Research Article

Dynamic Transmission of Correlation between Investor Attention and Stock Price: Evidence from China’s Energy Industry Typical Stocks

Yajie Qi,1,2 Huajiao Li,1,2 Sui Guo,1,2 and Sida Feng1,2

1School of Economics and Management, China University of Geosciences, Beijing, China
2Key Laboratory of Carrying Capacity Assessment for Resource and Environment, Ministry of Land and Resources, Beijing, China

Correspondence should be addressed to Huajiao Li; babyproud@126.com

Received 27 April 2019; Revised 28 September 2019; Accepted 26 October 2019; Published 3 December 2019

Copyright © 2019 Yajie Qi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The relationship between investor attention and stock prices has been a topic of interest in economics. Previous studies have shown that the correlation relationship between the two changes with time. However, there are few studies to explore the time-varying evolution of the relationship, as well as the transmission characteristics under important cycles. Thus, this paper is dedicated to discover the dynamic transmission characteristics of the correlation between investor attention and stock price. We selected the typical stocks of China’s energy industry, PetroChina and Sinopec, as the research objects, as they occupy a large market share and are representative. And the transaction data and attention data are used to build investor attention indicator. In order to reproduce the dynamic transmission process of correlation at different cycles, sliding time window and complex network are applied. The results show that PetroChina and Sinopec stocks have a weakly negative correlation between investor attention and stock price from 2017 to 2018. However, from the perspective of different cycles, the correlation has time-varying characteristics. As the cycle grows, the types of transmission patterns of the five consecutive days of correlation between the two become less, but the transmission intensity between the modes increases and the transition becomes more regular and inclined. In addition, by mining the important transmission modes and main transmission paths under important periods, we find that the series modes of uncorrelated or weakly positive correlation for five consecutive days dominate the transition of modes in the networks. Also, the closed loop formed by these two important modes and related modes is the main transmission path. These findings can reveal the rules of the typical stock market in China’s energy industry and help investors with different investment cycle preferences make sound decisions.

1. Introduction

Among the major global stock markets, energy industry stocks occupy a huge market share. Investors are increasingly concerned about energy stocks because they can play a hedging risk role in the portfolios [1–3], especially typical stocks, such as Exxon Stock [4, 5], Royal Dutch Shell Plc [6], and China’s two major oil giants (PetroChina Company Limited and China Petroleum And Chemical Corporation) [7]. In the long run, the stock prices of energy companies are determined by their intrinsic values, but in the middle- and short-term price analyses, they are determined by the relationship between supply and demand. Meanwhile, as a base financial product, energy stock prices are also influenced by energy prices [8–10], inflation [11], interest and exchange rates [12], government policy uncertainty [13], and so on. Actually, these factors will lead to changes in investors’ behaviors, finally causing stock prices to fluctuate by influencing the relationship between supply and demand.

Investor behavior has become a research topic of interest in finance and is driven by many factors. Grinblatt and Keloharju [14] found that investors with different performances are prone to have different investment strategies and domestic investors tend to be conservative. He and Casey [15] held that investor sentiment had a substantial
forecasting ability for oil service stock returns and crude oil prices. In addition, investor attention could also reflect the psychological changes of investors [16] and encourage investors to take action. Highly volatile stock prices will increase the attention of investors. Rising stock prices will enhance investor confidence, obtain more investor attention, and thus attract more investors to enter the market. Additionally, the continuous decline in stock prices will also obtain investor attention but may lead to the withdrawal of funds. Therefore, there are interdependent and mutually influential relationships between investor attention and stock prices.

However, since investor attention and stock prices fluctuate over time, the correlation between the two is not static [17, 18]. That is, the correlation relationship between investor attention and stock price is dynamically changing over time and presents different correlation characteristics at different time periods. Da et al. [19] focused on Russell 3000 stocks and revealed that increasing investor attention predicted higher stock prices in the next two weeks and a final reversal within the year. Fang et al. [20] collected data from ChiNext market and found that high individual investor attention would lead to high returns on the first day but low returns over the long term. The conclusions of these studies all reflect this phenomenon, but few studies focus on the dynamic evolution process and specific dynamic mechanism of the correlation between investor attention and stock price. Understanding the dynamics of the correlation between the two can reveal more interactive information between investor attention and stock prices and help us to master the market rules.

Previous studies mostly measure the correlation relationship between investor attention and stock price based on linear models such as univariate linear regression equations and vector autoregressive models [21–23]. However, the correlation is dynamic and continuous. Linear models are no longer able to satisfy the dynamic process of characterizing the relationship between the two. Actually, many studies have witnessed the superiority of sliding time windows and complex networks in characterizing nonlinear transmission relationships [7, 24–26]. Hao et al. [27] investigated the sentiment diffusion of public opinions about hot events. Gao et al. [17] proposed an approach to measure the transmission relation of the linear relation between time series. Jiang et al. [28] characterized the causality evolution behavior of multivariate series.

