Network correlation between investor's herding behavior and overconfidence behavior*

Mao Zhang(张昴)[1,†] and Yi-Ming Wang(王一鸣)[1,2]

[1] School of Economics, Peking University, Beijing 100871, China

[2] Key Laboratory of Mathematical Economics and Quantitative Finance, Peking University, Beijing 100871, China

(Received 2 December 2019; revised manuscript received 13 January 2020; accepted manuscript online 18 February 2020)

It is generally accepted that herding behavior and overconfidence behavior are unrelated or even mutually exclusive. However, these behaviors can both lead to some similar market anomalies, such as excessive trading volume and volatility in the stock market. Due to the limitation of traditional time series analysis, we try to study whether there exists network relevance between the investor's herding behavior and overconfidence behavior based on the complex network method. Since the investor's herding behavior is based on market trends and overconfidence behavior is based on past performance, we convert the time series data of market trends into a market network and the time series data of the investor's past judgments into an investor network. Then, we update these networks as new information arrives at the market and show the weighted in-degrees of the nodes in the market network and the investor network can represent the herding degree and the confidence degree of the investor, respectively. Using stock transaction data of Microsoft, US S&P 500 stock index, and China Hushen 300 stock index, we update the two networks and find that there exists a high similarity of network topological properties and a significant correlation of node parameter sequences between the market network and the investor network. Finally, we theoretically derive and conclude that the investor's herding degree and confidence degree are highly related to each other when there is a clear market trend.

Keywords: complex network, time series, herding behavior, overconfidence behavior

PACS: 89.65.Gh, 64.60.aq **DOI:** 10.1088/1674-1056/ab7740

1. Introduction

The study of complexity science has been a research hotspot in recent years, and complex networks and time series provide two theoretical ways to describe the complex system. As we gradually have a more in-depth understanding of the complex systems, we start to study time series from the perspective of network science, which not only riches the means of characterizing time series but alsoconduces to understanding the internal evolutionary mechanism of time series.[1]

Zhang *et al.* first pointed out that white noise sequences can be transferred into random networks which exhibit the small world and scale-free characteristics.[2] After that, different approaches have been put forward to studying the fundamental properties of time series in the view of complex networks. According to various definitions of network nodes and edges, current conversion methods can be roughly divided into phase space reconstruction,[3–7] visibility graph,[8–13] and probabilistic transfer method.[14,15] However, these methods only re-examine the time series on the surface, which lacks strong theoretical foundations as their support. Also, they do not retain time information in the network structure. To solve these problems, Zhao *et al.* presented a dynamically equivalent transformation method based on the coarse geometry theory and proved the equivalence property between complex networks and time series under the perspective of topological structure, geometrical characteristic, dynamic characteristic, and correlation dimension.[16]

In addition to the significant theoretical progress, complex networks have also been used to describe and characterize nonlinear systems in the real world, such as stock prices,[17] exchange rates,[18] traffic flows,[19] power,[20] social system,[21] macroeconomics,[22] *etc*. However, no one has studied the decision-making behavior of irrational investors in the stock market by using complex networks.

Herding behavior and overconfidence behavior are two kinds of irrational behaviors that kept attracting attention from academic researchers in behavioral finance in recent years. We can find the literature about the herding behavior in different investment bodies, the market conditions that may lead to herding, and the influence of herding behavior on financial markets. People tend to imitate others because they assume that others have information that can justify their actions. Sias found that institutional investors follow each other into and out of the same securities.[23] Barber *et al.* studied the trading of individual investors by using transaction data and documented the result that individual investors' herd.[24] Gontis *et al.* suggested that human herding is so strong that it persists even when other evolving fluctuations perturb the financial system.[25] There is also some literature studying the conditions that lead to herding behavior. Kremer and

*Project supported by the Youth Program of the National Social Science Foundation of China (Grant No. 18CJY057).

†Corresponding author. E-mail: zhangmao@pku.edu.cn

Nautz showed that herding intensity of institutions depends on stock characteristics, including past returns and volatility.[26] Galariotis *et al.* found that US investors tend to herd during days when important macro data are released.[27] The influence of herding has also attracted much attention. Chiang and Zheng argued that crisis triggers herding activity in the crisis country of origin and then produces a contagion effect, which spreads the crisis to neighboring countries.[28] Bikhchandani and Sharma,[29] Gervais and Odean,[30] Park and Sabourian,[31] Blasco *et al.*[32] all held the view that herding generates more volatile prices.

Overconfidence is a psychological phenomenon that people overestimate the accuracy of their information and their investment analysis. Studies related to overconfidence behavior mainly focus on manifestations of overconfidence, the comparison of overconfidence extent among various groups, and the impact of overconfidence on financial markets. Hoelzl and Rustichini observed that people may be overconfident in many ways.[33] For example, they may perceive themselves more favorably than others, or they may perceive themselves more favorably than they perceive others. Benoî t and Dubra found that it is common for most of people to rank themselves as better than the median.[34] Some studies focused on the performance differences among different investment groups. Barber and Odean proved that men trade more excessively than women.[35] Chuang and Susmel showed that individual investors are more confident than institutional investors.[36] Overconfidence behavior may play a role in excessive market volume and volatilities.[37–39]

Above all, as is well known, the herding investors tend not to value their private information, while the overconfident investors always overemphasize their own information. However, these two seemingly opposite behaviors can both lead to some common market phenomena, such as excessive trading volume and price volatility. Recent studies tended to examine herding behavior or overconfidence behavior alone. Because the traditional time series method mainly focuses on trends, fluctuations, and periodic characteristics of data, in this paper we try to explore whether there exist any network connections between these two kinds of behaviors based on the method of complex networks, which may provide us with a more comprehensive perspective to understand investors' behaviors.

In this paper, we first introduce the improved mapping approach to transform a time series into a network, which is suitable for the transaction data in the stock market. Then, we briefly describe the financial background of stock market equilibrium. Next, we construct a static market network representing the market trends in the eyes of a representative investor and a static investor network reflecting the past performance of the investor. As the stock market equilibrium moves with the arrival of new information, we make rules for the evolution of complex networks. After that, we find that the weighted in-degrees of nodes in the market network and the investor network can represent the herding degree and the confidence degree of the investor, respectively. Using the transaction data of Microsoft, we present the images of the market network and the investor network before and after updating. Finally, we compare the network topological properties and the node parametric properties between the two networks, finding that there do exist correlations between them at the same moment. For further discussion, we theoretically derive the conditions for the convergence between the investor's herding degree and confidence degree, that is, the coefficient of variation of price changes has to be low.

