Research Article **JTAE (2020) 4:5**

Journal of Theoretical and Applied Economics
(ISSN:2576-3970)

Studying the Correlation of Stocks via Copula Function

Yishuai Tian[1], Boying Lv[1], Botao Liu[1]

[1]College of Hydraulic & Environmental Engineering, China Three Gorges University, Yichang 443002, China.

ABSTRACT

Many factors affect the value of stocks, but there are few studies on the correlation between stocks and stock indexes. Therefore, this paper selects the closing prices of 4 industries related to Everbright Bank (601818) and 15 sectors related to the market index from 2017 to 2018, carries on the correlation analysis based on the Copula function, and compares the correlation coefficient by calculation. The stocks closely related to Everbright Bank are mainly bank stocks such as Bank of China and Ningbo Bank, while those greatly affected by the Shanghai Composite Index are Yunnan germanium industry, Tongji Technology, Futian Automobile and other technology and manufacturing stocks. Based on the Copula function to explore the correlation between stocks and the influencing factors, to provide a specific research basis for stock correlation analysis.

Keywords: Copula function; Correlation analysis; Stock market index

***Correspondence to Author:**
Yishuai Tian
University of Chinese Academy of Social Sciences, 11th, Beijing Fangshan District.

How to cite this article:
Yishuai Tian, Boying Lv, Botao Liu. Studying the Correlation of Stocks via Copula Function. Journal of Theoretical and Applied Economics, 2020; 4:5

eSciPub LLC, Houston, TX USA.
Website: https://escipub.com/

1. Introduction

The market price of stocks is not only determined by the value of shares, but also influenced by many other factors, such as the profitability of listed companies, major emergencies, and product market competition [1-4]. Besides, the law of stock price changes is complicated, and a large number of technical indicators and methods are often used for frying stocks. In China, there are k-line charts, market indexes, trend analysis, and morphological Analysis [5], while in the United States, there are dow jones industrial average index and NASDAQ composite index [6]. To discuss the stocks affected by other commodities and the stock market index is conducive to the investment of listed companies and to materialize the relationship between stocks and the stock market index. Therefore, it is of considerable significance to dig deeply into the correlation between stocks.

There have been many studies on the correlation and impact of the Analysis of stocks. For example, Song Jia [7] deeply discussed the integration of R & D and innovation, market competition, and stock performance into a unified whole. Based on the research of A-share listed companies, this paper gives a general conclusion that applies to the Chinese market. Yilun Qian [3] takes Vanke as a typical example, makes an in-depth analysis of the factors related to the changes in the stock prices of real estate companies, and finds that the Shanghai Composite Index has a far-reaching impact on it. Hao Ding [8] used the multiple linear regression method to study the factors that affect the stock price of Guizhou Moutai from 2001 to 2016. This paper analyzes the impact of GDP, Shanghai Composite Index, liquor industry output, and the net profit margin of Guizhou Moutai Co., Ltd on Guizhou Moutai stock price. Zongjing Liang et al. [9] aimed at the "clenbuterol" food safety incident that occurred on March 15, 2011, and used Baidu Index as an online attention index. To study the impact of food safety incidents based on the Baidu index on food stock prices in Shanghai and Shenzhen. Van et al. [10] found that the influence of high-frequency traders on stock price differences is not uniform and varies with the toxicity level of order flow and the number of stocks. The test of Blau [11] shows that religious belief and a lower degree of religious belief and religious belief will harm the volatility of ADR. Gao et al. [12] applied the linear correlation coefficient to quantify the dynamic evolution of the correlation structure between the American market and the Asian market. This paper studies the impact of the 2008 financial crisis on Asian markets. Rubin [13] and others use the latest innovative technologies in the field of network science to analyze the correlation of stock returns and how they evolve. Predecessors have studied how the prices of many stocks are affected by various factors, but at present, there is a lack of case studies on stocks in which a stock is affected by other stocks and related indexes.