Many studies have shown that the sliding window method has a good effect for extracting time series segments and can provide a solid foundation for investigating the structural characteristics of the whole time series [7]. The sliding window method itself needs to determine the window length and the sliding step size, which, respectively, represent the data length of the calculation correlation and the distance of the window sliding once. By determining the sliding of the length window, the dynamic evolution of investor attention and stock price correlation over a period of time can be achieved, and a complex network is an active method that is composed of agents (nodes) and relations (edges). If we set time windows as nodes and the chronological order between windows as edges, we can construct a network of correlations between investor attention and stock price that are transmitted in the chronological order. Through complex networks, we can portray the consecutive transmission process and explore the dynamic transmission mechanisms of the correlation between investor attention and stock price [29–32].

This paper focuses on discovering the dynamic transmission characteristics and mechanisms between investor attention and stock prices taking China’s energy industry typical stocks as an example. To achieve this goal, first we obtained the transaction records of PetroChina stock and Sinopec stock that are automatically generated by the Xueqiu social networking site and set the transaction volume and the attention behavior as investor attention. Second, we set the window length and sliding step size of the sliding window and calculated the correlation between investor attention and stock price in each window length. Then, correlation transmission networks between investor attention and stock price of PetroChina and Sinopec were constructed based on the complex network method. Finally, the dynamic transmission characteristics of correlation between investor attention and stock price were explored by means of the network’s topological features.

The contribution of this paper consists of two aspects. On the one hand, we focus on the time-varying characteristics of correlation between investor attention and stock price, which is almost unexplored. Many studies have discovered the overall lead-lag relationship and linear correlation between the two, but neglect the nonlinear laws and characteristics of correlation fluctuations from the perspective of dynamics. Although some studies have investigated the difference of correlation under different cycles, few focuses on the transmission of correlation between investor attention and stock price and the evolutionary features of the transmission. But our work used the combination of sliding time window and complex network to portray the dynamic transmission process of correlation between investor attention and stock price. Furthermore, discovering not only the evolutionary characteristics of correlation, but also the specific law of transmission under important cycles.

On the other hand, we construct an investor attention index that is more in line with China’s financial markets. Compared with previous studies [19, 33] using the Google search volume index (SVI), we use first-hand data collected from Xueqiu, which is the most popular professional social networking site of the Chinese stock market [34], to accurately construct an index representing the investor attention for the Chinese energy market. Unlike the SVI, which measures the search volume of one topic for anyone who can search the internet, the Xueqiu website is aimed at investors and automatically generates transaction records and the “follow to stock” behaviors, which can provide a more accurate and exhaustive measurement of investor attention. In addition, in this paper, we choose the top two listed companies of China’s oil and gas industry as typical stocks of the Chinese energy industry. Also, unlike the SVI, the users of Xueqiu are almost Chinese investors, which is more in line with the research objects, PetroChina and Sinopec stocks. Thus, the Xueqiu website is a better choice to
understand the extent to which Chinese stockholders are concerned about PetroChina and Sinopec stocks.

The remainder of this paper is organized as follows. Section 2 describes the data and methods that are used in this paper. Section 3 provides the results and analysis. Section 4 concludes this paper.

2. Data and Method

2.1. Data. We chose two well-known listed companies, PetroChina Co Ltd. and Sinopec Corp., as the typical stocks of China’s energy industry. The reasons are as follows: first, oil and natural gas, as an extremely important one-time energy source, are indispensable material basis for the development of human society and dominate the energy landscape. The document (ScienceNet.cn. Energy transformation ushers in an unprecedented period of opportunity. [EB/OL]. http://news.sciencenet.cn/sbhtmlnews/2018/11/341157.shtml?id=341157) pointed that “considering that energy transformation is a long process, the main energy status of oil and natural gas will not change until 2050. Especially natural gas, as a clean, low-carbon energy in traditional fossil energy, has been effectively utilized in a wide range of fields.” The main businesses of PetroChina and Sinopec listed companies include oil, gas exploration, mining, and sales. Second, PetroChina stock (stock code: 601857) and Sinopec stock (stock code: 600028) are the largest oil and gas producers and sellers in China’s oil and gas industry. The total market capitalization and total share capital of PetroChina stock and Sinopec stock account for 36% and 24%, respectively, in the A-share energy market, ranking first and second. Thus, PetroChina and Sinopec are the stocks that investors are generally concerned about.