This paper has made the following three main contributions. Firstly, existing researches on investor behavior usually use traditional econometric methods[26,36] or simulations.[30,40] To the best of our knowledge, we are the first to analyze the behaviors of investors based on complex networks and conclude their evolution patterns, which cannot be obtained by the traditional methods. Secondly, the common mapping method is to convert an unchangeable time series into a static network or do comparative static analysis of multiple networks in chronological order. We design the rules that allow networks to evolve dynamically with the update of time series. Finally, existing researches on market anomalies are single and isolated. When it comes to herding behavior and overconfidence behavior, no one has theoretically or empirically studied their possible connections. Some people tend to believe that they are mutually exclusive due to their seemingly contradictory definitions. To fill this void in the literature, we transform time series to networks and find a high similarity for the indicators of network betweenness centralization, network clustering degree and modularity and significant correlations for the node parameter sequences of weighted in-degree, weighted out-degree and weighted degree between the market network which is the basis of the investor's herding behavior and the investor network which is the foundation of the investor's overconfidence behavior. Moreover, we derive and obtain the convergence conditions of the herding degree and confidence degree of the investor.

The remainder of the paper is organized as follows. In Section 2, the construction of the market network and the investor network are introduced from the perspective of a representative investor. In Section 3, the dynamic evolution of networks with the arrival of new information is developed and the economic meaning of weighted in-degree is pointed out. In Section 4, the statistical properties between the two networks at the network level and node level are compared. In Section 5, some conclusions are drawn from the present study.. All the variables are defined in Appendix A, and all the proofs are contained in Appendices B and C.

2. Transformation

2.1. Complexity analysis of time series

The key to studying time series from the perspective of complex network theory is to choose an effective conversion method. As mentioned before, Zhao *et al.* presented a practical transformation approach between time series and complex networks based on the amplitude difference of data points.[16] Let $\{p_t|t=1,2,\ldots,N\}$ be a scalar time series of N observations and ε be a threshold. The adjacency matrix of the transformed network is defined as $A=\{a_{st}|s=1,2,\ldots,N;t=1,2,\ldots,N\}$, where $a_{st}=1$ if $|p_s-p_t|<\varepsilon$, otherwise $a_{st}=0$. That is, there exists a link from node s to node t if the distance between the nodes is less than the threshold. The basis of the equivalence theorem of dual characterization between time series and complex networks is the topological homeomorphism theorem and quasi-isometric theorem. They found that the waveform of the reconstructed time series and the original time series are exactly the same, which means that this conversion method implements an equivalently mathematical description of time series under network representation.

2.2. Brief introduction to stock market equilibrium

Suppose that there is only one financial asset in the stock market. The market consists of I heterogeneous active traders who take long position or short position in this security. The movement from the $(t-1)$-th to the t-th Walrasian equilibrium is driven due to the arrival of new information. Let P_t denote the current market price, and P_{ti} the i-th trader's reservation price which is his belief about the stock price in the future at time t, where $i=1,2,\ldots,I$. Assume that there are no transaction costs and the I traders differ only in their beliefs (Since different traders only differ in their beliefs, we only study the investment behavior of the i-th active trader for convenience), then the inter-trader differences in their beliefs will arise from different expectations about the future and from different needs to transfer risk through the market.

Consequently, all price risk is viewed as falling on terminal wealth. The agent is assumed to choose position Q_{ti} to maximize the expected utility of terminal wealth. The gain attributable to this speculative position and hence its contribution to terminal wealth is

$$\pi=Q_{ti}(P_{ti}-P_t). \tag{1}$$

Then, we obtain the desired position function for the risky security based on the criteria proposed by Tauchen and Pitts[41]

$$Q_{ti}=\alpha_{ti}(P_{ti}-P_t), \tag{2}$$

where α_{ti} is the risk preference coefficient of the i-th trader at time t. A positive value for Q_{ti} represents a desired long position in the contract, while a negative value represents a desired short position.

Equilibrium requires $\sum_{i=1}^{I}Q_{ti}=0$, which indicates that the weighted average of the reservation prices

$$P_t=\sum\nolimits_{i=1}^{I}\alpha_{ti}P_{ti}\Big/\sum\nolimits_{i=1}^{I}\alpha_{ti} \tag{3}$$

clears the market.

Now, we consider the movement from the $(t-1)$-th to the t-th equilibrium. A piece of news arrives at the market and changes the individuals' reservation prices. Since in equilibrium, the total number of buy orders matches the total number of sell orders, we have

$$\sum\nolimits_{i=1}^{I}Q_{ti}-\sum\nolimits_{i=1}^{I}Q_{(t-1)i}=0. \tag{4}$$

The resulting change in the market price is the weighted average of the increments with respect to the traders' reservation prices, which implies

$$\Delta P_t=\sum\nolimits_{i=1}^{I}\alpha_{ti}\Delta P_{ti}\Big/\sum\nolimits_{i=1}^{I}\alpha_{ti}, \tag{5}$$

where $\Delta P_t=P_t-P_{t-1}$ and $\Delta P_{ti}=P_{ti}-P_{(t-1)i}$. The proof is presented in Appendix B. From Eq. (5), we know that the market price change is consistent with the reservation price changes of most of position holders in the market.

We have introduced the decision-making behavior of rational investors above, where market prices and investor's reservation prices are both exogenous. However, we know that investors cannot be rational all the time in real life, so their investment decisions may be affected by their sentiment. For example, if the market trend is particularly apparent, the trader will be sure of his belief, and hence increasing his positions. That is what we see as the "buying the winners" strategy. Moreover, a trader will become more optimistic when he tends to judge correctly for the market price changes recently or after a high level of return,[42] which will lead to excessive transactions.

We can see that investors' seemingly opposite behaviors, such as following others and believing in themselves, can both lead their desired positions to increase. This raises the question whether an investor can behave both herding and confidence at the same time. If so, does any connection exist between them? Moreover, what is their combined effect on the investor's position? To answer these questions, we construct a market network that represents the market trend and an investor network that represents the past judgments from the perspective of a representative investor in the following subsection.

2.3. Construction of market network and investor network

Following the work of Zhao *et al.*,[16] we will build complex networks based on the time series of price changes (Due to the relatively large fluctuations in the time series of stock prices, which will interfere with the study of statistical properties of the data, used in this paper are the differenced stock prices, that is, the time series of price changes) and past judgments for the representative i-th investor.

Firstly, in the market network, we define node t as the market price change at time t and the edge from node $(t-m)$ to node t as the influence of the market price change at time $(t-m)$ on the market price change at time t. Considering that the current market price change cannot have an influence on the past market price changes, and the memory length of the trader is limited (assuming the i-th investor's memory length is M) (Although each one observes the same market prices, their herding behavior may be different due to the various memory lengths), we give definitions below.

Definition 1 In the market network,

(i) $a'_{(t-m)t} = 1$ if $|\Delta P_{t-m} - \Delta P_t| < \varepsilon$ and $1 \leqslant m \leqslant M$; otherwise, $a'_{(t-m)t} = 0$, $t = 1, 2, \ldots, N$ (Since there are only N nodes in the network, the in-degree of the first M nodes (*i.e.*, $t = 1, \ldots, M$) is t-1, the same as that in Definition 2).

(ii) The size of node t is the market price change at time t, that is, ΔP_t.