In-depth mining of the correlation analysis between stocks has a significant impact on the judgment of the stock market. For this reason, we establish a stock correlation model based on Copula function and take stock of China Everbright Bank (601818) as an example to analyze the correlation. Based on the Copula function model, the correlation coefficients between stocks are calculated and compared. And through the same method, the Shanghai Stock Index and the Shenzhen Stock Index of the large-cap index are selected for correlation analysis. It provides a specific research basis for the correlation analysis of stock price.

2. Establishment of Stock correlation Model based on Copula function

As the joint distribution function of uniform distribution, Copula function [14] can connect the edge distribution of several random variables and get their joint delivery. The correlation between the two sets of variables can also be expressed. Therefore, this paper is based on the Copula function to describe the association between stocks and the correlation of the representative index.

The correlation between stocks can be measured by constructing the Copula function of the stock sequence. Because the stock data has the characteristics of

time-variability, non-stationarity, and complexity, it is difficult to make assumptions about the Copula function describing the structure between variables. Therefore, the nonparametric kernel function is used to estimate the Copula function.

(1) Determining N-element kernel function，Let $X=(x_1,x_2,...,x_n), Y=(y_1,y_2,...,y_n)$ variables of two-time series, respectively，$z=(z_1,z_2,...,z_n)$ is a random vector. $k_n(z_n)$ Is the corresponding unary kernel function.

Among them, $n=1,2,...,N$，And the kernel function satisfies $\int_{-\infty}^{+\infty} k(x)dx=1$，Then the N-element kernel function is as follows

$$k(z;h)=\prod_{n=1}^{N} k_n(\frac{z_n}{h_n})$$

Where h_n are the width of the kernel function and a positive function of the sample size. The h_n matrix is as follows:

$$h_n=\begin{pmatrix} h_1 & 0 & \cdots & 0 \\ 0 & h_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_n \end{pmatrix}$$

(2) Determine the joint distribution density function. The edge density function of time series X, as follows:

$$\hat{f}_x(x)=\frac{1}{nh_x}\sum_{t=1}^{n} k_x(\frac{x-x_t}{h_x})$$

The edge density function of time series Y, as supports:

$$\hat{f}_y(y)=\frac{1}{nh_y}\sum_{t=1}^{n} k_y(\frac{y-y_t}{h_y})$$

The estimated value of the density function of the edge distribution at the point (x_t,y_t), as follows:

$$\hat{f}(x,y)=\frac{1}{n|h|}\sum_{t=1}^{n}(k_y(\frac{y-y_t}{h_y})\bullet k_x(\frac{x-x_t}{h_x}))$$

The joint distribution function at a point (x_t,y_t) is:

$$\hat{F}(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}\hat{f}(x,y)dxdy$$

(3) Determine the Copula function. According to the estimated value of the joint distribution function, the value of the Copula function at the point (u,v) can be expected to be

$$\hat{C}(u,v)=\hat{F}(f_x^{-1}(u),f_y^{-1}(v))$$

Among them,

$f_z^{-1}(z)=\inf_{a\in R}(a\mid \hat{F}=(a)\geq z)$.

(4) After determining the correlation function of the time series and obtaining the Copula function of the time series, we can calculate that the correlation between the two-time series is as follows:

$$\tau_{xy} = 4\int_0^1\int_0^1 C(u,v)dC(u,v) - 1$$

3. Case Analysis based on Everbright Bank and Market Index

By selecting the stocks of the industries related to the capital or the stock index, and calculating the correlation coefficient between the capital or the stock index and the stocks of the relevant industrial index, we can judge the correlation of the capital. The following is to take the shares of Everbright Bank (601818) as an example to discuss the impact of other stocks on its prices and which stocks are affected by the market index. First of all, select the industries related to Everbright Bank, as well as other influential stocks in the industry. Through the random selection of various types of stocks with a wide range of influence, so as to measure the impact of stocks on the stock index, and select the market index to measure the effect. The Copula function obtains a fluctuation trend [14], which can use different index data, so only the index, the closing price of the stock, is selected to represent the impact on Everbright stock.

