\pi}\arcsin(\rho), \quad (7)$$

and Spearman's rho is given by the following equation:

$$\rho_s = \frac{6}{\pi}\arcsin(\frac{\rho}{2}) \quad (8)$$

Gaussian copula does not capture the dependence in the tails of the distribution.

The Kendall's tau for the student's t copula is:

$$\frac{2}{\pi}\arcsin(\rho) \quad (9)$$

2.3 Archimedean Copulas

A $2D$ copula is Archimedean if it is expressed as follows:

$$C_\varphi(u,v) = \varphi^{-1}(\varphi(u) + \varphi(v)), 0 < u, v < 1 \quad (10)$$

Where $\varphi^{-1}: [0,1] \times [0,1] \to [0,\infty]$ is the inverse generator with $\varphi^{-1}(0) = \inf\{t: \varphi = 0\}$

C_φ is a copula if and only if φ is convex.

For a bivariate random variable there is one to one correspondence between the copula and Kendall's tau τ, which is given by

$$\tau = 1 + 4\int_0^1 \frac{\varphi(t)}{\varphi\prime(t)} dt, \quad (11)$$

Where φ is a continuous, strictly decreasing function from $[0,1]$ to $[0,\infty]$ such that $\varphi = 1$. This is called the generating function.

There are several other Archimedean copulas for measuring the dependence. Here we present only three of them as follows.

• **Clayton copula**

For $\theta > 0$, the Clayton copula is given by

$$C(u,v) = (u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}, where \theta \in (0,\infty) \quad (12)$$

The Clayton copula provides the dependence in the lower parts of the tails for $\theta > 0$ of a joint distribution.

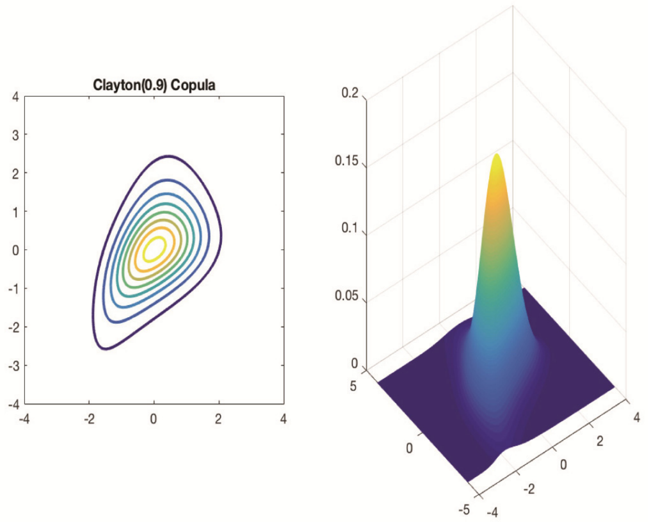

Figure 2. Contour and 2D Plots of Clayton Copula with Parameter *θ*=0.9

For Clayton copula, Kendall's tau is given by

$$\tau = \frac{\theta}{\theta+2} \quad (13)$$

• **Gumbel**

The Gumbel copula (1960) is used to model asymmetric dependence in the data. The bivariate Gumbel copula is given by

$$C^{Gu}(u,v) = \exp\{-[(-\log u)^\theta + (-\log v)^\theta]^{\frac{1}{\theta}}\}, \quad (14)$$

Where θ is the copula parameter restricted on the interval $[1, \infty)$.