(iii) The weight of the edge from node $(t-m)$ to node t is the product of the transition probability φ_{tm} and the size of node ΔP_{t-m}, that is, $\varphi_{tm}\Delta P_{t-m}$.

Similarly, we can construct an investor network based on the past judgments of the i-th investor.

Every time when a piece of new information arrives at the market, each trader will update his reservation price and then obtain the market clearing price. If the directions of the two price changes are the same, the trader will make the right investment decision because his judgment is consistent with the judgment of the majority. Therefore, for the investor network, we employ node t to represent the product of the market price change and the trader's reservation price change at time t and the edge from node $(t-m)$ to node t to represent the effect of the trader's judgment at time $(t-m)$ on the judgment at time t. Being the same as the market network, the trader's current judgment cannot influence the past judgment, and the memory lengths of different traders are different (assuming that the i-th investor's memory length is M), we give another definition below.

Definition 2 In the investor network,

(I) $a''_{(t-m)t} = 1$ if $\left|\Delta P_{t-m}\Delta P_{(t-m)i} - \Delta P_t \Delta P_{ti}\right| < \varepsilon$ and $1 \leqslant m \leqslant M$; otherwise, $a''_{(t-m)t} = 0$. $t = 1, 2, \ldots, N$.

(II) The size of node t is the product of the market price change and the trader's reservation price change at time t, that is, $\Delta P_t \Delta P_{ti}$.

(III) The weight of the edge from node $(t-m)$ to node t is the product of the transition probability ϕ_{tm} and the size of node $\Delta P_{t-m}\Delta P_{(t-m)i}$, that is, $\phi_{tm}\Delta P_{t-m}\Delta P_{(t-m)i}$.

To have an intuitive understanding of the mapping from time series to complex networks, we present the network structures of the stock price series (P_0), the price change series and the investor judgment series together based on the sample of Microsoft's stock prices during the first 20 trading days in July 2009. Specifically, we transfer the price change series into the market network (M_0) according to Definition 1, which is shown as the right sub-network in Fig. 1. Meanwhile, we map the investor judgment series into the investor network (I_0) by Definition 2 in the left sub-network of Fig. 1. The 20 vertical nodes (P_0) in the center of Fig. 1 correspond to the 20 observations of the stock price series, which are connected to the nodes at the same time point in the market sub-network and investor sub-network, respectively. That is to say, there are 20 nodes in each sub-network. It is supposed that the investor's memory length is 5 trading days and the threshold is big enough so that each node has 5 directed edges from its previous 5 nodes. The nodes in gray represent the stock prices. The other nodes are in green if their sizes are above 0, otherwise in red. Since we define the edges as the influence of past nodes on future nodes, the color of the edge is the same as that of its leaving node.

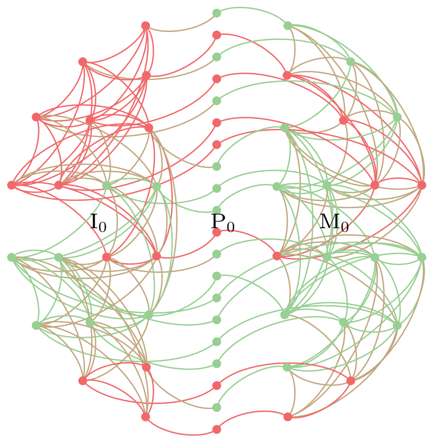

Fig. 1. Network mapping of time series of stock prices, price changes, and investor's past judgments.

3. Dynamic evolution of networks

In this section, we will first describe how the market network and the investor network evolve with the change of recent market fluctuation and the trader's past performance, respectively. Also, we will intuitively present the evolution patterns of these networks to have a more in-depth understanding of the changing process.

3.1. Dynamic evolution of networks

The networks constantly evolve as the stock market moves from one equilibrium to the next one. Due to the similar structure between the market network and the investor network, we take the market network for example to explain the dynamic evolution of the network structure in detail.

It is supposed that there are always N nodes in chronological order of a network. When a piece of news arrives at the stock market, we obtain a new market clearing price, thus a new market price change. At this time, we add a new node to this network and meanwhile delete the first node of the network, so there are still N nodes. Since the memory of investors is limited, we assume that only the previous M nodes can affect the t-th node in the eyes of investors. That is, there will be M directed edges from the previous M nodes to the new t-th node. Then, how is the influence of former nodes on the t-th node judged?

Assuming that every time when a node k is added to the network, the investor will compare the size of the k-th node ΔP_k with the sizes of the previous M nodes ΔP_{k-1}, ΔP_{k-2}, …, ΔP_{k-M}. If the difference between ΔP_k and ΔP_{k-m} is smallest, then $x_k = m$, where x is defined as a discrete variable taken on values 1, 2, …, M. This means that the $(k-m)$-th node has the most significant influence on the k-th node. In other words, the status of node k is more likely to be transformed from the status of node $(k-m)$ than the other previous M nodes. Let k traverse from $(t-N)$ to $(t-1)$, then $\sum_{k=t-N}^{t-1}\delta(x_k = m)$ will represent the number of times for the status of node k to be transformed from the status of node $(k-m)$ for the previous N nodes of node t. Therefore, we can use the frequency of $x = m$ to represent the transition probability from the $(t-m)$-th node to the t-th node, which can be expressed as

$$\varphi_{tm} = \frac{\sum_{k=t-N}^{t-1}\delta(x_k = m)}{N}, \tag{6}$$

where $\delta(x_k = m) = 1$ if $x_k = m$, otherwise $\delta(x_k = m) = 0$.

Meanwhile, the i-th investor compares the size of the k-th node $\Delta P_k \Delta P_{ki}$ with the sizes of the previous M nodes $\Delta P_{k-1}\Delta P_{(k-1)i}$, $\Delta P_{k-2}\Delta P_{(k-2)i}$, …, $\Delta P_{k-M}\Delta P_{(k-M)i}$ in the investor network. If the difference between $\Delta P_k \Delta P_{ki}$ and $\Delta P_{k-m}\Delta P_{(k-m)i}$ is smallest, then $y_k = m$, where y is defined as a discrete variable taken on values 1, 2, …, M. Therefore, the transition probability from the the $(t-m)$-th node to the t-th node in the investor network can be expressed as

$$\phi_{tm} = \frac{\sum_{k=t-N}^{t-1}\delta(y_k = m)}{N}, \tag{7}$$

where $\delta(y_k = m) = 1$ if $y_k = m$, otherwise $\delta(y_k = m) = 0$.

3.2. Investor's herding extent

The weighted degree is a critical statistical index in the complex network. The higher the weighted in-degree or out-degree of the node, the more critical the node is in the complex network.[43] We explore the economic implications of weighted in-degree of the market network in this subsection and weighted in-degree of the investor network in the next subsection.