3.1 Application of Everbright Bank (601818)

3.1.1Selection of related stocks

CHINA Everbright BANK[15], established in August 1992, is one of the subsidiaries of China Everbright Group and a ministerial company directly under the State Council. It is a financial enterprise approved by the State Council and approved by the people's Bank of China. It is the first national joint-stock commercial bank in China that is state-controlled and has shares in international financial organizations. This paper selects the closing prices of nine typical stocks, such as technology, real estate, energy, and similar banks, such as Bank of China (601988), Shenzhen Huaqiang (000062), and so on. The details are shown in Table 1.

Table 1 selection of stocks related to Everbright Bank

Serial number	Stock name	Stock type
1	Bank of China	Financial category
2	PetroChina	Energy category
3	Shenzhen Huaqiang	Manufacturing industry
4	Shenzhen Energy	Energy category
5	Quantum Biology	Science and technology category
6	Ningbo Bank	Financial category
7	Shanghai Electric Power	Energy category
8	Wayne Share	Science and technology category
9	Zhongfang Shares	Manufacturing industry

3.1.2 Based on the solution result of Everbright Bank

The closing prices of shares related to Everbright Bank in 2017 and 2018 are substituted into the Copula function to calculate the correlation of the time series. The calculated relationship of each stock is shown in Table 2.

Table 2 Stock correlation

Ranking	Stock name	Copula correlation coefficient
1	Bank of China	0.7852
2	Ningbo Bank	0.6659
3	Shanghai Electric Power	0.4561

4	Wanye Share	0.4114
5	PetroChina	0.4001
6	Shenzhen Huaqiang	0.3910
7	Shenzhen Energy	0.1992
8	Quantum biology	0.1692
9	Zhongfang shares	0.0757

As can be seen from Table 2:

(1) The correlation coefficient of the Bank of China is the largest, and its correlation is the strongest. The shares of companies that go with Everbright Bank have the most significant impact on Everbright Bank shares. The stocks most closely related to Everbright Bank are Bank of China 601988, Ningbo Bank 002142, and Shanghai Electric Power 600021. The closer stocks are Wanye Enterprise 600641, PetroChina 601857, and Shenzhen Huaqiang 000062.

(2) The related shares of enterprises in the energy sector have a significant impact on Everbright Bank, and it is the layman's stock that has the most significant effect on Everbright Bank shares in other industries except for banks.

(3) The influence of industrial stocks that have an effect on Everbright Bank shares is ranked from large to small, followed by banking, energy, technology, and real estate.

3.2 Application of the market index

3.2.1 Selection of related stocks

The market index generally refers to the Shanghai Composite Index of the Shanghai Stock Exchange and the Shenzhen component Index of the Shenzhen Stock Exchange. It can scientifically reflect the situation of the whole stock market, such as the overall rise and fall of the stock or the trend of the stock price. If the market index rises gradually, it can be judged that most stocks are rising. On the contrary, if the index slowly falls, that is, most stocks are falling. [16, 17] The market index is calculated by the index of individual stocks because the weight allocated to individual stocks is different, the impact of shares on the index is small and large, affected by a variety of industries. This paper selects 30 representative stocks in 15 main fields, such as the Yunnan germanium industry 002428, Jianglong boat 300589, and so on. The details are shown in Table 3.

