## The Structural Difference of Shanghai Stock Index before and after 2008: A Copula Based Analysis

Chuan Wu<sup>1</sup>, Ke Huang<sup>2</sup>, Xue Tian<sup>2</sup>, Weikang Geng<sup>1</sup>, H. J. Cai<sup>1</sup>

<sup>1</sup>International School of Software, Wuhan University, Wuhan, China <sup>2</sup>School of Economics and Management, Wuhan University, Wuhan, China Email: hjcai@yahoo.com

Received September 5, 2012; revised October 9, 2012; accepted October 16, 2012

## ABSTRACT

The year 2008 witnessed the greatest joint stock reform and financial crisis in Chinese history. After these two cases, significant changes have taken place in investors' behaviors worldwide, along with which is the occurrence of structure change in stock market. In this paper, we employ Copula model to simulate the joint distribution between Shanghai Stock Index (SSE) and Chinese Shanghai Index 300 (CSI 300), to find out structure change in Chinese stock market before and after 2008. From results of empirical studies, we get conclusions that the main nature of Chinese stocks market is symmetric, in both marginal and joint distributions. Via the changes of Copula types, upper and lower tail coefficients and Kendall coefficients, we can measure the structure change in Chinese stock market, and get further conclusion about investors' behaviors change. Before 2008, there is an equal power in quitting market and longing, while diversified investors adjusted their expectation uniformly after this year. Testing results show that the general dependence structure of CSI 300 and SSE is highly dependent and symmetric in most cases. From the distribution of upper and lower tail coefficients, we can draw the conclusion that stratified investors are mainly focused on two tasks, after this year, to close the position on stocks with high correlated stocks market and to maintain market value of stocks.

Keywords: Copula Distribution; EM Algorithm; Kendall Tau Upper; Lower Tail

## **1. Introduction**

The year 2008 has imposed remarkable influence on stock market worldwide, because of two significant cases, the joint stock reform and financial crisis. The 2008 financial crisis created strong and deep psychological influence on investors, while serious problems were induced by the incomplete joint stock reform and unreasonable reallocation of shares between stock holders. CSI300 is a major reflection of overall price trend of Chinese stock market. It experienced a sharp decrease in 2008, from 5062.26 to 1606.73 in just 10 months. Similarly, another comprehensive index, SSE, also declined severely, from 4695.80 to 1664.93. Such remarkable change in Chinese stock market undoubtedly generated investors' distinct behaviors before and after 2008. Since the main nature of Chinese stocks market is the symmetric marginal and joint distributions, and great changes of investors' behaviors would contribute to the changes of marginal and joint distribution of stocks. In this paper, we will introduce Copula model to analyze the structure difference of Shanghai Stock Index before and after 2008.

Since CAPM and Black-Scholes models were proposed to simulate the joint distribution, many following researchers have found an obvious deviation of empirical

Copyright © 2012 SciRes.

evidence from the theoretical model [1]. Many papers tend to ascribe it to the assumption of normality [2]. Based on overwhelming empirical evidence and discussions, various alternative models with non-normality are proposed, one of which is the Copula model proposed by Sklar in 1959 [3]. It is an exact and overall statistical method to describe the dependent structure between random variables by linking the marginal distribution of each [4,5]. The introduction of Copula model can improve the properties of estimation to fit the volatility, provide flexibility to depict the fat tail, high skewness as well as other non-normality in distribution, and solve the dimension disaster when estimating multivariate joint distributions. Its special fitness in dealing with monotone mapping variables, describing linear or non-linear dependence structure, symmetric or asymmetric between upper and lower tail and showing time series characteristics of data dynamically, makes it a promising method and gaining more and more popularity in financial researches such as price derivatives, financial calculations, risk management and financial fluctuations description [6-8].

Gauss Copula, *t*-Copula, Frank Copula, Clayton Copula and Gumbel Copula are five candidates for bivariate static



Copula, and Gaussian mixture Copula is estimated as dynamic mixture Copula to improve the fitting. With fluctuations and changes in stock market, any solo Copula model fails to simulate the joint distribution of two stocks without loss of accuracy in long term. A structure change in stock market can be reflected in the changes of Copular types. In this paper, we will develop the Copular model to describe the structure changes before and after 2008 due to the influences of joint stock reform and financial crisis. Based on results of empirical study, we get further conclusion about changes of investors' behaviors.