Based on the dynamic evolution process described above, the weighted in-degree of node t in the market network can be expressed as the sum of the influence of nodes which have direct links to node t on node t, that is,

$$\text{trend}_t = \varphi_{t1}\Delta P_{t-1} + \cdots + \varphi_{tM}\Delta P_{t-M} = \sum_{m=1}^{M}\varphi_{tm}\Delta P_{t-m}. \tag{8}$$

We can see from Eq. (8) that the magnitude of the weighted in-degree is larger when there is a considerable rise or fall with duration increasing. However, if the prices fluctuate, the magnitude of the weighted in-degree will be smaller. Meanwhile, the herding extent of investors has a similar pattern. When the market is in a booming stage, the herding investors will continue to be bullish because individuals holding most positions are bulls. However, in a recession, herding investors will continue to be bearish because individuals holding most positions are bears.[44] In a word, the herding extent of investors will accelerate if there is an evident market trend.[45]

Considering the consistency between the size of the weighted in-degree and the investor's herding degree with the change of market trend, we can use the absolute value of the weighted in-degree to indicate the investor's herding extent, that is,

$$h_t = |\varphi_{t1}\Delta P_{t-1} + \cdots + \varphi_{tM}\Delta P_{t-M}| = \left|\sum_{m=1}^{M}\varphi_{tm}\Delta P_{t-m}\right|. \tag{9}$$

At the beginning of each period t, the i-th trader forms his herding degree based on the past market trend. The higher the absolute value of the weighted in-degree, the stronger the investor's herding degree will be.

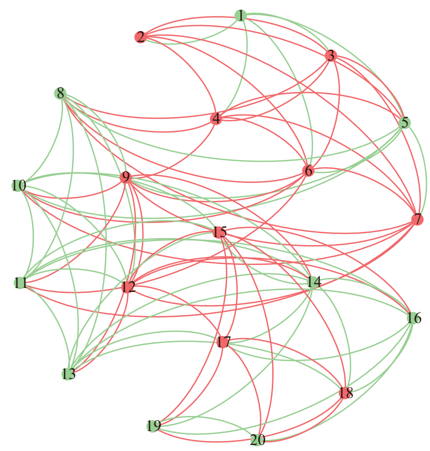

Fig. 2. Market network before being updated.

To clearly show the changes of the market network before and after being updated, we construct the market network by using the Microsoft's transaction data spanning from July 1, 2009 to June 30, 2019 which consists of 2517 observations. Suppose that there are always 20 nodes in a static network, *i.e.*, $N = 20$, and the memory length of a typical investor is one week, that is, $M = 5$, then we will update the network by using the method described above. Figure 2 shows the market network before being updated, which is the same as the right half part of Fig. 1, while figure 3 shows the market network after being updated for 10 years. Being the same as Fig. 1, the node is in green if its size is above 0, otherwise in red. The color depth is correlated to the magnitude of the node.

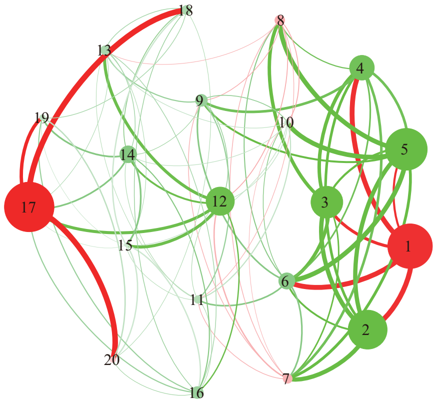

Fig. 3. Updated market network.

3.3. Investor's confidence extent

Similarly, we can express the weighted in-degree of node t in the investor network as

$$\begin{aligned} c_t &= \phi_{t1}\Delta P_{t-1}\Delta P_{(t-1)i} + \cdots + \phi_{tM}\Delta P_{t-M}\Delta P_{(t-M)i} \\ &= \sum_{m=1}^{M} \phi_{tm}\Delta P_{t-m}\Delta P_{(t-m)i}. \end{aligned} \tag{10}$$

If the i-th investor judges correctly at time $(t-m)$, then $\Delta p_{t-m}\Delta p_{(t-m)i} > 0$, otherwise $\Delta p_{t-m}\Delta p_{(t-m)i} \leqslant 0$. The weighted in-degree will increase with the right extent of judgment increasing and decrease with the wrong extent of judgment increasing. Meanwhile, the confidence extent of investors also has a similar change pattern to the weighted in-degree in the investor network. Odean argued that investors' degree of overconfidence can be measured by their early excess net return.[37] Many researchers believed that in the stock market, most of investors tended to take too much credit for the excess profits they had obtained, which would produce more trading and volatility.[30,36,46] In a word, the key to overconfidence is that investors have made the correct judgments many times, and they believe that it is their own abilities that help them make the right judgments. Therefore, we can use the value of the weighted in-degree in the investor network to indicate the confidence extent of investors. At the beginning of each period t, the i-th trader forms his confidence extent based on past judgments. The higher the weighted in-degree, the stronger the investor's confidence degree will be.

We still use Microsoft's stock prices to calculate market price changes. However, the data of the traders' reservation prices are not available, so we simulate the i-th trader's reservation prices as follows:

$$\Delta P_{ti} = \theta_i \Delta P_t + (1-\theta_i)\sum_{m=1}^{M} \phi_{tm}\Delta P_{t-m}, \tag{11}$$

where θ_i takes the value of 0.5 (We get similar results when θ_i takes the value of 0.2 and 0.8). Still supposing $N = 20$ and $M = 5$, we present the investor network as follows. Being the same as the left half part of Fig. 1, figure 4 shows the investor network before being updated, and Fig. 5 illustrates the network after 10-year update.

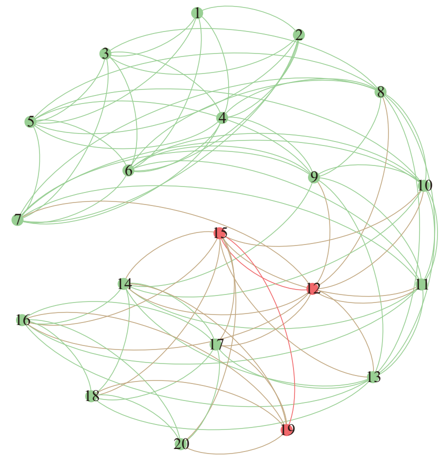

Fig. 4. Investor network before being updated.

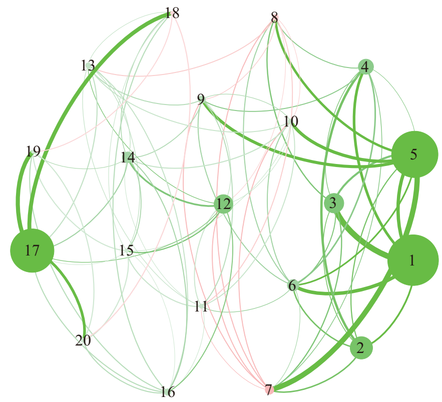

Fig. 5. Updated investor network.

4. Comparison between market network and investor network

To explore the potential correlations between the market network and the investor network, we will analyze the properties of the two networks at the network level and the node level in this section.

4.1. Topological index at network level

The common practice of comparing two networks at network level is to compare some main topological indicators, including the network betweenness centralization, network clustering coefficient, and modularity degree.