To measure the dynamic transmission of correlation between investor attention and stock price of the two companies, four series of data are used in this paper: the investor attention series of PetroChina and Sinopec listed companies include oil, gas exploration, mining, and sales. Second, PetroChina stock (stock code: 601857) and Sinopec stock (stock code: 600028) are the largest oil and gas producers and sellers in China’s oil and gas industry. The total market capitalization and total share capital of PetroChina stock and Sinopec stock account for 36% and 24%, respectively, in the A-share energy market, ranking first and second. Thus, PetroChina and Sinopec are the stocks that investors are generally concerned about.

To measure the dynamic transmission of correlation between investor attention and stock price of the two companies, four series of data are used in this paper: the investor attention series of PetroChina (IA-P), the stock price of PetroChina (SP-P), the investor attention series of Sinopec (IA-S), and the stock price of Sinopec (SP-S).

2.1.1. Investor Attention Index. Since the research object of this paper is China’s energy market, this paper aims to build an investor attention index that is in line with the Chinese security market. Thus, unlike numerous studies, we construct investor attention index through the most popular investment platform in Chinese security market—Xueqiu website other than Google search volume index.

Xueqiu website provides real-time quotes, actual exchanges, and real trades for users and is a popular platform for investors to communicate and share their experiences regarding investment strategies. Because of the low handing fees, the Xueqiu website attracts millions of investors by the end of 2015 [35]. According to research reports from Industrial Securities, one of the top 15% of Chinese securities companies (Securities Association of China. China Securities Industry Association announces the ranking of operating performance indicators of securities companies in 2018. [EB/OL]. https://www.sac.net.cn/hysj/zqgsyjpm/201906/t20190618_139094.html), Xueqiu has a total of 4,194,659 active users and 41,402,009 attention to behavior, posting behavior, and creating portfolio behavior from December 2010 to March 2017. From the coverage of A-shares by monthly and weekly Xueqiu user behaviors, the monthly coverage average reached 99.3% and the weekly reached 93.9% (Industrial Securities. Xueqiu knows about stocks? One of the series: I have an agreement with the attention factor [EB/OL]. https://xueqiu.com/1795556147/85899158). Thus, Xueqiu is undoubtedly one of the best material delivery platforms in many Internet platforms.

As just mentioned, Xueqiu website provides transaction service and every transaction behavior (sell or buy) will be recorded. More importantly, different from other transaction platforms, each investor can follow interested stocks in Xueqiu website. The behaviors of following stocks are the same as those of following interested people in the social networking sites, which means investors pay attention to the stocks following. The behavior of following stocks will also be recorded. Actually, whether buying or selling behavior is a result of attention to stocks. And the act of following stocks is a direct concern and does not overlap with actual trading behaviors. Therefore, we both crawl the transaction record data and following record data of a stock in the Xueqiu website (https://xueqiu.com/S/SH601857; https://xueqiu.com/S/SH600028) from April 1, 2017, to April 1, 2018 (downloaded on April 2, 2018), and use the sum of the two as the investor attention index for the stock. At this point, we obtained two series of daily investor attention data: investor attention of PetroChina and Sinopec stock (IA-P and IA-S, respectively).

2.1.2. Stock Price Data. The stock data were calculated using the daily closing prices that were downloaded on April 2, 2018, from NetEase Finance (http://quotes.money.163.com/stock) for the period from April 1, 2017, to April 1, 2018. To facilitate the calculations and analysis, we used the average of the previous five days’ closing prices instead of the missing daily closing prices. Thus, we obtained two stock price series: SP-P and SP-S.

2.2. Method. This paper focuses on the dynamic transmission of correlation between investor attention and stock price taking typical stocks of the energy industry. As shown in Figure 1, we used the sliding window method twice. For the first time, we calculated the Pearson correlation coefficient between investor attention and stock price in each time window and converted the resulting value into a symbolic representation. The second time, we applied the sequence of symbols obtained in the previous step to the sliding window and treated the continuous symbol sequence of each time window as a mode. At last, we set each mode as a node and the chronological transmission relationship between modes as edges to construct the dynamic correlation transmission network of investor attention and stock price.

2.2.1. Calculating Correlation of Investor Attention Series and Stock Price Series. In order to reflect the dynamics of the correlation between investor attention and stock price, we
initially cut the whole time series into subseries by means of the sliding time window approach. The advantage of a sliding window is that it can contain historical information and transitivity of the time series [7, 28, 36]. First, we need to set the window length of the time window $w_1$ and the step size for each slide $l_1$. Actually, the value of $w_1$ is important because it determines the data length of the calculation of correlation coefficient each time, that is, the diversity of the correlation mode, and $l_1$ is generally set to 1 to include more historical information in the correlation mode. Through the sliding time window, we can discover the correlation characteristics between investor attention and stock price under different time periods.