#### 2. Copula Model Selection

Various Copula models are chosen to simulate the joint distribution of two stocks. They go into two major kinds: Elliptical and Archimedean Copulas.

#### **2.1. Elliptical Copulas**

#### 2.1.1. Bivariate Gaussian Copula

Bivariate Gaussian Copula is a symmetry distribution with asymptotically independent in both upper and lower tail. Because of the relatively easy estimation of parameters in this model, it fits well when the market does not fluctuate much. However, it ignores extreme incidence in the financial market.

Let  $\Phi^{-1}(u)$  be the inverse function of standard normal accumulative probability function and  $\rho \in (-1, +1)$ be the correlation coefficient. Bivariate Gaussian Copula model can be expressed as below. The Probability Distribution Function (PDF) and the Cumulative Distribution Function (CDF) of Bivariate Gaussian Copula for Shanghai Stock Index are as shown in **Figure 1**.

$$C_{g}(u,v,\rho) = \int_{-\infty}^{\Phi^{-1}(u)\Phi^{-1}(v)} \frac{1}{\sqrt{1-\rho^{2}}} e^{\left(-\frac{r^{2}+s^{2}-2\rho rs}{2(1-\rho^{2})}\right)} dr ds$$

#### 2.1.2. Gaussian Mixture Copula

Gaussian mixture Copula is a linear combination of Gaussian Copulas and also one of the applications of Hidden Markov Model (HMM). The HMM model describes a stochastic process with several possible states. The random variable jumps between these states randomly. The HMM with Gaussian mixture is divided into two parts: a Markov chain and a stochastic process. The Markov chain measures the state switching process between different states while a stochastic process describes a dynamic process in a certain state. Since no one can specify which state the variable is experiencing, the model is called Hidden Markov Chain.

Let  $\mu$ ,  $\sigma$  denote the volatility and mean of variables, which can vary in different states. Then  $\mu = (\mu_{11}, \mu_{12}, \mu_{13})$ means that there are three major states in the market for variable  $\mu_1$ , namely: good, medium and low condition. The random variable jumps between states with probabil-

ity  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$  and  $\sum_{i=1}^n \alpha_i = 1$ . Gaussian mixture Copula model can be expressed as below.



$$f(x_1, x_2) = \sum_{k=1}^{n} \alpha_k \frac{1}{2\pi\sigma_{1k}\sigma_{2k}\sqrt{1-\rho_k^2}} e^{-\frac{1}{2(1-\rho_k^2)} \left[\frac{(x_1-\mu_{1k})^2}{2\sigma_{1k}^2} - \rho_k \frac{(x_1-\mu_{1k})(x_1-\mu_{1k})}{2\sigma_{1k}\sigma_{2k}} + \frac{(x_2-\mu_{2k})^2}{2\sigma_{2k}^2}\right]}$$

#### 2.1.3. Bivariate t-Copula

Bivariate *t*-Copula is a symmetric distribution with two equal tails, and is one of the substitute models for Gaussian distribution. With  $\Gamma_v^{-1}(.)$  representing the inverse of *t* distribution with degree of freedom *v*. *t*-Copula have two parameters *v*,  $\rho$  determining the structural dependence. These two parameters can control the peaks and tails, thus enabling more freedom in depicting the non-normality in return distributions. Expression of this Copula model can be described as below. The PDF and CDF of Bivariate *t*-Copula for Shanghai Stock Index are shown in **Figure 2**.

$$C_t(u,v,\upsilon,\rho) = \int_{-\infty}^{\Gamma_v^{-1}(u)\Gamma_v^{-1}(v)} \left\{ 1 + \left(\frac{r^2 + s^2 - 2\rho rs}{\upsilon(1-\rho^2)}\right) \right\}^{\frac{\nu+2}{2}} drds$$

#### 2.2. Archimedean Copulas

#### 2.2.1. Gumbel Copula

Gumbel Copula is asymmetric with a probability density function fatter in the upper tail and thinner in the lower tail like a "J", which makes it very sensitive in the upper tail. In other words, there exists a better ability for it to describe the bull market. Let  $\alpha \in [0,1]$  be a coefficient of dependence, it can be expressed as below. If  $\alpha \rightarrow 0$ ,

then the two assets are highly correlated in the upper tail and vice versa. The PDF and CDF of Gumbel Copula for Shanghai Stock Index are shown in **Figure 3**.