To verify the generality of the experimental results, apart from employing Microsoft's transaction data spanning from July 1, 2009 to June 30, 2019 in the previous article, we present the corresponding results of Microsoft's transaction data from July 1, 2009, respectively, to September 30, 2015, December 31, 2016, and March 31, 2018. We choose different months as the cut-off time so as to exclude the interference of quarterly effects in the stock market. Also, we show the network topological indexes and the correlation coefficients between the node parameter sequences based on the data of the US S&P 500 stock index and the China Hushen 300 stock index. We use M_0 and I_0 to denote the market network and the investor network before being updated, M_1 and I_1 to represent the two networks after being updated. The results are shown in Table 1.

Table 1. Network topological indexes.

	Network betweenness centralization	Network clustering coefficient	Modularity
Panel A: Microsoft transaction data spanning from July 1, 2009 to June 30, 2019			
M_0	0.023	0.382	0.510
I_0	0.023	0.382	0.523
M_1	0.023	0.382	0.413
I_1	0.023	0.382	0.371
Panel B: Microsoft transaction data spanning from July 1, 2009 to September 30, 2015			
M_1	0.023	0.382	0.362
I_1	0.023	0.382	0.351
Panel C: Microsoft transaction data spanning from July 1, 2009 to December 31, 2016			
M_1	0.023	0.378	0.329
I_1	0.023	0.378	0.404
Panel D: Microsoft transaction data spanning from July 1, 2009 to March 31, 2018			
M_1	0.023	0.382	0.416
I_1	0.023	0.382	0.336
Panel E: Hushen 300 transaction data spanning from July 1, 2009 to June 30, 2019			
M_0	0.023	0.382	0.448
I_0	0.023	0.382	0.484
M_1	0.023	0.382	0.418
I_1	0.023	0.382	0.406
Panel F: S&P 500 transaction data spanning from July 1, 2009 to June 30, 2019			
M_0	0.023	0.382	0.435
I_0	0.023	0.382	0.491
M_1	0.023	0.382	0.434
I_1	0.023	0.382	0.483

Network betweenness centralization is the average of the ratios of the number of paths that pass through each node to the total number of shortest paths, which can be used to measure the average bridge intermediation of network nodes. Table 1 shows that the values of the network betweenness centralization remain constant after being updated for all the three types of stocks. Besides, the values of the indicator for the market network are the same as those for the investor network at a single point, indicating that the average financial information transfer effects of the two networks are similar to each other.

Network clustering coefficient is a measure of the degree to which the nodes in a graph tend to cluster together. We can see from Table 1 that the value of the network clustering coefficient declines only for the Microsoft data spanning from July 1, 2009 to December 31, 2016. For the three kinds of stocks, the aggregation characteristics in the market network and the investor network are still similar, whether they are updated or not.

Modularity is used to measure the strength of the division of a network into modules that are also called groups, clusters, or communities. Networks with high modularity have dense connections between the nodes within modules but sparse connections between nodes in different modules. From Table 1 it may follow that the values of the modularity of both networks all decline, whether for different updated intervals or classes of stocks. At the same point, there is no much difference in the value between the market network and the investor network.

4.2. Correlation analysis at node level

In this part, we focus on the parameters related to the nodes such as the weighted in-degree, weighted out-degree, and weighted degree (The weighted in-degree in the market network is taken as its absolute value in the following calculation due to its significant economic meaning discussed in Subsection 3.2). The detailed interpretations of each parameter used here are presented as follows.

Weighted in-degree is the sum of the weights of the edges entering a node.

Weighted out-degree is the sum of the weights of the edges leaving a vertex.

Weighted degree is the sum of weighted in-degree and weighted out-degree.

Using the updated networks, we first compute the corresponding parameters of each node. Then, we obtain two parameters series, and thus we can calculate their correlation coefficients. To confirm that the correlation coefficient is only significant between the two networks at the same time, we also calculate the correlation coefficients between the updated market network (M_1) and the investor network before being updated (I_0), and the updated investor network (I_1) and the market network before being updated (M_0) for each node parameter. The results are shown in Table 2.

Table 2. Correlation coefficients for node parameter sequences.

Correlation coefficients	Weighted in-degree	Weighted out-degree	Weighted degree
Panel A: Microsoft transaction data spanning from July 1, 2009 to June 30, 2019			
$M_1 \sim I_1$	0.662	0.919	0.837
$M_1 \sim I_0$	−0.054	0.243	−0.041
$I_1 \sim M_0$	−0.100	0.040	−0.103
Panel B: Microsoft transaction data spanning from July 1, 2009 to September 30, 2015			
$M_1 \sim I_1$	0.602	0.942	0.927
$M_1 \sim I_0$	−0.233	−0.086	−0.359
$I_1 \sim M_0$	−0.096	0.165	−0.105
Panel C: Microsoft transaction data spanning from July 1, 2009 to December 31, 2016			
$M_1 \sim I_1$	0.733	0.950	0.957
$M_1 \sim I_0$	−0.355	−0.236	−0.323
$I_1 \sim M_0$	0.122	−0.360	−0.150
Panel D: Microsoft transaction data spanning from July 1, 2009 to March 31, 2018			
$M_1 \sim I_1$	0.402	0.746	0.503
$M_1 \sim I_0$	0.556	0.110	0.282
$I_1 \sim M_0$	0.226	−0.066	−0.150
Panel E: Hushen 300 transaction data spanning from July 1, 2009 to June 30, 2019			
$M_1 \sim I_1$	0.828	0.958	0.926
$M_1 \sim I_0$	0.182	0.155	0.101
$I_1 \sim M_0$	0.077	0.221	0.250
Panel F: S&P 500 transaction data spanning from July 1, 2009 to June 30, 2019			
$M_1 \sim I_1$	0.754	0.922	0.852
$M_1 \sim I_0$	0.118	−0.348	−0.226
$I_1 \sim M_0$	0.078	−0.297	−0.170

From Table 2, we can see that the correlation coefficient of the weighted in-degree series, weighted out-degree series, and weighted degree series between the updated market network and investor network are all around 0.6 in these cases, which indicates a high correlation of the node parameters between the two networks. Correspondingly, we cannot find distinct patterns between the updated market network and the investor network before being updated, nor the updated investor network and the market network before being updated, which further illustrates that the correlation relationship only exists between the market network and the investor network at the same time point.

As demonstrated in Section 3, the weighted in-degrees of the nodes in the market network and the investor network can represent the investor's herding degree and confidence degree. Due to their critical economic meanings, we focus on the correlation of the weighted in-degree series of the two networks and theoretically derive the conditions for their convergence, which helps us analyze the internal mechanisms behind our conclusions. Then, we obtain

$$\lim_{CV \to 0} \rho = 1, \tag{12}$$

$$\lim_{CV \to \infty} \rho = 0, \tag{13}$$

where CV is the coefficient of variation of price change. Their proof is presented in Appendix C.