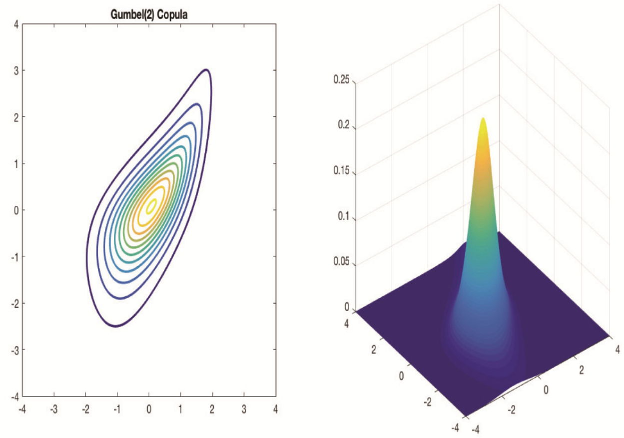

Figure 3. Contour and 2D Plots of Gumbel Copula with Parameter *θ*=2

Kendall's tau for a bivariate Gumbel copula is

$$\tau = 1 - \frac{1}{\theta} \tag{15}$$

• **Frank copula**

The Frank copula (1979) is given by

$$C^{Fr}(u, v) = -\theta^{-1}\log\{1 + \frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\}, \tag{16}$$

Where θ is the copula parameter that may take any value.

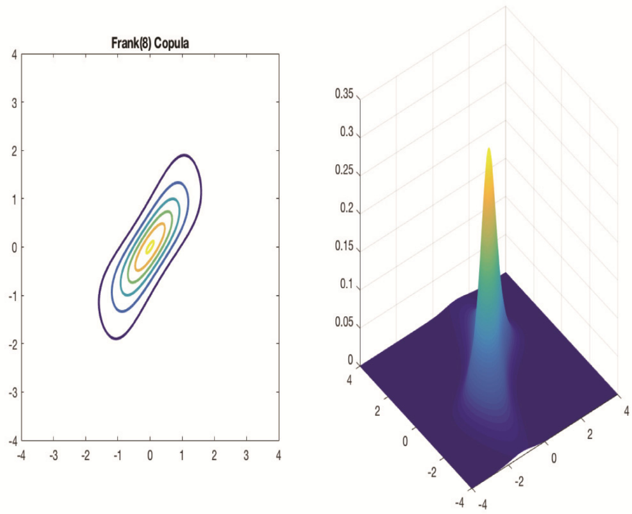

Figure 4. Contour and 2D Plots of Frank Copula with Parameter *θ*=8

Kendall's tau for a bivariate Frank copula is given by

$$\tau = 1 - \frac{4}{\theta}(1 - D_1(\theta)), \tag{17}$$

Where $D_k(x)$ is the Debya function, given by

$$D_k(x) = \frac{k}{x^k}\int_0^x \frac{t^x}{e^t-1}dx \quad (18)$$

2.4 Choosing the Optimal Bivariate Copula Model

In this paper we use the R-package Vine Copula by Schepsmeier et al. (2013) to analyze our data. The main functions for exploratory data analysis, selection and estimation of bivariate copulas are as follows.

R-vine Structure Select where R-vine trees are selected using maximum spanning trees w.r.t. some edge weights. The most commonly used edge weigth is the absolute value of the empirical Kendall's tau, say $\tilde{\tau}_{ij}$. Then, the following optimization problem is solved for each tree:

$$max \sum_{edgese_{ij} \in inspanningtree} \tilde{\tau}_{ij} \quad (19)$$

where a spanning tree is a tree on all nodes. The setting of the first tree selection step is always a complete graph. For subsequent trees, the setting depends on the R-vine construction principles, in particular on the proximity condition. Some commonly used edge weights are implemented: "tau" absolute value of empirical Kendall's tau, "rho" absolute value of empirical Spearman's rho, "AIC" Akaike information (multiplied by -1), "BIC" Bayesian information criterion (multiplied by -1) and "CAIC" corrected Akaike information criterion (multiplied by -1).

If the data contain NAs, the edge weights in "tau" and "rho" are multiplied by the square root of the proportion of complete observations. This penalizes pairs where less observations are used.

The criteria "AIC", "BIC", and "CAIC" require estimation and model selection for all possible pairs. This is computationally expensive and much slower than "tau" or "rho". The user can also specify a custom function to calculate the edge weights. The function has to be of type function (u_1, u_2, weights) and must return a numeric value. The weigths argument must exist, but does not has to be used.

For example, "tau" (without using weights) can be implemented as follows:

Function (u_1, u_2, weights)=absolute (correlation (u_1, u_2, method="kendall", use="complete.obs")).

