ISSN 1749-3889 (print), 1749-3897 (online)
International Journal of Nonlinear Science
Vol.32(2021) No.1,pp.24-29

Study on Stock Return Evolution Based on q Correlation Analysis

Wenjin Song, Mei Sun*,
Energy Development and Environmental Protection Strategy Research Center, Jiangsu University, Zhenjiang, Jiangsu, 212013, China

(Received 2 May 2021 , accepted 16 August 2021)

Abstract: By studying the correlation and evolution laws of the Chinese stock market, it is helpful to understand the internal complexity and change laws of the system, and to provide a reliable scientific basis for better national supervision and national investment. This paper divides the data into four fragments based on the big events and actual trend of the SSE 180 stock market, calculates the correlation of the stock data, constructs the network by threshold method, analyzes the correlation and fluctuation causes, and analyzes the correlation and evolution trend of the SSE 180 stock return sequence.

Keywords:q-detrended cross-correlation coefficient; complex network; stock market

1 Introduction

The correlation of the stock market is the mutual influence of the changes among the correlated stocks. This kind of relationship changes over time. The correlation research of stock market has been quite mature, there are many researches on the correlation of stock market. Sun[1] proposed a new method based on the detrend correlation coefficient (DCCA) to improve the portfolio optimization of stocks based on the stock correlation, and to evaluate the risk and return at the same time. Zhang[2] studied the correlation and long-term and short-term trading behavior of Chinese stock indexes under "One Belt And One Road" and stock indexes of other three countries based on the multi-fractal detrended moving average cross-correlation analysis method (MF-X-DMA). Dong[3] constructed the time-varying network of stock data of five BRICS countries from the perspective of complex network, and studied the evolution characteristics of its correlation, so as to guide the investment behavior of investors with different preferences. Zhao[4] analyzed the stock market by using stochastic matrix theory and complex network theory, and proved the superiority of q-detrended cross-correlation coefficient in the calculation of stock correlation based on portfolio theory, which deepened the understanding of the collective behavior of the stock market. By exploring the correlation relationship of the stock market over time, scholars can help to better understand the investment risks and returns of the stock market.

In previous studies, scholars tend to compute correlations based on pearson coefficient, but data in financial markets is often nonlinear, so we need to find correlation computations that are more suitable for processing nonlinear data. Jaroslaw[5] proposed q-detrended cross-correlation coefficient based on detrended cross-correlation analysis (DCCA), detrended fluctuation analysis (DFA) and pearson coefficients, and demonstrate the utility of the q-detrended cross-correlation coefficient in quantifying the correlation strength of nonlinear data and identifying the range of correlation fluctuation amplitude.

Based on the above analysis, this paper discusses the representative SSE market in Chinese stock market based on q-detrended cross-correlation coefficient and complex network theory.

2 Data

As one of the key indexes launched by the Shanghai Stock Exchange, the Shanghai 180 Index has 180 stocks includes different industries. Grasping the rules of its evolution can help us better understand the investment risks and returns of

*Corresponding author. *E-mail address*: sunm@ujs.edu.cn

IJNS.2021.09.15

Figure 1: Trend and segment of SSE 180.

Chinese stock market. This dataset includes the SSE 180 stock, spanning from September 30,2009 to January 1,2021. However, since the data sets of some stocks do not meet the length of the sample, the stock samples used in this paper are all the remaining 112 stocks except the 68 stocks whose length does not meet the requirements. Before calculation, all the 112 stocks that meet the requirements are sorted according to the industry.

Combined with the actual trend of SSE 180 stocks and the national macro policy, the four periods were determined as follows: (1) the volatile market from September 30,2009 to June 30,2014; (2) the bull market from July 1,2014 to December 31,2015; (3) the bear market from January 1,2016 to March 31,2017; (4) the market with rising volatility from April 1,2017 to January 1,2021.

For stock data, the logarithmic return rate is usually calculated to represent the stock data. Therefore, the data of 180 stocks in Shanghai Stock Exchange is used to calculate the logarithmic return

$$r_i(t) = ln(p_i(t+1)) - ln(p_i(t)), \tag{1}$$

where $p_i(t)$ is the daily adjusted closure price of stock i at time t.