We finish this section by analyzing the internal mechanisms behind our conclusions. We suggest that these two kinds of behaviors that seem to be mutually contradictory are in fact consistent when their variations with the mean of the price changing are low. In particular, when their variations with the market price changing is small, which means that the market price changes are stationary, the market will show a clear trend. According to Eq. (9), the investor's herding degree will go up because it is easy for him to seize the market trend. Equation (5) shows that the direction of the market price change is the same as that of the reservation price change for most of position holders in the market. Therefore, seizing the market trend means that the investor knows what the majority will do. At this time, it is easy for the investor to make the same decisions as most of position holders do, making him more likely to make the right judgments. Based on Eq. (10), the investor's confidence degree will increase with the right extent of judgments increasing. Finally, the overconfidence goes up with the accumulation of continuously growing confidence levels. For example, if the price of some stock rises steadily over a period of time, that is, there is a clear upward trend, the investors are more likely to enter into the market or increase their positions, showing herding effects. The herding behavior of investors will push the price to go up further, and investors will obtain positive return, which will enhance their confidence degrees.

However, when the coefficient of variation is large, investors do not have the chance to follow the crowd. At this time, whether investors have made their right judgments is entirely accidental, so a trader's herding degree is almost not correlated with his confidence degree. For example, if the price of a stock goes up and down like a seesaw for some time, then investors will not have any trends to follow. At this time, some investors will increase their positions, and others will reduce their positions. The irregular behaviors of investors will have an uncertain influence on the direction of price change, so investors cannot continuously obtain positive returns, and their confidence degrees cannot accumulate.

5. Conclusions

Since it is difficult to study the relationship between investors' herding behavior and overconfidence behavior by using traditional approaches, in this paper we propose a framework to study the mechanism between herding behavior and overconfidence behavior from the perspective of the complex network. In this framework, we transform a time series into a complex network based on the amplitude difference of data points. The nodes represent the market price change in the market network and the investor's past performance in the investor network, respectively. The stock market moves from one equilibrium to another with time, so we characterize the transition probability with the transition frequency of node status in the past, and then present the dynamic evolution processes of the two networks. Since the investor's herding degree depends on the market trend and the investor's confidence degree depends on his past performance, we use the weighted in-degrees of the nodes in the market network and those in the investor network to represent the investor's herding degree and confidence degree, respectively. To reflect the change of the investor's herding degree and confidence degree with the arrival of new information, we design the rules that the network can evolve dynamically as new information reaches the market. Finally, we find that the market network and the investor network at the same point are not only highly similar in the topological nature, but also their node parameter sequences have a certain correlation. Due to the economic significance of the weighted in-degree, we focus on it and find that an investor's herding degree is highly related to his confidence degree if there exists a clear market trend.

Our conclusions provide the support for explaining the price-volume relation. Karpoff reviewed previous empirical and theoretical research and proposed a model that further verifies the positive correlation between volume and the absolute value of the price change or the price change *per se* in stock markets.[47] As we introduced in Section 1, there exists evidence that overconfidence behavior has an influence on the market trading volume and volatility, and herding behavior can affect market volatility. In this paper, we prove the consistency between confidence and herding under certain conditions, which explains the same increase or decrease in trading volume and volatility to some extent. Therefore, the interaction between the two phenomena may also be an intrinsic mechanism of the price-volume relation.

Our findings also have some policy implications. Firstly, the investment decisions made by investors are inevitably affected by the global market environment, such as market price changing over a period of time. When the market conditions are good, investors should realize that the seemingly favorable market environment may be a kind of "illusion", that is, the result of malicious manipulation by investors with dominant capital. Otherwise, uninformed investors would believe that they have seized the market trend, thereby accumulating the excessive confidence in a short time. It may cause huge losses when market manipulators decide to harvest. Secondly, regulators should establish a perfect investor protection mechanism. Both professional investment institutions and individual investors should receive educations on securities law and risk warning. It is also necessary to popularize the investment techniques and investment strategies and encourage investors to learn psychological knowledge. In other words, regulators should guide investors in avoiding short-term overconfidence and establishing a sense of rational investment.

Appendix A

Table 1. Variable definitions.

Variables	Definitions
p_t	The t-th observation of a random scalar time series.
ε	The threshold used when transforming a time series into a complex network, and its value may be greater than or equal 0.
a_{st}	The element of the adjacency matrix, which equals 1 if there exists an edge from node s to node t, otherwise equals 0.
I	There are I heterogeneous active traders in the market in total.
i	We use the subscript i to represent the i-th active investor who takes long or short positions in the market, where $i = 1, 2, \dots, I$.
P_t	The market price at time t.
ΔP_t	The market price change at time t.
P_{ti}	The i-th trader's belief in the stock price in the future at time t.
ΔP_{ti}	The change of the i-th trader's belief in the stock price in the future at time t.
Q_{ti}	The desired position of the i-th trader at time t.
π	The contribution of the i-th trader's speculative position to his terminal wealth at time t.
α_{ti}	The risk preference coefficient of the i-th trader at time t.
$a'_{(t-m)t}$	The element of the adjacency matrix of the market network, which equals 1 if there exists an edge from node $(t-m)$ to node t, otherwise equals 0.
φ_{tm}	The transition probability from node $(t-m)$ to node t of the market network.
$a''_{(t-m)t}$	The element of the adjacency matrix of the investor network, which equals 1 if there exists an edge from node $(t-m)$ to node t, otherwise equals 0.
ϕ_{tm}	The transition probability from node $(t-m)$ to node t of the investor network.
P_0	The network structure of the stock price series before being updated.
M_0	The market network before being updated.
I_0	The investor network before being updated.
M_1	The market network after being updated.
I_1	The investor network after being updated.
$\delta(x_k = m)$	An indicator variable which equals 1 if $x_k = m$, otherwise equals 0.
$\delta(y_k = m)$	An indicator variable which equals 1 if $y_k = m$, otherwise equals 0.
trend_t	The weighted in-degree of node t in the market network.
h_t	The investor's herding extent.
c_t	The weighted in-degree of node t in the investor network, that is, the investor's confidence extent.
θ_i	The weight parameter when simulating reservation prices of the i-th trader.
CV	The coefficient of variation of price changes.
ρ	The correlation coefficient between the weighted in-degree series of the market network and the investor network.

Appendix B: Proof of Eq. (5)

The requirement for the i-th equilibrium, in which the total number of buy orders must match the total number of sell orders, is given by Eq. (4).