$$C_G(u,v,\upsilon,\rho) = e^{\left\{-\left[\left(-\log u\right)^{1/\alpha} + \left(-\log v\right)^{1/\alpha}\right]^{\alpha}\right\}}$$

Gumbel Copula fits to simulate the dependence of two assets that are highly correlated in a bull market. However, since it is insensitive to changes in the lower tail, Gumbel Copula is not suitable in a bear market, less able to model lower tail, which is independent.

#### 2.2.2. Clayton Copula

Let  $\theta \in (0,\infty)$  be the coefficient to depict dependence. If  $\theta \to \infty$ , it indicates the two assets are highly dependent in lower tails which is just the reversed case of Gumbel Copula like and "*L*". The PDF and CDF of Clayton Copula for Shanghai Stock Index are shown in **Figure 4**.

$$C_{C}(u,v,\theta) = \left(u^{-\theta} + v^{-\theta} - 1\right)^{-\frac{1}{\theta}}$$

In contrast with Gumbel Copula, it is more sensitive to bear market, and specially fits to simulate the dependence structure of two assets in bear market.



Figure 2. The PDF and CDF example of bivariate t-Copula.





#### 2.2.3. Frank Copula

$$C_{F}(u,v,\beta) = -\frac{1}{\beta} \log \left( 1 - \frac{(1 - e^{-\beta u})(1 - e^{-\beta v})}{1 - e^{-u}} \right)$$

where  $\beta \in (-\infty, +\infty)$ ,  $\beta \neq 0$ . When  $\beta \rightarrow 0$ , the two assets are asymptotically independent. If  $\beta > 0$ , they are positively correlated while if  $\beta < 0$ , they are negatively correlated.

Frank Copula is symmetric, and can model both the

upper tail and the lower tail. However, it cannot capture asymmetric dependence. Unlike *t*-Copula, it has both fat tails and higher correlation in the middle of the distribution while t is thinner.

## 3. The Measure of Dependence Structure

#### 3.1. Kendall Rank Coefficient $\tau$

To analyze the correlation between variables, we introduce Kendall rank correlation coefficient ( $\tau$ ) which is represented by corresponding Copula function as

$$\tau = 4 \int_{0}^{1} \int_{0}^{1} C(u, v) dC(u, v) - 1$$

#### 3.2. Upper and Lower Dependence

Upper and lower dependence statistics are introduced to measure the association between two quantities. Define the joint survival function of Copula as C(:,:) where

$$\hat{C}(u,v) = u + v - 1 + C(1 - u, 1 - v),$$

Let *X*,*Y* be two continuous random variables with marginal distribution  $F(\cdot)$  and  $G(\cdot)$ . Specify the Copula as  $C(\cdot, \cdot)$ , then the Upper tail dependence  $\lambda^{up}$  and lower dependence  $\lambda^{low}$  are defined as follows:

$$\lambda^{\text{up}} \equiv \lim_{u \to 1} \lim_{u \to 1} P\Big[Y > G^{-1}u \Big| X > F^{-1}u\Big] = \lim_{u \to 1} \lim_{u \to 1} \frac{C(1-u,1-u)}{1-u}$$
$$\lambda^{\text{low}} \equiv \lim_{u \to 0} \lim_{u \to 0} P\Big[Y < G^{-1}u \Big| X < F^{-1}u\Big] = \lim_{u \to 0} \lim_{u \to 0} \frac{\hat{C}(1-u,1-u)}{1-u}$$

If  $\lambda^{up}$  or  $\lambda^{low}$  exists and  $\lambda^{up}$  or  $\lambda^{low} \in (0,1]$ , they measure the dependence structure in extreme cases; if  $\lambda^{up}$  or  $\lambda^{low} \to 0$ , then, they are statistically independent. It is very easy to observe whether the collapse of one asset's price will be followed or lead to the others' collapse with these measure dependences.