In this paper, we use the Pearson correlation coefficient to measure the correlation of investor attention and stock price. The Pearson correlation coefficient, also known as the linear correlation coefficient, can be used to measure the similarity of two time series in the trend of change. The principle is that, given two time series, it is judged whether the trends of the values at the same time point and the respective means are consistent. The calculation formula for the Pearson correlation coefficient between investor attention and stock price within a certain sliding window length is as follows:

$$
corr_{w_1,l_1} = \frac{n \sum IA_{w_1,l_1}SP_{w_1,l_1} - \sum IA_{w_1,l_1} \sum SP_{w_1,l_1}}{\sqrt{n \sum IA_{w_1,l_1}^2 - (\sum IA_{w_1,l_1})^2} \sqrt{n \sum SP_{w_1,l_1}^2 - (\sum SP_{w_1,l_1})^2}},
$$

where $i$ represents the $i$th sliding time window, $IA_{w_1,l_1}$ represents the investor attention sequence in the $i$th time window, and $SP_{w_1,l_1}$ is the stock price series in the $i$th time window. The value of $corr_{w_1,l_1}$ is between $-1$ and 1 and can be roughly divided into five categories of correlation degrees according to the coefficient. And we use five symbols to represent the five correlation categories, namely, $P$, $p$, $U$, $n$, and $N$ [37]. The division rules are as follows. From this we convert the correlation sequence between investor attention and stock price into correlation symbol series:

$$
\begin{align*}
P(0.8 \leq corr_{w_1,l_1} < 1), & \quad \text{strongly positive correlation}, \\
p(0.3 \leq corr_{w_1,l_1} < 0.8), & \quad \text{weakly positive correlation}, \\
U(-0.3 \leq corr_{w_1,l_1} < 0.3), & \quad \text{no correlation}, \\
n(-0.8 \leq corr_{w_1,l_1} < -0.3), & \quad \text{weakly negative correlation}, \\
N(-1 \leq corr_{w_1,l_1} < -0.8), & \quad \text{strongly negative correlation}.
\end{align*}
$$

2.2.2. Constructing Correlation Transmission Networks. After symbolization, we again use the sliding window to split the sequence of correlation coefficient symbols into segments of the same length which we call the correlation modes. As shown in Figure 1, $w_2$ and $l_2$ represent the window length of the sliding time window and the step size of each sliding. Every correlation mode is formed by concatenating correlation symbols from time $t$ to time $(t + w_2 - 1)$. Since the stock market has five days of trading days (excluding holidays), we set $w_2$ to 5 days. The correlation modes represent the continuous five-day correlation transmission patterns between investor attention and stock price. At the same time, in order to keep more historical information during the sliding and obtain more correlation patterns, we set $l_2$ to 1 day. Thus, we obtain the correlation modes series.

In order to reflect the evolutionary nature of correlations, we use a complex network theory to construct correlation transmission networks of investor attention and stock price. As shown in formula (3), the networks are constructed with correlation modes as the nodes, the transmission relationship between two chronological modes as the edges, and the transmission frequency as the weights. The direction of the edges is the transmission direction. Finally, two networks are constructed, which are the correlation transmission network between investor attention and stock price for PetroChina stock (CTN-P) and Sinopec stock (CTN-S):

$$
G = (N, E),
$$

where $G$ represents the networks and $N$ is the set of correlation modes. $E$ is the set of edges and can be expressed as the following formula:

$$
E = \begin{bmatrix}
e_{i1} & \cdots & e_{ij} \\
\vdots & \ddots & \vdots \\
e_{i1} & \cdots & e_{ij}
\end{bmatrix}, \quad i, j \in N,
$$

where $e_{ij}$ represents the transmission frequency between mode $i$ and mode $j$. We define the transmission relationship between the two modes as the chronological relationship.
2.2.3. Analyzing Characteristics of Transmission Processes Based on Network Topological Features. Based on the theory of complex networks, we analyze characteristics of transmission processes from three aspects: networks, nodes, and edges. For networks, the density and average weighted degree are vital network feature indicators. From the perspective of nodes, the weighted degree and closeness centrality are commonly used to identify the important nodes in the network. In addition, the weights of edges are indicators that measure the importance of edges.

(1) Closeness of Networks. The density of networks can be used to characterize the degree of closeness between nodes from the perspective of the network. It is defined as the ratio of the number of actual existing edges and the number of edges that can be accommodated in the network. The density is calculated by the following formula:

\[ d(G) = \frac{2L}{N(N-1)} \]

where \( L \) is the number of actual existing edges and \( N \) is the total number of nodes in the network. For CTNs, the densities represent the tightness of transmission between different correlation relationship modes. The higher the density is, the closer the correlation relationship transmission networks will be.