Table 3 selects all kinds of stocks related to the market index

NO.	Domain	Stock name	NO.	Domain	Stock name
1	Non-ferrous metal	Yunnan germanium industry	16	Environmental tourism	Changbai Mountain
2	Construction business	Tongji science and technology	17	Waste utilization	Dongjiang environmental protection
3	Automobile manufacturing	Fukuda Motor	18	Communication equipment	Shenyu shares
4	Non-ferrous metal	Guiyan platinum industry	19	Food manufacturing	Yili shares
5	Financial industry	Shanxi Securities	20	Delivery equipment	Xinri shares
6	Education	Kaiyuan shares	21	Waste utilization	Pengqi science and technology
7	Environmental tourism	Shanghai environment	22	energy	Shanghai Electric Power
8	Textile industry	Xinlong Holdings	23	Construction business	Kolida
9	Real estate	Wanye enterprise	24	Energy	PetroChina
10	Automobile manufacturing	Xinquanshares	25	Health	American year health
11	Communication	Jiuzhiyang	26	Education	Kevin education

	equipment				
12	Media industry	Hubei Radio and Television	27	Real estate	Yuetaishares
13	Financial industry	Wuxi Bank	28	Health	Tiger Medicine
14	Delivery equipment	Jianglong boat	29	Media industry	Visual China
15	Textile industry	Norbangshares	30	Food manufacturing	Haitian flavor industry

3.2.2 The solution result of the model

Because the market index includes Shanghai index and Shenzhen index, this paper calculates the correlation coefficient between Shanghai index and Shenzhen index and the selected stocks respectively, and adds the sum of Shanghai index and Shenzhen index, obtains the Copula value of each stock and the market index, and measures the influence degree of the market index by the combined value. Using MATLAB to achieve the solution of the correlation coefficient, get all the stocks related to the market index, as well as the stocks greatly affected by the market index, the specific ranking and correlation are as follows.

Table 4 the relationship between the market index and all kinds of stocks

Rank	Stock name	value	Rank	Stock name	value
1	Yunnan germanium industry	1.3441	16	Changbai Mountain	0.9072
2	Tongji science and technology	1.3420	17	Dongjiang environmental protection	0.8202
3	Fukuda Motor	1.3183	18	Shenyushares	0.7779
4	Guiyan platinum industry	1.2276	19	Yili shares	0.7664
5	Shanxi Securities	1.2162	20	Xinri shares	0.7097
6	Kaiyuan shares	1.2103	21	Pengqi science and technology	0.7054
7	Shanghai environment	1.1604	22	Shanghai Electric Power	0.6835
8	Xinlong Holdings	1.1393	23	Kolida	0.6742
9	Wanye enterprise	1.1361	24	PetroChina	0.6108
10	Xinquan shares	1.1299	25	American year health	0.6045
11	Jiuzhiyang	0.9915	26	Kevin education	0.5952
12	Hubei Radio and Television	0.9790	27	Yuetaishares	0.5936
13	Wuxi Bank	0.9716	28	Tiger Medicine	-0.0490
14	Jianglong boat	0.9323	29	Visual China	-0.4240
15	Norbang shares	0.9178	30	Haitian flavor industry	-0.5311

It can be seen from Table 4 that stocks in non-ferrous metals, construction, automobile manufacturing, non-ferrous metals, financial industry, education, environmental tourism, textile industry, real estate, and other industries are most affected by the market index. Among them, the Copula value of the Yunnan germanium industry, Tongji science and technology, Futian Automobile, your research platinum industry, Shanxi Securities stocks, and the market index reached more than 1.2, which was greatly affected. The stock correlation values of health, media industry, and food manufacturing industry are negative, indicating that these stocks may have a negative correlation with the market index; that is, the market index rises and stock prices fall.

4. Result Analysis

4.1The result Analysis of the correlation of Everbright Bank

The correlation value calculated by the Copula function of the stock closing price of the selected industry is arranged in descending order, the Abscissa represents the ranking number, and the corresponding industrial stock draws a two-dimensional scatter chart as shown in Fig.1.