#### 4. The Estimation of Copula

#### 4.1. Estimation of Static Copula

As assumed in Sklar's theorem, if the marginal distributions of variables are all continuous, then there must exists one Copula that fits their joint distribution. Suppose F(u,v) is the joint distribution with marginal distribution  $F_1(\cdot)$ ,  $F_2(\cdot)$ . A unique Copula C(u,v) exists satisfying: F(u,v) = C(u,v).

The estimation of Copula and the marginal distribution can be very time consuming and cumbersome with ordinary MLE method. We employ a two stage MLE method to improve the accuracy and optimize the algorithm of static Copula, include Gauss Copula, *t*-Copula, Frank Copula, Clayton Copula and Gumbel Copula. Suppose there are n samples and the likelihood function is:

$$L = \prod_{i=1}^{n} \frac{\partial^{2} C_{IP}(u, v) | u = F_{iI}, v = F_{ip}}{\partial u \partial v} f_{iI} f_{iI}$$
$$c_{i} = \frac{\partial^{2} C_{IP}(u, v) | u = F_{iI}, v = F_{ip}}{\partial u \partial v}$$

where

and 
$$\ln L = \sum_{i=1}^{n} \ln c_i (u, v, \theta) + \sum_{i=1}^{n} \ln f_{ip} + \sum_{i=1}^{n} \ln f_{il}$$

The first stage is to estimate  $\sum_{i=1}^{n} \ln f_{ip}$  and  $\sum_{i=1}^{n} \ln f_{il}$ 

respectively with ordinary MLE methods. The second step is to transform the estimated  $f_{il}, f_{iP}$  into cumulative probability value in each sample and substitute them

into  $\sum_{i=1}^{n} \ln c_i(u, v, \theta)$ . Then the only unknown vector is  $\theta$ , which can be estimated by ordinary MLE.

#### 4.2. Estimation of Gaussian Mixture

The estimation method for Gaussian mixture different from that of static Copula. Here the Expectation-Maximization (EM) algorithm is employed to estimate the parameters of Gauss mixture. It is a MLE method where the model depends on unobserved latent variables. In addition, it is also an iterative method and therefore a computer program can be designed to solve them efficiently. In this case, the data is just like the shadow projected by many normal regimes and the target is to find how many regimes and what each regimes looks like exactly [9].

$$\mu_{k} = \frac{\sum_{i=1}^{n} x_{i} p(k | x_{i}, \theta_{t-1})}{\sum_{i=1}^{n} p(k | x_{i}, \theta_{t-1})}$$

where

$$\sigma_{k}^{2} = \frac{\sum_{i=1}^{n} (x_{i} - \mu_{k})^{2} p(k|x_{i}, \theta^{t-1})}{\sum_{i=1}^{n} p(k|x_{i}, \theta^{t-1})}$$
$$\sigma_{k}^{2} = \frac{\sum_{i=1}^{n} (x_{i} - \mu_{k})^{2} p(k|x_{i}, \theta^{t-1})}{\sum_{i=1}^{n} p(k|x_{i}, \theta^{t-1})}$$

where

$$\sigma_{k}^{2} = \frac{\sum_{i=1}^{n} (x_{i} - \mu_{k})^{2} p(k|x_{i}, \theta^{i-1})}{\sum_{i=1}^{n} p(k|x_{i}, \theta^{i-1})}$$
$$\sigma_{k}^{2} = \frac{\sum_{i=1}^{n} (x_{i} - \mu_{k})^{2} p(k|x_{i}, \theta^{i-1})}{\sum_{i=1}^{n} p(k|x_{i}, \theta^{i-1})}$$

The upper limit of regimes of Gaussian mixture is 10. For the determination of regimes of Gaussian mixture, the following procedures are suggested [10]:

1) Suppose the original regime umber is M (M starts from 1).

2) Estimated the parameters with assumption of M regimes.

3) Test the model with Kolmogorov-Smirnov (K-S) test under confidence level of 95%.

4) If the test is past, then the result is what we need. If not, let M = M + 1 and return to Step 1.

Results of univariate Gaussian mixture distribution will be candidates of marginal distribution, while that of Bivariate Gaussian mixture is used to simulate the joint distribution in our empirical study. In this paper, we use AIC, BIC and Chi square test to choose a more appropriate one among these models [11,12].