Inserting Eq. (2) into Eq. (4), we obtain

$$\sum_{i=1}^{I} \alpha_{ti}(P_{ti} - P_t) = \sum_{i=1}^{I} \alpha_{(t-1)i}(P_{(t-1)i} - P_{t-1}).$$

It can be rewritten as follows:

$$\sum_{i=1}^{I} \alpha_{ti} P_{ti} - \sum_{i=1}^{I} \alpha_{(t-1)i} P_{(t-1)i} = \sum_{i=1}^{I} \alpha_{ti} P_t - \sum_{i=1}^{I} \alpha_{(t-1)i} P_{t-1}.$$

We subtract and add $\sum_{i=1}^{I} \alpha_{ti} P_{(t-1)i}$ on the left-hand side of the equation and $\sum_{i=1}^{I} \alpha_{ti} P_{t-1}$ on the right-hand side of the equation, and obtain

$$\begin{aligned} &\sum_{i=1}^{I} \alpha_{ti} P_{ti} - \sum_{i=1}^{I} \alpha_{ti} P_{(t-1)i} \\ &+\sum_{i=1}^{I} \alpha_{ti} P_{(t-1)i} - \sum_{i=1}^{I} \alpha_{(t-1)i} P_{(t-1)i} \\ =& \sum_{i=1}^{I} \alpha_{ti} P_t - \sum_{i=1}^{I} \alpha_{ti} P_{t-1} \\ &+ \sum_{i=1}^{I} \alpha_{ti} P_{t-1} - \sum_{i=1}^{I} \alpha_{(t-1)i} P_{t-1}. \end{aligned}$$

Define $\Delta P_t = P_t - P_{t-1}$,$\Delta P_{ti} = P_{ti} - P_{(t-1)i}$, and $\Delta\alpha_{ti} = \alpha_{ti} - \alpha_{(t-1)i}$, then we will have

$$\sum_{i=1}^{I} \alpha_{ti} \Delta P_{ti} + \sum_{i=1}^{I} \Delta\alpha_{ti} P_{(t-1)i} = \sum_{i=1}^{I} \alpha_{ti} \Delta P_t + \sum_{i=1}^{I} \Delta\alpha_{ti} P_{t-1}.$$

Although we assume that each trader's risk preference coefficient changes over time, it cannot change all of a sudden, which means $\Delta\alpha_{ti} = 0$. Then, we have

$$\Delta P_t = \sum_{i=1}^{I} \alpha_{ti} \Delta P_{ti} / \sum_{i=1}^{I} \alpha_{ti}.$$

Appendix C: Proof of Eqs. (12) and (13)

We first consider the relationship between market price change and traders' reservation price change. From Eq. (5), we have

$$\lim_{\mathrm{Var}(\Delta P_t)\to 0} \mathrm{Var}(\Delta P_{ti}) = 0, \forall i = 1, 2, \dots, I.$$

From the above equation, we know that both market price change and each individual's reservation price change can be derived from the same distribution. Supposing that the market price change is independently normally distributed, we have

$$\begin{aligned} &\Delta P_t = \mu + \sigma Z_t, Z_t \sim N(0,1), \quad \text{that is,} \quad \Delta P_t \sim N(\mu, \sigma), \\ &\Delta P_{ti} = m_i \mu + n_i \sigma Z_t, \quad \text{that is,} \quad \Delta P_{ti} \sim N(m_i \mu, n_i \sigma). \end{aligned}$$

Then, we obtain

$$E(Z) = 0,\ E(Z^3) = 0,\ D(Z) = 1,\ E(Z^2) = 1,\ \ D(Z^2) = 2.$$

Next, we calculate the covariance of h_t and c_t as follows.

If $\sum_{m=1}^{M} \varphi_{tm} \Delta P_{t-m} \geqslant 0$, then $h_t = \sum_{m=1}^{M} \varphi_{tm} \Delta P_{t-m}$ and $\mu \geqslant 0$.

$$\begin{aligned} &\mathrm{cov}(h_t, c_t) \\ =& \mathrm{cov}\left(\sum_{m=1}^{M} \varphi_{tm} \Delta P_{t-m}, \sum_{m=1}^{M} \phi_{tm} \Delta P_{t-m} \Delta P_{(t-m)i}\right) \\ =& \sum_{m_1=m_2=1}^{M} \mathrm{cov}(\varphi_{tm_2} \Delta P_{t-m_2}, \phi_{tm_1} \Delta P_{t-m_1} \Delta P_{(t-m_1)i}) \\ &+ \sum_{m_1\neq m_2=1}^{M} \mathrm{cov}(\varphi_{tm_2} \Delta P_{t-m_2}, \phi_{tm_1} \Delta P_{t-m_1} \Delta P_{(t-m_1)i}) \\ =& \sum_{m_1=m_2=1}^{M} \mathrm{cov}(\varphi_{tm_2} \Delta P_{t-m_2}, \phi_{tm_1} \Delta P_{t-m_1} \Delta P_{(t-m_1)i}) \\ =& \sum_{m=1}^{M} \phi_{tm} \varphi_{tm} \mathrm{cov}\{(\mu + \sigma Z_{t-m}) \\ &\times [m_i \mu + n_i \sigma Z_{t-m}], \mu + \sigma Z_{t-m}\} \\ =& \sum_{m=1}^{M} \phi_{tm} \varphi_{tm} [(m_i + n_i) \mu \sigma^2 \mathrm{cov}(Z_{t-m}, Z_{t-m}) \\ &+ n_i \sigma^3 \mathrm{cov}(Z_{t-m}^2, Z_{t-m})] \end{aligned}$$

$$= (m_i+n_i)\mu\sigma^2 \sum_{m=1}^{M} \phi_{tm}\phi_{tm}.$$

The third equation holds true because ΔP_{t-m_2} is independent of ΔP_{t-m_1} or $\Delta P_{(t-m_1)i}$ when $m_1 \neq m_2$.

If $\sum_{m=1}^{M} \varphi_{tm}\Delta P_{t-m} < 0$, then $h_t = -\sum_{m=1}^{M} \varphi_{tm}\Delta P_{t-m}$ and $\mu \leqslant 0$. So, we can obtain the covariance of h_t and c_t similar to previous values. Combining the above two situations, we have

$$\mathrm{cov}(h_t, c_t) = (m_i+n_i)\,|\mu|\,\sigma^2 \sum_{m=1}^{M} \phi_{tm}\varphi_{tm}.$$

We have already calculated the covariance of h_t and c_t under different conditions above; now we deduce the variance of h_t and c_t, respectively.

$$\begin{aligned} D(h_t) &= \mathrm{cov}\left(\sum_{m=1}^{M} \varphi_{tm}\Delta P_{t-m}, \sum_{m=1}^{M} \varphi_{tm}\Delta P_{t-m}\right) \\ &= \sum_{m_1=m_2=1}^{M} \mathrm{cov}(\varphi_{tm_1}\Delta P_{t-m_1}, \varphi_{tm_2}\Delta P_{t-m_2}) \\ &\quad + \sum_{m_1\neq m_2=1}^{M} \mathrm{cov}(\varphi_{tm_1}\Delta P_{t-m_1}, \varphi_{tm_2}\Delta P_{t-m_2}) \\ &= \sum_{m=1}^{M} \mathrm{cov}(\varphi_{tm}\Delta P_{t-m}, \varphi_{tm}\Delta P_{t-m}) \\ &= \sigma^2 \sum_{m=1}^{M} \varphi_{tm}^2. \end{aligned}$$