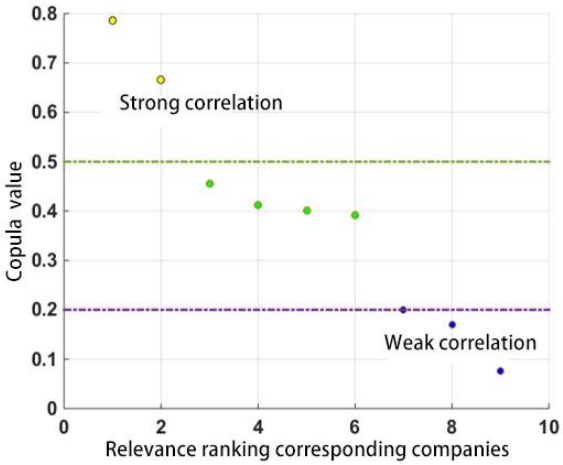

Fig. 1 scatter graph of the correlation between industrial and Everbright stocks

It can be seen from Fig.1 that the yellow distribute represents the industrial stocks of the same industry, that is, the shares of banks. From the figure, we can see that the Copula values of bank stocks are all above 0.6, and the correlation is powerful. This is because the value transmission power between peer industries is relatively stable, the financial transaction rate is high, and the competition is fierce, resulting in a series of transmission systems, such as stock price influence between each other, market arbitrage, and so on. The blue scatter represents the stocks of the science and technology industry, and it has a weak correlation with the majority of banks, that is, it shows that the market value transmission rate between the technology industry and Everbright bank is low, and the influence-oriented lag is substantial. Therefore, due to the influence between listed companies, the impact between stocks is also little.

4.2 The result Analysis of the correlation of Market Index

The correlation value calculated by the Copula function of the stock closing price of the selected industry is arranged in descending order, and the Abscissa represents the ranking of stock correlation. According to the range of the value, it can be divided into strong correlation influence, weak correlation influence, and negative correlation influence, that is, [-1，0] is negative correlation influence,[0，0.8] is a weak correlation influence, and[0.8，1.5] is strong correlation influence. The stock of the corresponding industry draws a two-dimensional scatter diagram, as shown in Fig.2.

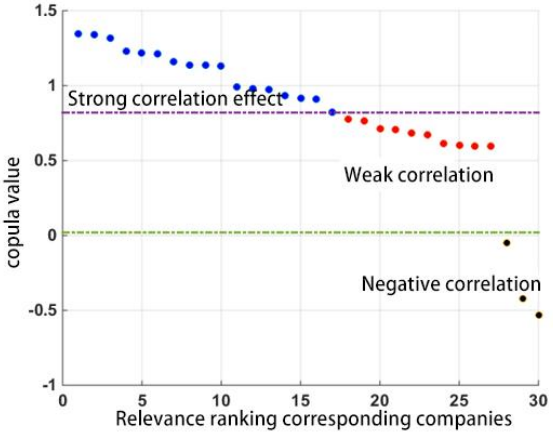

Fig. 2 correlation between industrial stocks and market index

According to figure 2, we can see that:

(1) The industry categories represented by the blue scatter are the science and technology industry, energy industry, service industry, and manufacturing industry. Because of its emerging characteristics or the nature of crucial social supply, this kind of production has a robust mutual reaction and mutual value transmission with all aspects of society. The market index has a significant impact on the stocks of these industries.

(2) The industry categories represented by the black scatter are the health care industry, food industry, and other related primary industries, which have a stable market value because of their essential contribution characteristics and small elasticity of product demand. Under the background of the weakening Engel coefficient, it is easy to be affected by the negative impact of the stock market environment, that is, the negative correlation with the market index.

5. Model Evaluation of relationship based on Copula function

5.1 Advantages of the model

(1) Under the condition that it is impossible to determine whether the traditional linear correlation coefficient can correctly measure the correlation between variables, the appearance of the Copula function makes the description of dependence between variables more perfect.

(2) Based on the Copula function model, it has been widely used in the Analysis of relationship between returns on financial assets and financial risk management.

(3) After the establishment of the stock correlation model, combined with the application of specific examples, it shows that the model has good feasibility and can be applied to the correlation analysis of stocks.

5.2 Model shortcomings

(1) Lacking the stability analysis of the Copula function model, the testing of the model needs to be strengthened.