(2) Transmission Strength. In weighted networks, the weighted degree of nodes summarizes the total weights of edges, including in-weighted degree and out-weighted degree [38]. In-weighted degree is the sum of the times that other nodes transmit to this node, and the out-weighted degree is the sum of the number of times this node conducts to other nodes. Due to the characteristics of the network construction in this paper, except the first node and the last node, the in-weighted degree of other nodes is equal to the out-weighted degree. Therefore, we use weighted degree indicators to reflect the transmission capacity and strength of the nodes.

(3) Transmission Tendency. The average weighted degree is the average of all nodes’ weighted degrees and reflects the transmission strength of the whole network. Because the weight between the modes in the networks we have constructed is the number of times the two modes are transmitted, it is clear that the larger the average weighted degree of the network, the larger the average frequency of the two modes transmit in the network which representing the transmission of the modes has a transmission tendency.

(4) Transmission Medium. The betweenness centrality represents the ability to control the resources in the network [39, 40]. Betweenness centrality is the ratio of the shortcuts between mode \( X \) and mode \( Z \) through mode \( Y \) and all shortcuts of mode \( X \) and mode \( Z \). If a mode is on shortcuts for many node pairs, the node has a high betweenness centrality and acts as a transmission medium for connecting other nodes.

3. Results and Analysis

The main purpose of this paper is to measure the transmission characteristics of correlation relationship between investors’ attention and stock price. To achieve this goal, our analysis is engaged from two aspects. The first is to observe the dynamic evolutionary characteristics of correlation between the two under different time periods. The second is to take the important periods as examples to reflect the characteristics of correlation transmission from vital transmission correlation modes and main transmission paths. This can help investors better understand the typical stocks of the Chinese energy market.

3.1. Evolutionary Characteristics of Correlation between Investor Attention and Stock Price under Different Time Periods. We have known that the correlation relationship of investor attention and stock price is not static all the time due to the fluctuation of two time series. And Granger representation theorem states that there is a process of transitioning from short-term fluctuations to long-term equilibrium [41, 42]. Thus, we need to investigate the evolutionary characteristics of investor attention and stock price under different time periods.

Initially, we examined the normality and Pearson correlation of investor attention and stock price series of PetroChina and Sinopec. We used two common and efficient methods to test normality, namely, Shapiro–Wilk test and Lilliefor test [43, 44]. It can be concluded from Table 1 that the four time series obey normal distribution and the Pearson correlation coefficients between investor attention and stock price are \(-0.2064\) and \(-0.0959\), respectively. The results indicate that there is a negative correlation between investor attention and stock price of PetroChina stock, that is, a unit of stock price decline (up) will cause investors to increase (decrease) their attention by 0.2064 units, vice versa. However, for Sinopec stock, there is no obvious correlation between the two. Thus, the relationship between the stock price of typical stocks and investor attention is not exactly the same.

The previous results show a roughly negative correlation between investor attention and stock price across the sample from April 2017 to April 2018. So, is there a negative correlation between the two in each cycle? We cut the entire time series into sequence segments of different periods using a sliding window approach. The longer \( w_1 \), the closer the correlation coefficient obtained is to that of the whole sample, and the shorter \( w_1 \), the bigger the random error. So, we observe changes of transmission processes from one month to three months (nearly 20 days to 60 days). Figure 2 shows the proportion of correlation symbols both for PetroChina and Sinopec. Weakly positive correlation (symbol \( p \)), weakly negative correlation (symbol \( n \)), and no correlation (symbol \( U \)) are the correlation categories between investor attention and stock price, especially the symbol \( U \) and symbol \( p \), and we found an interesting phenomenon. As the length of the sliding window increases, the negative correlation between the two in the sliding window gradually disappears. This is in contrast to the negative correlation between the two variables in the whole sample, indicating
that the correlation relationship varies with the length of the data, and the correlation based on the long-term between investor attention and stock price for investment decisions is almost no value for reference.

By changing the length of the window $w_1$, we get the evolutionary features of network indicators. The average shortest path is used to measure the average distance between each two correlation patterns in the network. The clustering coefficient of a node refers to the connectivity of the adjacent nodes of a node. And the average clustering coefficient of a network is the average of the clustering coefficients of all nodes. From Figures 3 and 4, we found that the trend of change of the number of nodes and the number of edges are exactly the same, while the average shortest path length and the average clustering coefficient are basically irregular. As a whole, with the length of the window increasing, the number of nodes and the number of edges show a downward trend, while the average weighted degree and density of the network tend to increase. This phenomenon shows that as the cycle grows, the relationship between the correlation modes in the network becomes more and more clear and the transmission intensity gradually increases. Based on the above results, we conclude that the correlation relationship transmission between investor attention and stock price is more and more regular and has a preference. That is, the transmission of correlation between investor attention and stock price experiences a process of change from short-term fluctuation to long-term stability, which also confirms the Granger representation theorem.