#### 4.3. The Estimation of t Copula

For the fitting extent of t Copula, there is a mature and precise way to estimate the parameters so that the time series effect can be taken into consideration. Patton firstly set up a research about how a Copula evolution equation, like ARMA (1,10) model, describes the actual trend of certain price process, leading to the study of time-varying Copula. Kole, Koedijk and Verbeek proposed some ways of choosing Copula functions. Following these studies, a Copula-GARCH model can be obtained as follows.

Previous research strongly suggest that GARCH (1,1)-*t* can reflect the lag in time series [13]. We continue to employ the same model in this paper. Let  $r_i$  be the returns of a stock, then GARCH (1,1)-*t* model can be solved as below:

$$r_{t} = C + \hat{u}_{t}^{2}$$
  $\hat{\sigma}_{t}^{2} = K + \alpha \hat{u}_{t-1}^{2} + \beta \hat{\sigma}_{t-1}^{2}$ 

If  $F(\cdot, \cdot, \dots, \cdot)$  is a *N*-dimension joint distribution function with marginal distribution  $F_1(\cdot)$ ,  $F_2(\cdot)$ ,  $\dots$ ,  $F_N(\cdot)$ . There exits a Copula function  $C(\cdot, \cdot, \dots, \cdot)$  such that

$$F(X_1, X_2, \cdots, X_N) = C(F_1(X_1), F_2(X_2), \cdots, F_N(X_N))$$

While studying time-varying situation, we usually consider correlation parameters of Copula evolution equation [14]:

$$\rho_{t} = \hat{\Lambda} \left( \omega_{\rho} + \beta_{\rho} \rho_{t-1} + \alpha_{\rho} \times \frac{1}{q} \sum_{i=1}^{q} \phi^{-1} \left( u_{t-i} \right) \phi^{-1} \left( v_{t-i} \right) \right)$$

#### 4.4. Kernel Smoothing Density Estimation

However, there is a possibility that both t and Gaussian mixture are not capable of passing the test, to simulate marginal distribution of stocks exactly. In this case, we will specify them as inability to find real model. But it

does not mean there is no approximate one to estimate. We use Kernel smoothing density estimation as an asymptotic distribution of those ones [15]. In essence, it is a weighted average between sample points by using kernel function as weighted functions. Generally, we employ the standard normal distribution as the kernel function for it is widely used in distribution [16]. It can also be shown that the estimated parameter is asymptotically unbiased estimator of the real distribution as n approaches positive infinity [17].

## 5. Modeling the Return

#### 5.1. Sample Description

In this paper, we choose Shanghai Stock Index and CSI 300 as a sample. Data from period 2005.6-2007.12 are treated as sub-sample before 2008 while 2008.1-2010.6 as the one after 2008. To calculate a specific return rate, we use  $r_i = \log P_{i,t} - \log P_{i,t-1}$ , where *i* is the trading code during CSI 300 and *t* is the specific trading day. Description of return rate data is shown in **Table 1**.

We employ two approaches to test whether the returns are normal, Q-Q plot and K-S test. If their distribution is normal, they should be consistent with the line y = x in the Q-Q plot. Moreover, in the Kolmogorov-Smirnov test, if the null hypothesis is true, namely the returns are normal, the test value should be greater than 0.01. According to the results in **Figure 5** and **Table 2**, both Q-Q plot and K-S test show a deviation from normality of return rate of SSE and CSI 300. In other words, these results contradict the traditional assumption of Brownian motion.

#### 5.2. Modeling the Marginal Distribution

From the above discussion, non-normal models are appropriate to simulate marginal distributions. In this paper, we mainly use t distribution and univariate Gaussian mixture distribution model. The former one can better model the fat tails if they are symmetric while the latter is more suitable to model the asymmetric tails. In t distribution model, the fatness of the tails is variable and controlled by the value of parameter nu. A small nu gives heavy tails. While for Gaussian mixture, it mainly depends on the different regimes to model the asymmetric tails. Theoretically, the more regimes employed in the model, the more precisely the simulation will be (of course it is possible to over specifying, so that AIC and

Table 1. The summary statistics of return rate.