The root nodes of C-vine trees are determined similarly by identifying the node with strongest dependencies to all other nodes. That is we take the node with maximum column sum in the empirical Kendall's tau matrix. Note that a possible way to determine the order of the nodes in the D-vine is to identify a shortest Hamiltonian path in terms of weights $1 - |\tilde{\tau}_{ij}|$. This can be established for example using the package TSP in R program.

For model selection, we can use Akaike's (1973) Information Criterion (AIC), Schwarz (1978) Bayesian Criterion (SBC) known also as BIC, and Bozdogan's (1987) consistent AIC (CAIC) to choose the best fitting copula. All the available copulas are fitted using the Maximum Likelihood (ML) method.

If u_1 and u_2 (u_1=u,u_2=v) are negatively dependent, Clayton, Gumbel, Joe, BB1, BB6, BB7 and BB8 and their survival copulas are not considered. The family with the minimum of the criterion is chosen as the best fitting model.

2.5 Model Selection Criteria, Information Complexity

For a given n observations, AIC for a bivariate copula family density c (u_1, u_2) with parameter (s) θ is defined as:

$$AIC = -2\sum_{i=1}^{n} \ln[c(u_{i1}, u_{i2}|\theta)] + 2k, \quad (20)$$

Where k=1 for one parameter copulas and k=2 for the two parameter t-, BB1, BB6, BB7, and BB8 copulas. Similarly the SBC (or BIC) is given by

$$BIC = -2\sum_{i=1}^{n} \ln[c(u_{i1}, u_{i2}|\theta)] + \ln(n)k, \quad (21)$$

Bozdogan's (1987) CAIC is given by

$$CAIC = -2\sum_{i=1}^{n} \ln[c(u_{i1}, u_{i2}] + k[\ln(n) + 1], \tag{22}$$

We note that the penalty for two parameter families is much heavy when we use SBC and CAIC as compared to AIC.

Based on these results, for the first time, in this paper we introduce Bozdogan's (2010, 2016) consistent information complexity (CICOMP) criterion given by

$$CICOMP = -2\sum_{i=1}^{n} \ln[c(u_{i1}, u_{i2}|\theta)] + k[\ln(n) + 1] + 2C_{1F}\left(\widehat{Cov}(\hat{\theta})\right)$$

$$= CAIC + 2C_{1F}\left(\widehat{Cov}(\hat{\theta})\right), \tag{23}$$

Where

$$\widehat{Cov}(\hat{\theta}) = \widehat{H}^{-1}(\hat{\theta})\hat{C}(\hat{\theta})\widehat{H}^{-1}(\hat{\theta}) \tag{24}$$

is the robust estimated covariance matrix of the parameters, and where

$$C_{1F}\left(\widehat{Cov}(\hat{\theta})\right) = \frac{1}{4\bar{\lambda}_a^2}\sum_{j=1}^{s} (\lambda_j - \bar{\lambda}_a)^2 \tag{25}$$

is the quadratic complexity of the estimated covariance matrix.

We note that $C_{1F}(.)$ is a scale-invariant measure of complexity and $C_{1F}(.) \geq 0$ with $C_{1F}(.) = 0$ when all eigenvalues $\lambda_j = \bar{\lambda}_a$, the arithmetic mean of the eigenvalues of $\widehat{Cov}(\hat{\theta})$. Also, $C_{1F}(.)$ measures the relative variation in eigenvalues. In short, we have

$$CICOMP = -2\sum_{i=1}^{n} \ln[c(u_{i1}, u_{i2}|\theta)] + k[\ln(n) + 1] + 2[\frac{1}{4\bar{\lambda}_a^2}\sum_{j=1}^{s} (\lambda_j - \bar{\lambda}_a)^2]. \tag{26}$$

A model with minimum CICOMP is chosen to be the best fitting model among all possible alternative models.