### 3.2. Transmission Characteristics of Correlation between Investor Attention and Stock Price under Important Time Periods

By observing the evolutionary characteristics of the correlation relationship between investor attention and stock price in different periods, we find some interesting periods in which the network features will have special performance. For example, in the 24-day and 32-day cycles of PetroChina and Sinopec, the five-day correlation patterns between investor attention and stock price has the most types and the most transmission frequency, but the least close transmission relationship in the whole, which can be concluded from the number of nodes, the number of edges, average weighted degree, and density of network. In the 36-day and 47-day cycles, respectively, the opposite is true. Therefore, we select these four important cycles to investigate the transmission characteristics of the correlation between investor attention and stock price.

Figure 5 displays the Pearson correlation coefficient in each sliding window, and Figure 6 is the network visualization under four important periods. The size of the node in the figure represents the weighted degree of the node in the network. The larger the node, the more frequent the transmission of the correlation mode and other modes and the more active in the network. And the thickness of the edge represents the number of transmissions between the two modes, and the thicker the edge, the more the number of transmissions. We find that the types of correlation between investor attention and stock price in the four important periods are basically weakly positive correlation and irrelevant. In theory, there are up to 243 types of modes, but the actual number of modes in the networks is far less than this value. And in combination with the visualization of the network, the series of five consecutive uncorrelated symbols (mode UUUUU) and the series of five consecutive weak positive correlation symbols (mode ppppp) are the main transmission modes of the CTN. This indicates that the transmission of correlation between investor attention and stock prices has a tendency and preference.

In addition, we can observe that these four important networks tend to form closed loops. For example, Figure 6 shows two large closed loops in (a), red and green, respectively, a large closed loop labeled green in (b), red and green “large loops” in (c), and a large green closed loop in (d). The visualization of the networks can intuitively show us the regularity of the correlation transmission between investor attention and stock price. And in order to obtain a comprehensive understanding of the transmission characteristics of correlation between the two, we analyze the transmission process from the main transmission mode and the main transmit path.

#### 3.2.1. Vital Transmission Correlation Modes

Vital transmission correlation modes can reduce randomness in the transmission process and help us to summarize how the correlation between investor attention and stock price is transmitted. To discover the vital transmission correlation modes in the CTN-P and CTN-S, we analyzed the active transmission modes and transmission mediums in the transmission process. Actually, as shown in Figure 6, the visual networks can help us to more intuitively identify the active fluctuation modes, such as correlation modes “UUUUU” and “ppppp” both in the CTN-P and CTN-S. These modes are easy to identify because of their large weighted degrees. In networks, the weighted degrees of nodes represent the transmission strength from one mode to other modes. The larger the weighted degree is, the stronger the transmission ability of the mode is. With respect to the transmission mediums, they are the modes that function as bridges in the transmission process. The betweenness centrality measures the number of shortest paths of all pairs of modes and is an effective indicator that can identify the transmission medium in the network. Modes with large betweenness centralities are the transmission mediums in the transmission process and control the transmission between correlation modes.
Figure 2: Proportion of correlation symbols.

Figure 3: Evolution in network indicators under different window lengths (1).
As shown in Figures 7–10, we use different colors to indicate the transmission and control capabilities of all correlation modes present in the four important periodic networks. Correlation mode “UUUUU” and mode “ppppp” are the correlation modes with the strongest transmission capabilities. And combined with the distribution of investor attention and stock price correlation coefficient, we know that the weakly positive correlation and irrelevant relationship between them not only appear many times, but also have a high probability of appearing for five consecutive days. At the same time, the conversion between these two correlation modes and other modes is also very frequent.

Modes with high betweenness centrality are often on multiple shortest paths, which have strong control over the transmission between correlation modes. For PetroChina stock, the mode “UUUUU” and the mode “pUUUU” are the mediators of the correlation mode transition in the network. For the mode “UUUUU,” it is the only common node between the two “large loops” with the “UUUUU” as the center, as shown in Figure 6(a). Therefore, the shortest path of a mode belonging to two “major loops” largely needs to go through this mode. Similarly, the mode “UUUUU” is also the only connection of “short chain” and “large loop” in Figure 6(b). And the same is true for the mode “pUUUU,” which is at the junction of several closed loops, controlling the transmission between the closed loops.