(2) The model selects the primary industries related to banks and selects the influential stocks in each sector to analyze the correlation directly. The subjective factors are substantial and can be improved by the algorithm.

6. Conclusion

We established a correlation analysis model based on the Copula function. Taking Everbright Bank as an example, it concludes that Bank of China has the strongest correlation with Everbright Bank stock. The stocks of Everbright Bank have the most substantial influence on the shares of banking enterprises, and the related shares of enterprises in the energy sector are the layman stocks that have the most significant impact on them in other industries. Similarly, for the Analysis of the market index, the shares of non-ferrous metals, construction, automobile manufacturing, non-ferrous metals, financial sector, education, environmental tourism, textile industry, real estate, and other industries are most affected by the market index. The stock correlation values of health, media industry, and food manufacturing industry are negative, indicating that these stocks may have a negative correlation with the market index, that is, the market index rises, and stock prices fall. This provides a particular reference for studying the impact of stock prices among stocks and the stocks affected by the market index.

References

1. Yuan Yao, Beibei Yao, Qi Zhong. A study on the influence of Investor sentiment on Stock return--an empirical Analysis based on Shanghai A-share data [J]. Price theory and practice,2019 (05):88-91.
2. Rongyu Deng. The impact of profitability of listed companies on stock prices [J]. The Chinese market,2019(25):42-44.
3. Yilun Qian. Analysis on the influencing factors of Vanke Stock Price-- based on the empirical Analysis of listed companies in Real Estate Industry [J]. Paying taxes, 2019,13(05):166-167.

4. Huixian Zhang. Study on the relationship between Shanghai Composite Index and other Shanghai Classification Index [J]. Modern Marketing (Information Edition), 2019(01):12.
5. Hanlin Ji, Meng Xu. Research on the dynamic correlation between Financial sector and Shanghai Stock Index-- based on VAR-DCC-GARCH Model [J]. Prices in China, 2020(01):83-85.
6. S&P Dow Jones Indices; NeoGenomics Set to Join S&P SmallCap 600[J]. Genetics & Environmental Business Week, 2019.
7. Jia Song. Product market competition, R & D investment and stock return [D]. Shanxi University, 2018.
8. Hao Ding. Empirical Analysis on the influencing factors of stock price change in Guizhou Moutai [J]. Modern economic information,2018(12):304.
9. Zongjing Liang, Yonghui Pang, Yun Kuan. Study on the influence of Food Safety incident Network on Food Stock Price [J]. Commercial economy 2019(10):177-180.
10. Van Ness B F, Van Ness R A, Yildiz S. The role of HFTs in order flow toxicity and stock price variance, and predicting changes in HFTs' liquidity provisions[J]. Journal of Economics and Finance, 2017,41(4).
11. Blau B M. Religiosity and the Volatility of Stock Prices: A Cross-Country Analysis[J]. Journal of Business Ethics, 2017,144(3).
12. Gao H, Mei D. The correlation structure in the international stock markets during global financial crisis[J]. Physica A: Statistical Mechanics and its Applications, 2019,534.
13. Rubin D N, Bassett D S, Ready R. Uncovering dynamic stock return correlations with multilayer network analysis[J]. Applied Network Science, 2019,4(1).
14. Boako G, Tiwari A K, Ibrahim M, et al. Analysing dynamic dependence between gold and stock returns: Evidence using stochastic and full-range tail dependence copula models[J]. Finance Research Letters, 2019,31.
15. Xiaojun Dong. Macro-economy continues to decline and beware of systemic risk agglomeration-- from the listing of China Cinda and Everbright Bank in Hong Kong [J]. Administrative management reform, 2014(04):33-37.
16. Weichen Xing. Prediction of Shanghai Stock Index based on BP neural network [J]. Enterprise Science and Technology and Development, 2019(12):124-125.
17. Bing Xu, Juan Chen. An empirical study on the tail of Index return Distribution in Shanghai and Shenzhen Stock Market [J]. The practice and understanding of mathematics, 2006(09):49-54.