	Mean	Variance	Skewness	Kurtosis	Max	Min
SSE	0.0025437	0.0003866	-0.8096452	6.970561	0.0690344	-0.1107323
CSI300	0.0029191	0.0004248	0.0696865	-0.1187816	-0.8062138	7.147056



Figure 5. The Q-Q plot of the return rate series.

Table 2	. Kolmogorov-	Smirnov tes	t for no	rmality.
---------	---------------	-------------	----------	----------

Skewness/Kurtosis Tests for Normality						
Variable	Pr (Skewness)	Pr (Kurtosis)	adj chi2 (2)			
SSE	0.000	0.000	0.000			
CSI 300	0.000	0.000	0.000			

Table 3. Types of marginal distribution of constituent stocks of CSI300 before and after 2008.

Marginal Distribution					
t Gaussian Mixture Invalid					
Before	41	158	101		
After	84	207	9		

BIC will set up punitive function to deal with over specifying problem). For t distribution, ordinary MLE method is enough to estimate parameters while a simpler EM algorithm estimates the parameters of the subsequent one.

We apply these distribution models to simulate the marginal distribution of constituent stocks of CSI300 before and after 2008, and results of their types are shown in **Table 3**. From the "invalid" figures, it can be easily inferred that the ability of t and Gaussian mixture to model the return after 2008 is better than that before 2008. Actually, neither of t distribution nor Gaussian Mixture distribution model fits to simulate the marginal distribution of these 300 constituent stocks before this year, thus we use Kernel smoothing density estimation for simulation. After 2008, we see that the numbers of both t and Gaussian mixture distributions increase greatly and they simulate their distributions well.

The changes of marginal distribution reflect the structure change in the stock market, which can be explained by the hedging strategy and other behaviors of investors in this particular year. Before 2008, we were not able to simulate returns of a large proportion of stocks owing to the lack of uniform behavior from traders and other investors. Owing to different expectation towards stocks, there might be different peaks in the distribution, while a slight decline will cause different reactions: some quit but some continue to hold long stocks. When the dominating power switches, the divergence from normality was hard to capture. However, after 2008, the behaviors of investors are much more uniform. The increased number of t distributions illustrates the symmetric characteristics in both upper and lower tails, reflecting the equal power in quitting market and longing. Fewer investors stay put during 2008. For Gaussian mixtures, a sudden increase might suggest that stratified investors start to dominate the market after 2008. Although they have different preference for risk, after the crisis, diversified investors adjusted their expectation uniformly so that the regimes switching model are specified. We simulate the marginal distribution of CSI300 constituent stocks, and find that large part of them change marginal distribution types after 2008. Detailed information of a small sample is shown in **Table 4**.

#### 5.3. Modeling the Joint Distribution of Copula

After fitting the marginal distribution, then the joint distribution can be found by the two stage MLE method. The general results are shown in the **Table 5**.

From **Table 5**, one interesting fact is that both Clayton and Gumbel Copula that are good at modeling bull and bear market are not specified here. On the contrary, the results of model fitting mainly show the symmetric dependence structure both before and after 2008. Before 2008, the *t*-Copula is most frequent, while after 2008, *t* and Frank Copula dominate. The symmetric dependence between SSE and CSI300 constituent stocks has one more characteristics, which can be described by Frank Copula model, after 2008. Then, it can be concluded that the general dependence structure of CSI300 and SSE is symmetric in most cases.

#### 5.4. Measure of Correlation in Overall Level

Besides, the positive correlation can also be found in the Kendall tau rank correlation. The distribution of Kendall tau rank correlation of CSI300 and SSE is shown in **Figure 6**. In fact, the rank correlation coefficients show a highly explicit dependence structure of CSI 300 and SSE, which gives the reason why symmetric Copula is prevailing.

Table 4. Marginal distributions of constituent stocks ofCSI300 before and after 2008.

Code	Before	After	Code	Before	After
SH600016	-	t	SZ000002	GausMix	t
SH601328	GausMix	t	SZ000858	GausMix	t
SH600000	GausMix	t	SH600837	GausMix	t
SH601166	GausMix	t	SH601398	GausMix	t
SH600519	t	GausMix	SH601601	-	GausMix
SH601088	-	GausMix	SH601668	-	t

Table 5. Frank and Gumbel Copula dominate after 2008 and not *t* and Frank.