3. Result

3.1 A Real Numerical Example and Computational Results

For our numerical example, we consider a real data set using only the first d=4 variables and the first n=250 observations of a data set that contains transformed standardized residuals of daily log returns of the original d=15 major German stocks represented in the index DAX observed from January 2005 to August 2009 for our illustration purposes in this paper due to space considerations. We fitted the best copula for the first four stocks, called Allianz SE (ALV.DE), BASF SE (BAS.DE), Bayer AG (BAYN.DE), Bayerische Motoren Werke AG (BMW.DE).

The time series and the scatter plot of the stock returns data is shown in Figure 5.

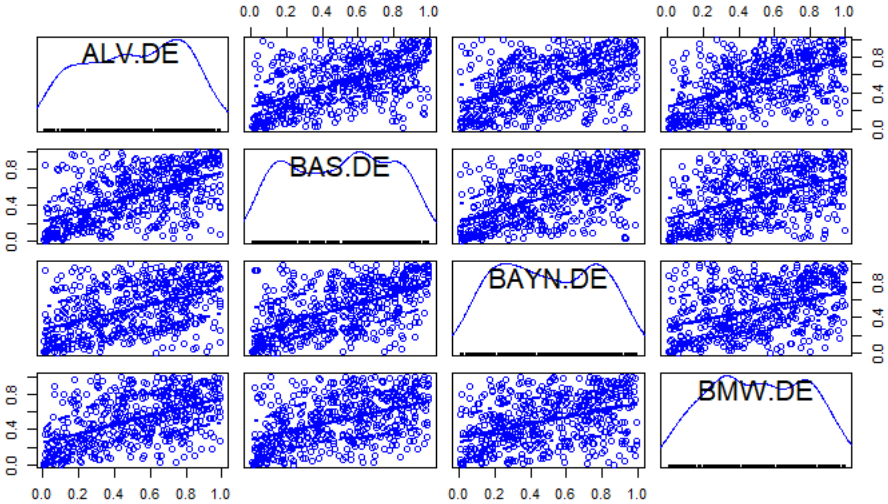

Figure 5. Time Series and Scatter Plot of Stock Returns Data

We note that the stock returns data set is not normally (i.e., Gaussian) distributed. Therefore, using copulas and copula modeling is appropriate to relax the classical distributional assumption to take into account the dependency and the tail behavior of the stock returns data set.

Since we considered the first d=4 variables of d=15 dimensional stock return data set, that is, we considered ALV.DE, BAS.DE, BAYN.DE, and BMW.DE, we have a combinations of 6 pair of copula vector of returns. For d=15 stock returns, we would have in total 105 combinations to consider, which we will not pursue here for space considerations. As it can be seen, for large dimensional data, we would have combinatorial explosion to construct bivariate copula models. Because of this, for this example we only chose the first d=4 stock returns. In a separate paper, for high dimensional data, we will generalize these results using the clever Genetic Algorithm (GA) to choose the best fitting copula model with information complexity (ICOMP) criterion.

3.2 Scores of Model Selection Criteria

Our results from the analysis of stock returns data set are presented in Tables 1 to 6 below for different copula distributions for each 6 pairs of copula vectors along with their values of the model selection criteria.

Looking at the results presented in Table 1, we see that Frank copula is the best fitting copula for vectors u_1 and u_2 (i.e., ALV.DE and BAS.DE) with correlation, $\tau = 0.41$. The contour and 2D plots of the Frank copula model is shown in Figure 4. As it can be seen from the plot of the Frank copula has heavy tails.

Further, looking at Table 2, we see that the rotated survival Clayton copula is the best fitting copula model for vectors u_1 and u_3 (i.e., ALV.DE and BAYN.DE) with correlation, $\tau = 0.15$. The contour and 2D plots of the Clayton copula is shown in Figure 2. Continuing with our pairwise analysis, from Table 3, we see that Frank copula is the best fitting copula model for vectors u_2 and u_3 with correlation, $\tau = 0.41$. From Table 4 results, we see that the rotated survival Gumbel copula is the best fitting copula model for vectors u_1 and u_4 with tau correlation, $\tau = 0.38$. Table 5 results show that the rotated survival Gumbel copula is also the best fitting copula model for vectors u_2 and u_4 with tau correlation, $\tau = 0.15$. Finally, from Table 6 results, we see that Gumbel copula is the best fitting copula model for vectors u_3 and u_4 with tau correlation, $\tau = 0.05$. The contour and 2D plots of the Gumbel copula is shown in Figure 3. This copula has more probability concentrated in the tails than does Frank copula. It is also asymmetric, with more weight in the right tail.