For CTNs-S, the betweenness centrality of mode “UUUUU” is very low in the transmission network under 47-day cycle, which is different from the other three important cycle networks. The reason for this result is that the mode is no longer a connection point except for the large loop and self-loop. In addition, in the 32-day cycle transmission network, the correlation symbol “n” appears in a few important transmission mediums, for example, modes “Unnn,” “nnnU,” and “nnnn.” Although the weakly...
negative correlation relationship accounts for a small proportion of all correlation relationships between investor attention and stock price, the correlation modes containing weakly negative symbols assume the role of transmission intermediary, indicating that they are still correlation modes that cannot be ignored in the network.

Figure 5: Pearson correlation coefficient between investor attention and stock price in each sliding window under four important periods.

Figure 6: Network visualization under four important periods (notes: (a) represents the network of PetroChina under 24 days; (b) represents the network of PetroChina under 36 days; and (c) and (d) represent the networks of Sinopec under 32 days and 47 days.).
3.2.2. Main Transmission Paths. The transmission path reflects the transit trends of the modes in the transmission process. The main transmission path is the path that is most often transmitted by modes, which represents the most possible transmission direction in the transmission process. Exploring the main transmission paths of CTNs-P and CTNs-S can help us obtain a deeper understanding of the transmissions of correlation between investor attention and stock price.

In networks, the weights of edges represent the transmission strength between two modes. The larger the weight is, the stronger the transmission relationship is. In Figure 6, we use the thickness of the edges to represent the weight between the two correlation modes. And it is not difficult to see that in four important transmission networks; the self-transitions of mode “UUUUU” and mode “ppppp” are the most important transmission paths. However, the main transmission paths vary for different cycle lengths.

The networks with the most transmission patterns and the largest number of edges of PetroChina and Sinopec (24-day cycle transmission network and 32-day cycle network, respectively) often form two distinct transmission paths, namely, the red closed loops and the green closed loops in Figures 6(a) and 6(c). Among them, the red closed loops are the sequential transitions of the mode combinations of the weak negative correlation and the uncorrelated symbols and the green closed loops are the sequential transmissions between the modes of the weak positive correlation and the uncorrelated symbols. And as the cycle grows, the networks (36-day cycle transmission network and 47-day cycle network, respectively) have less correlation modes and transmission relationships, but they exhibit a tighter and more concise transition law. As shown in Figures 6(b) and 6(d), from a 24-day to 36-day periodic transmission network, there are basically only transmissions between modes of combination of weakly positive correlation and uncorrelated symbols. The same is true for transmission networks from 32 days to 47 days.

Furthermore, there is a common transmission path of all four important networks, that is, the green closed loop in
And we also find an interesting phenomenon in the weight ordering of the edges; the top 13 edges exactly form this common transmission path, except for the 32-day cycle transmission network of Sinopec (the 13 edges forming the closed loop are located in the top 16 of ranking in this network). The results indicate that the self-transition of the mode consisting of five consecutive uncorrelated symbols, the self-transition of the mode consisting of five consecutive weak positive correlation symbols, and the cyclic transmission between the two modes are the main transmission paths. Therefore, for the typical stocks of the China’s energy industry, the correlation between investor attention and stock price has long-term continuous irrelevant or weak positive correlation, and the relationship of alternately conversion between irrelevant and weak positive correlation.

4. Discussion and Conclusions

In this paper, our purpose is to contribute to the literature on the research of the characteristics of the typical stocks of the Chinese energy stock market by discovering the transmission characteristics of relationship between energy stock investor attention and energy stock prices. The existing studies on investor attention and stock price mainly focus on the correlation between the two, the lead-lag relationship and the prediction effect of investor attention on the stock price. However, since the investor attention series and the stock price series are constantly changing over time, the correlation between the two should not be static. Although some studies have found the differences of correlation between investor attention and stock price in short term and long term, few literatures focus on the relationship under different cycles and explore the dynamic transmission characteristics of correlation relationships between the two. Therefore, this paper selects two typical Chinese energy industry stocks, PetroChina and Sinopec, as the research object. By establishing the investor attention index based on social network data, the evolutionary characteristics of correlation relationships under different periods and the transmission features of correlation between investor attention and stock price under important cycles are investigated by means of the method of sliding window and
complex network. And we reached the following conclusions.