	Copula types					
	Gaussian	t	Clayton	Frank	Gumbel	
before	4	162	46	15	2	
after	8	150	14	116	4	



statistics interval

Figure 6. The Kendall tau of CSI300 and SSE.

# 5.5. Measure of the Dependence in Extreme Cases

The dependence structure in extreme cases is a very important issue in empirical studies. For example, the joint stock reform and financial crisis are extreme cases, and here we employ upper and lower tails coefficient to figure out the changes in structure of Chinese stock market. Statistics about a small sample from CSI300 constituent stocks are shown in **Table 6**.

Collect test results of CSI 300 constituent stocks with SSE, we get the statistical results in **Table 7**. It shows that the value of upper and lower tails coefficients increase after 2008, which implies a stronger tendency that CSI300 and SSE are more correlated with each other in extreme cases. In both cases, upper tails and lower tails are scattered similarly, echoing to the prior finding of the symmetric Copula. The distribution of upper tail and lower tail are very similar, as shown in **Figure 7**. Before 2008, they are most aggregated in 0.2 - 0.4, while after 2008; there are two intervals that are aggregated 0 - 0.2 and 0.4 - 0.6.

From other points of view, if the same categories of coefficient were put together, it might bring our attention that another structural change is the extreme cases for some stocks, like those stocks with lower tail lying in 0.6 - 0.8. The main nature of CSI 300 after 2008 can be divided into two types: the first one is that a majority of constituent stocks in CSI300 continue to perform against the trend of SSE, or simply suffer from a small slump. For some additional stocks, the second large proportions of them are highly correlated with SSE, which in turn suffer a great loss, just like SSE.

Therefore, one important fact deduced from it is that the formerly mentioned stratified investors are mainly focused on two tasks. The first one is to close the position on these stocks with high correlated stocks market.

Table 6. Acomparison of sample stocks' t	test results before and after 2008.
--	-------------------------------------

Code	Copula	Copula type		Upper tail		Lower tail		Kendall rank	
Coue	before	after	before	after	before	after	before	after	
SH600036	Gumbel	t	0.5377	0.4793	0.1171	0.4793	0.6222	0.8061	
SH600016	Frank	t	0.0562	0.5342	0.0562	0.5342	0.7068	0.8307	
SH601318	Clayton	t	0.0201	0.5060	0.5120	0.5060	0.4868	0.8118	
SH600000	Frank	t	0.0562	0.4661	0.0562	0.4661	0.7070	0.7971	
SH601088	-	t	-	0.5501	-	0.5501	-	0.8064	
SH601398	Gaussian	t	0.3652	0.5000	0.3652	0.5000	0.7765	0.7920	
SH601601	-	Frank	-	0.0869	-	0.0869	-	0.8488	
SH601668	-	t	-	0.3418	-	0.3418	-	0.6935	
SH601939	-	t	-	0.5519	-	0.5519	-	0.8098	
SH601169	-	t	-	0.5126	-	0.5126	-	0.8112	
SZ000001	t	Frank	0.3165	0.0723	0.3165	0.0723	0.6563	0.7957	

 Table 7. A comparison of upper and lower tails coefficient before and after 2008.

Value	Bet	fore	After		
value	Upper tail	Lower tail	Upper tail	Lower tail	
0 - 0.2	93	49	141	132	
0.2 - 0.4	130	136	62	65	
0.4 - 0.6	6	41	88	89	
0.6 - 0.8	0	3	1	6	
0.8 - 1	0	0	0	0	



Figure 7. The comparison of upper and lower tail coefficient before and after 2008.

These strategies intensified the collapse of the market. The second one is to maintain market value of stocks so that these stocks perform a less correlated dependence structure.