Table 1. Scores of Model Selection Criteria for the Stock Returns $\{u_1, u_2\}$

Copula Distributions	AIC	BIC	CAIC	CICOMP
Gaussian Copula	-2.845998	1.36861	68.44165	68.44165
t-copula	-2.808817	5.620399	211.766480	211.766480
Clayton Copula	0.2379965	4.452605	71.52565	71.52565
Rotated "Survival Gumbel" Copula	-1.08705	3.127558	70.20060	70.20060
Frank Copula	-6.438022	-2.223414	64.84963	**64.84963**
Rotated "Survival BB8" copula	-4.281868	4.147348	210.29343	210.29343

Table 2. Scores of Model Selection Criteria for the Stock Returns $\{u_1, u_3\}$

Copula Distributions	AIC	BIC	CAIC	CICOMP
Gaussian Copula	-22.536358	-18.32175	48.75129	48.75129
t-copula	-17.797262	-9.368046	196.778035	196.778035
Rotated "Survival Clayton Copula	-27.6194068	-23.404799	43.66824	**43.66824**
Gumbel" Copula	-21.30115	-17.086545	49.98650	49.98650
Frank Copula	-26.541491	-22.326883	44.74616	44.74616
Rotated "Survival BB8" copula	-25.985152	-17.555936	188.59014	188.59014

Table 3. Scores of Model Selection Criteria for the Stock Returns $\{u_2, u_3\}$

Copula Distributions	AIC	BIC	CAIC	CICOMP
Gaussian Copula	-33.44094	-29.22633	37.84671	37.84671
t-copula	-28.30094	-19.87172	186.274357	186.274357
Clayton Copula	-24.67039	-20.45578	46.61726	46.61726
Rotated "Survival Gumbel" Copula	-24.96566	-20.75106	46.32199	46.32199
Frank Copula	-39.84042	-35.62581	31.44723	**31.44723**
Rotated "Survival BB8" copula	-30.14284	-21.71362	184.43246	184.43246

Table 4. Scores of Model Selection Criteria for the Stock Returns $\{u_1,u_4\}$

Copula Distributions	AIC	BIC	CAIC	CICOMP
Gaussian Copula	-186.049300	-181.83469	-114.76165	-114.76165
t-copula	-212.126015	-203.696799	2.449282	2.449282
Clayton Copula	-171.862643	-167.64804	-100.575	-100.575
Rotated "Survival Gumbel" Copula	-200.67034	-196.455733	-129.38269	**-129.38269**
Frank Copula	-182.243437	-176.08974	-110.95579	-110.95579
Rotated "Survival BB8" copula	-182.403041	-173.973825	32.17226	32.17226

Table 5. Scores of Model Selection Criteria for the Stock Returns $\{u_2,u_4\}$

Copula Distributions	AIC	BIC	CAIC	CICOMP
Gaussian Copula	-177.81815	-173.60355	-106.53051	-106.53051
t-copula	-205.42676	-196.99754	9.148537	9.148537
Rotated "Survival Clayton" Copula	-198.28020	-194.06559	-126.99255	-126.99255
Rotated "Survival Gumbel" Copula	-211.74919	-207.53458	-140.46154	**-140.46154**
Frank Copula	-180.30434	-176.08974	-109.0167	-109.0167
Rotated "Survival BB8" copula	-199.87818	-191.44896	14.69712	14.69712