Overall, the correlation between investor attention and stock price of typical stocks in China’s energy industry is weakly negative correlated. Among them, for PetroChina, a unit of stock price decline (up) will cause investors to increase (fall) 0.2064 units of attention, vice versa; for Sinopec, the weakly negative correlation between the two is not significant from April 2017 to April 2018. In reality, investors are paying more attention to the fall in stock prices than the rise in stock prices. Because when the price of stocks held by investors suddenly begins to fall and is expected to continue to fall, investors tend to increase their attention to stocks and sell short in time. Or investors will pay close attention to stocks that have fallen to the lowest point but have a tendency to pick up. These circumstances may lead to a weakly negative correlation between investor attention and stock price. However, by counting the Pearson correlation coefficients between investor attention and stock price under different periods (different sliding window lengths), we find that the relationship between the two has time-varying characteristics, and weakly positive correlation and irrelevance are the main relationship. The market participants are with different investment preferences. For example, short-cycle investors who tend to earn spread interest through high-frequency operations, companies affected by product cycles, and policy-making departments that have long-term effects on the market. Therefore, investors with different investment preferences can make reasonable decisions based on the correlation between investor attention and stock price in different cycles.

In addition, as the cycle grows, the transmission relationship between investor attention and stock price becomes more and more close. During the process of sliding window length from 20 days to 60 days, the type of correlation modes is gradually reduced, and the transmission between modes is less, but the overall transmission intensity of the network is enhanced. Therefore, as the cycle grows, the transmission law of the correlation between investor attention and stock price becomes more and more obvious. Therefore, the transmission of correlation also has a process
of transitioning from short-term chaos transition to long-term regular transmission.

Furthermore, in order to deeply observe the transmission characteristics of the correlation between the two, we investigated the important transmission correlation modes and main transmission paths of the stocks under two typical cycles. One is the cycle with the least transmission but the most modes (24-day cycle for PetroChina and 32-day cycle for Sinopec); the other is the cycle with the least modes type but the strongest transmission (36-day cycle for PetroChina and 47-day cycle for Sinopec). We find that the active correlation modes and the transmission mediation modes are prominent in the networks, and the main transmission paths of the correlation mode are clear. This shows that the transition of correlation between investor attention and stock price has transition preferences and tendencies.

For important transmission correlation modes, they can reduce the randomness in the transmission process and facilitate the mining of transition laws. The mode “UUUUU” and the mode “ppppp” are active for four important periods, and they are easy to transmit and be transmitted by other modes including the mode itself. And the mode “UUUUU” and mode “pUUUU” are the transmission intermediaries of PetroChina in two cycles. Unlike PetroChina, most modes with strong control capability have negative symbols in the 47-day cycle for Sinopec, such as modes “Unnnn,” “nnnnU,” and “nnnn.” We can conclude that the modes at the junction of multiple chains or loops are modes of strong control in the networks because they are the necessary modes when the modes transmit between different loops or chains.

The transmission networks of the correlation between investor attention and stock price present a closed-loop structure. The 24-day cycle transmission network and the 32-day cycle network have two main transmission loops. One is the cyclically transmitted loop among the modes combined by the uncorrelated relationship and weakly positive correlation; the other is a transmission loop among the modes combined by the uncorrelated and weakly negative correlations. However, the networks under important long-term cycles (36-day cycle and 47-day cycle) only have

![Diagram (a)](image-a)

![Diagram (b)](image-b)

**Figure 10:** Transmission ability and control ability of modes under 47-day cycle for Sinopec.
the latter transmission closed loop. Furthermore, it is interesting to note that the transmission loop (the green closed loop in Figure 6(d)) formed by the transmissions between the modes of five-day uncorrelated symbols and the modes of five-day weakly positive symbols is the only coexisting path for the four cycles and has the strongest transmission in the respective networks. The identified important transmission modes and main transmission paths can reveal the rules of the typical stocks in China’s energy industry and help investors with different investment cycle preferences make reasonable decisions.

In summary, this paper mainly discusses the dynamic transmission characteristics of the correlation between investor attention and stock prices of the typical stocks of China’s energy industry, with a focus on the time-varying characteristics of the correlation between the two. The research results in this paper illustrate the inherent regularity of the stock market. By grasping these rules, investors with different investment preferences can make better decisions. However, our study only focuses on the individual investors and has no concern about the institutional investors who have a scale advantage in influencing the stock prices. In future work, we will add the institutional investor into our studies.

Data Availability

The investor attention data for PetroChina stock and Sinopec stock used to support the findings of this study can be downloaded on the websites https://xueqiu.com/S/SH601857 and https://xueqiu.com/S/SH600028.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research is supported by grants from the National Natural Science Foundation of China (grant nos. 41701211 and 41871202), the Special Project of Geological Mineral Resources and Environmental Investigation (no. DD20190463), the Comprehensive Geological Survey of Beijing and Its Surrounding Areas (no. DD20190655), and the Fundamental Research Funds for the Central Universities (grant no. 2652018254). The authors would like to express their gratitude to Dr. Yang Li and Dr. Yanxin Liu who provided support for our research.

References