3 Method

3.1 *q*-detrended cross-correlation coefficient

The q-detrended cross-correlation coefficient adopted in this paper can deal with the nonlinear and non-stationary real data. When q-detrended cross-correlation coefficient is used, the dominant position of large and small fluctuations in the sequence can be controlled by changing the order q, and the division of sequence length can be changed by changing the scale s, which is helpful to reveal the hidden information in the sequence. The q-detrended cross-correlation coefficient is calculated as follows:

(i) We consider two time series x_i and y_i, $i = 1, \cdots, l$. We integrate these time series and we get two new time series

$$\chi^x(k) = \sum_{i=1}^{k} x_i - \langle x \rangle, k = 1, \cdots, l, \tag{2}$$

$$\chi^y(k) = \sum_{i=1}^{k} y_i - \langle y \rangle, k = 1, \cdots, l. \tag{3}$$

(ii) We divide $\chi^x(k)$ and $\chi^y(k)$ into $2M_s = 2 \times int(l/s)$ non-overlapping parts of length s of two integrated time series according to the order. Then we calculate the local trends for each part $v(v = 1, \cdots, 2M_s)$ by a least-square fit and subtract it from $\chi^x(k)$ and $\chi^y(k)$. So we find the residual signals X and Y, which are equal to the differences between the

integrated signals and the mth-order polynomials $P_{s,v}^{(m)}$ of these signals:

$$X_s(i,v) = \sum_{i=1}^{s} \chi^x[(v-1)s+i] - P_{X,s}^{(m)}(i,v), \tag{4}$$

$$Y_s(i,v) = \sum_{i=1}^{s} \chi^y[(v-1)s+i] - P_{Y,s}^{(m)}(i,v). \tag{5}$$

Then we get the covariance and variance of X and Y in part *v* are defined as:

$$f_{XY;s}^2(v) = \frac{1}{s}\sum_{i=1}^{s} X_s(i,v)Y_s(i,v), \tag{6}$$

$$f_{ZZ;s}^2(v) = \frac{1}{s}\sum_{i=1}^{s} Z_s^2(i,v), \tag{7}$$

where Z represents either X or Y.
(iii) Then we get the fluctuation functions at multifractal order q and detrending scale s:

$$F_{XY}^q(s) = \frac{1}{2M_s}\sum_{v=1}^{2M_s} sgn[f_{XY;s}^2(v)]|f_{XY;s}^2(v)|^{q/2}, \tag{8}$$

$$F_{ZZ}^q(s) = \frac{1}{2M_s}\sum_{v=1}^{2M_s} [f_{ZZ;s}^2(v)]^{q/2}. \tag{9}$$

Then we get the q-detrended cross-correlation coefficient between x_i and y_i is defined as:

$$\rho(q,s) = \frac{F_{XY}^q(s)}{\sqrt{F_{XX}^q(s)F_{YY}^q(s)}}. \tag{10}$$

When $q = 2$, large and small fluctuation parts influence the correlation coefficient $\rho(q,s)$ equally; When $q > 2$, large fluctuation parts influence the correlation coefficient $\rho(q,s)$ more; When $q < 2$, small fluctuation parts influence the correlation coefficient $\rho(q,s)$ more.

3.2 Network construction

We calculate the correlation coefficient between any two stock return sequence $r_i(t)$ and $r_j(t)$ based on q-detrended cross-correlation coefficient. All the correlation coefficients form the cross-correlation matrix $C_1(q,s)$,

$$C_1(q,s) = \begin{pmatrix} \rho_{11} & \cdots & \rho_{1N} \\ \vdots & \ddots & \vdots \\ \rho_{N1} & \cdots & \rho_{NN} \end{pmatrix}, \tag{11}$$

$$C_2(q,s) = |C_1(q,s)|. \tag{12}$$

$C_2(q,s)$ only reflects the strength of the correlation, ignoring the positive and negative correlation. We set the mean of the upper triangle matrix to be the threshold, set 0 for less than the threshold, otherwise set it to 1. By transforming the correlation matrix into a 0-1 matrix, four networks are built according to the adjacent matrix of four periods.

3.3 Centrality method

(1) Degree centrality
Degree centrality is an index to analyze the importance of nodes in the network, which describes the direct influence of nodes. The greater the degree value of a node, the more important it is, which is denoised as DC_i:

$$DC_i = \sum_{j=1,j\neq i}^{n} \delta_{ij}, \tag{13}$$

where δ_{ij} is the number of edges between nodes, n is the total number of nodes.
(2) Closeness centrality
Closeness centrality is measured by the reciprocal of the average of the shortest distances from a given node to all nodes, denoised asCC_i:

$$CC_i = (n-1)/\sum_{j=1,j\neq i}^{n} d_{ij}, \tag{14}$$

where d_{ij} represents the shortest distance from node i to the node j.
(3) Betweenness centrality
Betweenness centrality is defined as the proportion of the number of times that the shortest path of any pair of nodes passes through node i to the number of all shortest paths in the network, denoised as BC_i:

$$BC_i = \sum_{s,t\in V;s,t\neq i} \frac{\sigma(s,t|i)}{\sigma(s,t)}, \tag{15}$$

where s and t are any pair of nodes except node i in set V,$\sigma(s,t|i)$ is the number of shortest paths passing through node i , $\sigma(s,t)$ is the total number of shortest paths between s and t.

4 Results

This section will analyze the SSE 180 market in the case of $q = 2$ and $s = 20$. The stock lists in the table all represent the corresponding stock company codes.