Table 6. Scores of Model Selection Criteria for the Stock Returns $\{u_3,u_4\}$

Copula Distributions	AIC	BIC	CAIC	CICOMP
Gaussian Copula	-158.9632	-154.7486	-87.67559	-87.67559
t-copula	-163.1173	-154.6881	51.45795	51.45795
Rotated "Survival Clayton Copula	-184.6783	-180.4637	-113.39065	-113.39065
Gumbel" Copula	-187.7309	-183.5163	-116.4433	**-116.4433**
Frank Copula	-154.846	-150.6314	-83.55832	-83.55832
Rotated "Survival BB8" copula	-184.469	-176.0398	30.10626	30.10626

Table 7. Correlation Matrix

	ALV.DE	BAS.DE	BAYN.DE	BMW.DE
ALV.DE	1	0.41	0.15	0.38
BAS.DE	0.41	1	0.41	0.15
BAYN.DE	0.15	0.41	1	0.05

Table 8. Summary of the Results of CICOMP for the Best Fitting Pairs of Stocks

Copula distributions	Pairs	CICOMP	τ
Frank copula	$\{u_1,u_2\}$	64.84963	0.41
"Survival Clayton" copula	$\{u_1,u_3\}$	43.66824	0.15
Frank copula	$\{u_2,u_3\}$	31.44723	0.41
"Survival Gumbel" copula	$\{u_2,u_4\}$	-140.46154	0.15
Gumbel copula	$\{u_3,u_4\}$	-116.4433	0.05

Table 7 presents the tau correlation matrix among the four stock returns, whereas Table 8 summarizes our final results. Looking at Table 8, if we have to choose one best pair of stock returns from German stock prices, based on the minimum CICOMP criterion we would choose the survival Gumbel copula model with pairs of stocks $\{u_2,u_4\}$ as our best fitting copula model, followed by Gumbel copula model. In other words, we choose the pairs of stocks {BAS.DE, BMW.DE} and {BAYN.DE, BMW.DE} stock returns. This makes sense in that it shows the strength of the popularity of the BMW.DE stock return.

4. Discussion

Examining and mitigating market risk is really a challenge in financial institutions. But as we mentioned before, the most important parameter to measure the market risk is the correlation between the return prices of the underlying stocks. We took the portfolio of the return prices of the stocks and fitted the copula distributions to identify the best copula model. The results show that each bivariate copula of every two return stock prices vectors are best fitted by different copula models using the information criteria measures especially the CAIC and CICOMP, where CICOMP predominated the AIC-type criterion, both in the case when the fitted models are correctly specified.We found then the Kendall's tau correlation between the return prices.

This paper can be extended later into several areas that are interesting especially in the credit risk modeling. For example, it can be realized that the value at risk is estimated properly when using the copula technique and Kendall's tau correlation more than when using the simple Pearson correlation method and it can detect how the Kendall's tau correlation matrix contributes in pricing credit derivatives.

Acknowledgements

The author is grateful to Professor Bozdogan for his guidance and introduction of the information complexity and his valuable comments which improved the paper. We also thank Dr. Michel Ballings for his assistance on R. The first draft of this paper was presented at the International Conference on Information Complexity and Statistical Modeling (IC-SMHD-2016) during May 18-21, 2016 in Cappadocia, Turkey. This paper received a 2nd best paper presentation award at IC-SMHD-2016 among student presentations.

References

Akaikei, H. (1973). Information theory and an extension of maximum likelihood principle. In *Proc. 2nd Int. Symp. on Information Theory* (pp. 267-281).

Bartram, S. M., Taylor, S. J., & Wang, Y. H. (2007). The Euro and European financial market dependence. *Journal of Banking & Finance*, *31*, 1461-1481. https://doi.org/10.1016/j.jbankfin.2006.07.014

Bedford, T., & Cooke, R. M. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial intelligence*, *32*, 245-268. https://doi.org/10.1023/A:1016725902970

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

Boero, G., Silvapulle, P., & Tursunalieva, A. (2011). Modelling the bivariate dependence structure of exchange rates before and after the introduction of the euro: A semi-parametric approach. *International Journal of Finance and Economics*, *16*, 357-374. https://doi.org/10.1002/ijfe.434

Bozdogan, H. (1987). Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, *52*, 345-370. https://doi.org/10.1007/BF02294361

Bozdogan, H. (2010). A new class of information complexity (ICOMP) criteria with an application to customer profiling and segmentation. *Istanbul University Journal of the School of Business*, *39*, 370-398.