#### 5. Conclusion

This paper mainly discusses the structural changes before and after 2008 with the application of bivariate Copula to measure the joint dependence between CSI300 and SSE. Firstly, the non-normality characteristic is found in the marginal distribution. Before 2008, investors' behaviors are so various that marginal distribution models fail to simulate them well, but after it, the marginal distribution is better specified by t and Gaussian mixture distributions. Although there is a collapse of market, the return distribution remains relatively symmetric by comparison of types of Copula. It might be explained by the stratified investors' distribution. There is a group of investors who want to quit the market or close the position while another requires hedging strategies to maintain the value. In addition, we examine the value of Kendall tau, upper and lower tail to support the investor's constitution. The Kendall tau shows a medium correlation of CSI 300 and SSE. After 2008, the upper and lower tail also shows a two-peak in its value interval distributions. Thus, it can be concluded that the structural change, with a sudden change of the tendency of SSE, leads to a multidimensional influence on the behavior of CSI 300 and investors. The peer effect might gather uniformly distributed types of investors into several main groups that remain universal nature in a certain group. With this change in investors' expectations and strategies, the marginal distribution of return deviates from normal distribution symmetrically and asymmetrically. And the joint distribution is more symmetric with medium correlation of market because two powers are offset to a certain extent.

## REFERENCES

- C. Jacky, "Some Empirical Evidence on the Outliers and the Non-Normal Distribution of Financial Ratios," *Journal of Business Finance & Accounting*, Vol. 14, No. 4, 1987, pp. 483-496. doi:10.1111/j.1468-5957.1987.tb00108.x
- [2] X. Z. Bai, "Beyond Merton's Utopia: Effects of Non-Normality and Dependence on the Precision of Variance Estimates Using High-Frequency Financial Data," The University of Chicago, Chicago, 1999.
- [3] A. Sklar, "Fonctions de Repartition à *n* Dimensions et Leurs Marges," Publication de l'Institut de Statistique de l'Université de Paris, Paris, 1959, pp. 229-231.
- [4] R. B. Nelsen, "An Introduction to Copulas," 2nd Edition, Springer, Berlin, 2006.
- [5] E. Jondeau and M. Rockinger, "The Copula-GARCH Model of Conditional Dependencies: An International Stock-Market Application," *Journal of International Money and Finance*, Vol. 25, No. 5, 2006, pp. 827-853. doi:10.1016/j.jimonfin.2006.04.007
- [6] H. Joe, "Multivariate, Models and Dependence Concepts," Chapman & Hall, London, 1997.
- [7] A. J. Patton, "Estimation of Multivariate Models for Time Series of Possibly Different Lengths," *Journal of Applied Econometrics*, Vol. 21, No. 2, 2000, pp. 147-173. doi:10.1002/jae.865
- [8] A. J. Patton, "On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation," *Journal of Financial Econometrics*, Vol. 2, No. 1, 2004, pp. 130-168. doi:10.1093/jjfinec/nbh006
- [9] A. Dempster, *et al.*, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal*

Statistical Society, Series B, Vol. 39, No. 1, 1977, pp. 1-38.

- [10] C. Genest, J.-F. Quessy and B. Rémillard, "Goodnessof-Fit Procedures for Copula Models Based on the Probability Integral Transformation," *Scandinavian Journal of Statistics*, Vol. 33, No. 2, pp. 337-366. doi:10.1111/j.1467-9469.2006.00470.x
- [11] F. X. Diehoid, et al., "Evaluating Density Forecasts with Application Financial Risk Management," International Economic Review, Vol. 39, No. 4, 1998, pp. 863-883. doi:10.2307/2527342
- [12] I. Karatzas and S. Shreve, "Methods of Mathematical Finance," Springer-Verlag, Berlin, 1998.
- [13] H. J. Cai, H. Sun, C. Wang and Y. Wang, "A Copula-GARCH Analysis of Chinese Stock Market Dependence Structure," Advances in Intelligent and Soft Com-

puting, Vol. 136, 2012, pp. 589-596.

- [14] C. Jacky, "Some Empirical Evidence on the Outliers and the Non-Normal Distribution of Financial Ratios," *Journal of Business Finance & Accounting*, Vol. 14, No. 4, 1987, pp. 483-496. doi:10.1111/j.1468-5957.1987.tb00108.x
- [15] J. Durbin, "Weak Convergence of the Sample Distribution Function When Parameters Are Estimated," *The Annals of Statistics*, Vol. 1, No. 2, 1973, pp. 279-290. doi:10.1214/aos/1176342365
- [16] B. L. S. PrakasaRao, "Nonparametric Functional Estimation," Academic Press, Orlando, 1983.
- [17] B. W. Silverman, "Density Estimation for Statistics and Data Analysis," Chapman & Hall, London, 1986.