We use three centrality methods to sort the stocks in the network(**Table 1**). It can be observed that the list of top 10 stocks changes significantly in different periods, and they belong to different industries. The percentage of the top 10 stocks reflects the percentage of influential stocks in the industry. The four periods correspond to the period division in section 2. Before sorting, 112 stocks in the data set were classified, and the proportion in each industry was calculated(**Table 2**).

As we can see in table 1, the three centrality top 10 stocks overlapped in period 1, totaling 14 stocks. Some stocks in period 1 are ranked in the top 10 in three centricity methods at the same time (the following is the stock company code) : 600004, 600522, 600362, 600309, 601111, 600068. The three centrality top 10 stocks overlapped in period 2, totaling 11 stocks. Some stocks in period 2 are ranked in the top 10 in three centricity methods at the same time: 600809, 601607, 600801, 600183, 600068, 600004, 600741, 600827, 600406. The three centrality top 10 stocks overlapped in period 3, totaling 14 stocks. Some stocks in period 3 are ranked in the top 10 in three centricity methods at the same time: 601788, 600030, 601006, 601099, 600637, 601872. The three centrality top 10 stocks overlapped in period 4, totaling 13 stocks. Some stocks in period 4 are ranked in the top 10 in three centricity methods at the same time: 600705, 600637, 600030, 600352, 600109, 601788, 600690.

Table 2 is the summary table of ranked stocks in each industry. The first column of the table corresponds to 11 industries respectively. The second column of the table represents the proportion of stocks in each industry before the analysis. By segmenting the time, we can observe the change of the top 10 stocks in each industry over the long four periods.
(1) Changing trends from the perspective of the industry:

The manufacturing industry showed an upward trend in the first and second periods. After comparing with the trend of SSE 180, it was found that stocks also showed an upward trend in the second period, which indicated that leading stocks in the manufacturing industry promoted economic development, and the relationship between them was positively correlated. Stocks fell in the third period, and so did the percentage of the top 10 stocks in manufacturing companies, suggesting that leading stocks in manufacturing have become less important in the overall market during the economic downturn.

The ranking stocks of transportation, storage and postal service industries accounted for the highest proportion in the first period, and showed a slow downward trend over time, indicating that the influence of these industries in the overall market was gradually weakened.

The proportion of information transmission, software and information technology services showed a steady rising trend in the whole four periods, which can also be corresponding to the booming development of information technology services in real life in recent years.

Table 1: The top 10 stocks over four periods

Period 1				Period 2			
Rank	Betweeness	Closeness	Degree	Rank	Betweeness	Closeness	Degree
1st	600004	600362	600362	1st	600809	600183	600809
2nd	600522	600004	600004	2nd	601607	601607	601607
3rd	600362	600522	601111	3rd	600801	600004	600801
4th	600779	601111	600522	4th	600183	600068	600183
5th	601788	600068	600068	5th	600068	601006	600068
6th	600584	600029	600029	6th	600004	600801	600004
7th	600309	601088	600741	7th	600741	600809	600741
8th	600872	600741	601088	8th	601808	600827	601808
9th	601111	600309	600309	9th	600827	600406	600827
10th	600068	600089	601788	10th	600406	600741	600406
Period 3				Period 4			
Rank	Betweeness	Closeness	Degree	Rank	Betweeness	Closeness	Degree
1st	601788	601788	601788	1st	600705	600030	600030
2nd	601899	600637	600637	2nd	600637	600705	600705
3rd	600030	600089	600332	3rd	600030	600352	600352
4th	600208	600030	600089	4th	600352	600837	600837
5th	600859	600332	600030	5th	600016	600176	600176
6th	601006	600010	601872	6th	600109	601788	601788
7th	601099	601872	600010	7th	601899	600109	600109
8th	600637	600109	600109	8th	600038	600637	600637
9th	600196	601006	601006	9th	601788	600690	600690
10th	601872	601099	601099	10th	600690	600362	600362

The ranking stocks of the financial industry occupy a relatively low proportion in the first and second periods, and show a significant rising trend in the third and fourth periods. Therefore, the importance of the financial industry stocks in the overall economy is much higher than the normal market proportion of the financial industry.

It is also possible to find low correlation industries such as: Farming, Iorestry, Fishery; Mining Industry; Electricity, Heating and Supply Industries; Construction Industry; Real Estate; Health and Social work. These industries have a low degree of interaction between stocks and other stocks within the industry, which can also reflect that these industries have a small impact on other industries.