Bozdogan, H., & Pamukçu, E. (2016). Novel Dimension Reduction Techniques for High-Dimensional Data Using Information Complexity. *INFORMS*, 140-170. https://doi.org/10.1287/educ.2016.0154

Brechmann, E. C., & Schepsmeier, U. (2013). Modeling dependence with C-and D-vine copulas: The R-package CDVine. *Journal of Statistical Software*, *52*, 1-27. https://doi.org/10.18637/jss.v052.i03

Cherubini, U., Luciano, E., & Vecchiato, W. (2004). *Copula methods in finance*. John Wiley & Sons. https://doi.org/10.1002/9781118673331

Durante, F., & Jaworski, P. (2010). Spatial contagion between financial markets: A copula-based approach. *Applied Stochastic Models in Business and Industry*, *26*, 551-564. https://doi.org/10.1002/asmb.799

Embrechts, P., McNeil, A., & Straumann, D. (2002). Correlation and dependency in risk management: Properties and pitfalls. In *Risk management: Value at risk and beyond* (pp. 176-223).

Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, *96*, 1348-1360. https://doi.org/10.1198/016214501753382273

Frey, R., & McNeil, A. J. (2001). *Modelling dependent defaults*. ETH Zurich.

Frey, R., & McNeil, A. J. (2003). Dependent defaults in models of portfolio credit risk. *Journal of Risk*, *6*, 59-92. https://doi.org/10.21314/JOR.2003.089

Grønneberg, S., & Hjort, N. L. (2014). The copula information criteria. *Scandinavian Journal of Statistics*, *41*, 436-459. https://doi.org/10.1111/sjos.12042

Haff, I. H. (2013). Parameter estimation for pair-copula constructions. *Bernoulli*, *19*, 462-491. https://doi.org/10.3150/12-BEJ413

Joe, H. (1996). Families of m-variate distributions with given margins and m (m-1)/2 bivariate dependence parameters. In *Lecture Notes-Monograph Series* (pp. 120-141). https://doi.org/10.1214/lnms/1215452614

Jondeau, E., & Rockinger, M. (2006). The copula-garch model of conditional dependencies: An international stock market application. *Journal of international money and finance*, *25*, 827-853. https://doi.org/10.1016/j.jimonfin.2006.04.007

Li, D. X. (2000). On default correlation: A copula function approach. *The Journal of Fixed Income*, *9*, 43-54. https://doi.org/10.3905/jfi.2000.319253

McNeil, A. J., Frey, R., & Embrechts, P. (2005). *Quantitative risk management: Concepts, techniques and tools* (Vol. 3). Princeton: Princeton university press.

Min. A., & Czado, C. (2011). Bayesian model selection for D-vine pair-copula constructions. *Canadian Journal of Statistics*, *39*, 239-258. https://doi.org/10.1002/cjs.10098

Necula, C. (2010). Modeling the dependency structure of stock index returns using a copula function approach. *Romanian Journal of Economic Forecasting*, *13*, 93-106.

Nelsen, R. B. (2006). Introduction. In *An Introduction to Copulas* (pp. 1-5). https://doi.org/10.7208/chicago/9780226572055.003.0001

Patton, A. J. (2006). Modelling asymmetric exchange rate dependence. *International economic review*, *47*, 527-556. https://doi.org/10.1111/j.1468-2354.2006.00387.x

Rosenberg, J. V., & Schuermann T. (2006). A general approach to integrated risk management with skewed, fat-tailed risks. *Journal of Financial Economics*, *79*, 569-614. https://doi.org/10.1016/j.jfineco.2005.03.001

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, *6*, 461-464. https://doi.org/10.1214/aos/1176344136

Sklar, A. (1973). Random variables, joint distribution functions, and copulas. *Kybernetika*, *9*, 449-460.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*, 583-639. https://doi.org/10.1111/1467-9868.00353

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society*, *Series B* (Methodological), 267-288.