(2) Analyze specific enterprises and reasons:

Manufacturing industry: TBEA Company (600089), taking a place in the market at the peak of the manufacturing industry and the lowest point, which is due to the formation of the battery wafer - photovoltaic module - inverter - complete industrial chain system, it has a relatively complete industrial chain, with silicon wafer, battery, module production capacity. The coordinated and complete development of the industrial chain helps to enhance the ability of enterprises to resist risks. TBEA Company can rank the top 10 stocks in many periods, which is inseparable with its complete industrial chain and anti-risk ability. China Jushi Company (600176) entered the top 10 in the last period. The company is mainly engaged in the production and sales of glass fiber and its products, which are widely used in wind power, electronic circuit, transportation, construction and other fields. The glass fiber industry itself also has certain entry barriers, making it difficult for new entrants to participate in the competition, and enterprises with weak competitiveness are easy to be squeezed out of the market. Large glass fiber enterprises have obvious advantages, and the industry concentration is high. China boulder's capacity share ranks first in the world. Due to its share of more than one-fifth of the global glass fiber material production capacity, as well as the advantages of glass fiber material as an alternative to steel, wood and other traditional materials, under the guidance of the policy, its impact on other industries is huge, so the top 10 stocks are inseparable from these reasons.

Transportation and Postal Services: The top 10 stock changes over time appear the phenomenon of slow decline shows in the earlier years people tend to choose their aircraft, with the development of high-speed rail and other transportation development, to the airport industry has great influence in the importance of the overall market, it also accord with cognitive law, people are now way to travel is gradually tend to be diversified.

Table 2: The top 10 stocks as a percentage of each sector over four periods

sector	total ratio	period1	period2	period3	period4
Farming, Iorestry, Fishery	0.9%	0%	0%	0%	0%
Mining Industry	6.25%	7.1%	9.1%	7.1%	7.7%
Manufacturing Industry	44.6%	57%	36.4%	28.6%	38.5%
Electricity, Heating and Supply Industries	2.7%	0%	0%	0%	0%
Construction Industry	4.5%	7.1%	9.1%	0%	0%
Wholesales and Retail Trade	3.6%	0%	18.2%	7.1%	0%
Transportation and Postal Services	6.25%	21.4%	18.2%	14.3%	0%
Information Technology Service Industry	5.4%	0%	9.1%	7.1%	7.7%
Financial Industry	19.6%	7.1%	0%	28.6%	46.2%
Real Estate	5.4%	0%	0%	7.1%	0%
Health and Social work	0.9%	0%	0%	0%	0%

Financial Industry: In the third and fourth periods, there is a significant upward trend. Concrete reason is probably in recent years, the ”financial committee office of the State Council has issued a series of relevant finance policy, under this background, our country financial industry reaches a new fast track, in the current domestic economic development facing the critical moment of economic structural adjustment and transformation, improve the status of China's financial sector in the economy as a whole.This series of measures has expanded the coverage of China's financial services, so that it can better serve the real economy and drive the overall economy to flourish. Taking CITIC Securities Company (600030) as an example, CITIC Securities Company improved its domestic and international business layout through equity financing and grew into a leading domestic securities company. Compared with other companies in the financial industry, CITIC Securities Company has significant advantages in the accumulation of corporate customer resources and low-cost financing, and the income of the securities industry is closely related to the prosperity of the stock market. Therefore, it is reasonable for CITIC Securities Company to rank in the top 10 of the 180 stock market in Shanghai.

5 Conclusion

This paper divides the stock market into four periods based on the Shanghai Stock Market 180 data trend and major events. The correlation coefficient was calculated based on the q-detrended cross-correlation coefficient.For the stock return series of each period, the q-correlation matrix was calculated and the threshold method was used to build the network. We use three network centrality methods to rank the stocks in the network. It can be observed that in different periods, the top ten stocks have obvious changes, and they come from various industries, which deepens the understanding of sse 180 stock market changes with the period. The phenomenon and reason of the fluctuation of the top ten stocks in the ranking are analyzed.

References

[1] Sun X, Liu Z. Optimal portfolio strategy with cross-correlation matrix composed by DCCA coefficients: Evidence from the Chinese stock market. *Physica A: Statistical Mechanics and its Applications*, 444(2016):667–679.

[2] Zhang X, Zhu Y, Yang L. Multifractal detrended cross-correlations between Chinese stock market and three stock markets in The Belt and Road Initiative. *Physica A: Statistical Mechanics and its Applications*, 503(2018):105–115.

[3] Dong Z, An H, Liu S, et al. Research on the time-varying network structure evolution of the stock indices of the BRICS countries based on fluctuation correlation. *International Review of Economics & Finance*, 69(2020):63–74.

[4] Zhao L,Wei L,Andrea F,et al. The q-dependent detrended cross-correlation analysis of stock market. *Journal of Statistical Mechanics: Theory and Experiment*, (2018):023402.

[5] Jarosław K, Paweł O, Stanisław D. Detrended fluctuation analysis made flexible to detect range of cross-correlated fluctuations. *Physical Review E*, 92(2015):052815.