The third equation holds true as well because ΔP_{t-m_1} and ΔP_{t-m_2} are mutually independent when $m_1 \neq m_2$.

$$\begin{aligned} & D(c_t) \\ &= \mathrm{cov}\left(\sum_{m=1}^{M} \phi_{tm}\Delta P_{t-m}\Delta P_{(t-m)i}, \sum_{m=1}^{M} \phi_{tm}\Delta P_{t-m}\Delta P_{(t-m)i}\right) \\ &= \sum_{m_1=m_2=1}^{M} \mathrm{cov}(\phi_{tm_1}\Delta P_{t-m_1}\Delta P_{(t-m_1)i}, \phi_{tm_2}\Delta P_{t-m_2}\Delta P_{(t-m_2)i}) \\ &\quad + \sum_{m_1\neq m_2=1}^{M} \mathrm{cov}(\phi_{tm_1}\Delta P_{t-m_1}\Delta P_{(t-m_1)i}, \phi_{tm_2}\Delta P_{t-m_2}\Delta P_{(t-m_2)i}) \\ &= \sum_{m=1}^{M} D(\phi_{tm}\Delta P_{t-m}\Delta P_{(t-m)i}) \\ &= \sum_{m=1}^{M} \phi_{tm}^2\{[D[(m_i+n_i)\mu\sigma Z_{t-m}] + D(n_i\sigma^2 Z_{t-m}^2) \\ &\quad + 2\mathrm{cov}[(m_i+n_i)\mu\sigma Z_{t-m}, n_i\sigma^2 Z_{t-m}^2]\} \\ &= [(m_i+n_i)^2\mu^2\sigma^2 + 2n_i^2\sigma^4] \sum_{m=1}^{M} \phi_{tm}^2. \end{aligned}$$

The third equation holds true because ΔP_{t-m_1}, $\Delta P_{(t-m_1)i}$, ΔP_{t-m_2}, and $\Delta P_{(t-m_2)i}$ are mutually independent when $m_1 \neq m_2$.

we can calculate the correlation coefficient of the i-th trader's confidence degree and herding degree at time t

$$\begin{aligned} & \rho(h_t, c_t) \\ &= \mathrm{cov}(h_t, c_t)/\sqrt{D(h_t)D(c_t)} \\ &= \frac{\sum_{m=1}^{M} \varphi_{tm}\phi_{tm}}{\sqrt{\{1+2[n_i/(m_i+n_i)]^2(CV)^2\}\left(\sum_{m=1}^{M} \phi_{tm}^2 \sum_{m=1}^{M} \varphi_{tm}^2\right)}}, \end{aligned}$$

where $CV = \sigma/|\mu|$ is the coefficient of variation of the price change. When $CV \to 0$, the distributions of x and y tend to be the same, so we have $\varphi_{tm} = \phi_{tm}$.

Then,

$$\lim_{CV\to 0} \rho = 1.$$

When the coefficient of variation of market prices approaches to zero, a trader's herding degree is highly related to his confidence degree.

$$\lim_{CV\to\infty} \rho = 0.$$

When the coefficient of variation of market prices approaches to infinity, a trader's herding degree is almost unrelated to his confidence degree.

References

[1] Donner R V, Small M, Donges J F, Marwan N, Zou Y, Xiang R X and Kurths J 2011 *Int. J. Bifurcat. Chaos* **21** 1019

[2] Zhang J, Luo X and Small M 2006 *Phys. Rev. E* **73** 016216

[3] Zhang J, Sun J F, Luo X D, Zhang K, Nakamura T and Small M 2008 *Physica D* **237** 2856

[4] Marwan N, Donges J F, Zou Y, Donner R V and Kurths J 2009 *Phys. Lett. A* **373** 4246

[5] Gao Z K and Jin N D 2009 *Chaos* **19** 033137

[6] Eroglu D, Marwan N, Prasad S and Kurths J 2014 *Nonlinear Proc. Geoph.* **21** 1085

[7] Xu X K, Zhang J and Small M 2008 *Proc. Natl. Acad. Sci. USA* **105** 19601

[8] Lacasa L, Luque B, Ballesteros F, Luque J and Nuño J C 2008 *Proc. Natl. Acad. Sci. USA* **105** 4972

[9] Lacasa L, Luque B, Luque J and Nuno J C 2009 *Europhys. Lett.* **86** 30001

[10] Luque B, Lacasa L, Ballesteros F J and Luque J 2009 *Phys. Rev. E* **80** 046103

[11] Luque B, Lacasa L, Ballesteros F J and Robledo A 2012 *Chaos* **22** 013109

[12] Yu L 2013 *Physica A* **392** 3374

[13] Zhuang E Y, Small M and Feng G 2014 *Physica A* **410** 483

[14] Shirazi A H, Jafari G R, Davoudi J, Peinke J, Tabar M R R and Sahimi M 2009 *J. Stat. MechTheory E* **7** P07046

[15] Campanharo A S L O, Sirer M I, Malmgren R D, Ramos F M and Amaral L A N 2011 *Plos One* **6** e23378

[16] Zhao Y, Weng T F and Ye S K 2014 *Phys. Rev. E* **90** 012804

[17] Yang Y and Yang H 2008 *Physica A* **387** 1381

[18] Yue Y, Wang J, Yang H and Mang J S 2009 *Physica A* **388** 4431

[19] Pang M B and Huang Y M 2018 *Chin. Phys. B* **27** 118902

[20] Chen D, Shi D D and Pan G J 2019 *Acta Phys. Sin.* **68** 118901 (in Chinese)

[21] Wu L R, Li J J and Qi J Y 2019 *Acta Phys. Sin.* **68** 078901 (in Chinese)

[22] Wang N, Li D and Wang Q 2012 *Physica A* **391** 6543

[23] Sias R W 2004 *Rev. Financ. Stud.* **17** 165

[24] Barber B, Odean T and Zhu N 2009 *Rev. Financ. Stud.* **22** 151

[25] Gontis V, Havlin S, Kononovicius A, Podobnik B and Stanley H E 2016 *Physica A* **462** 1091

[26] Kremer S and Nautz D 2013 *J. Bank. Financ.* **37** 1676

[27] Galariotis E C, Rong W and Spyrou S I 2015 *J. Bank. Financ.* **50** 589

[28] Chiang T C and Zheng D 2010 *J. Bank. Financ.* **34** 1911

[29] Bikhchandani S and Sharma S 2000 *IMF Staff Papers* **47** 279

[30] Gervais S and Odean T 2001 *Rev. Financ. Stud.* **14** 1

[31] Park A and Sabourian H 2011 *Econometrica* **79** 973

[32] Blasco N, Corredor P and Ferreruela S 2012 *Quantum Finance* **12** 311

[33] Hoelzl E and Rustichini A 2005 *Econ. J.* **115** 305

[34] Benoît J P and Dubra J 2011 *Econometrica* **79** 1591

[35] Barber B and Odean T 2001 *Quart. J. Econ.* **116** 261

[36] Chuang W I and Susmel R 2011 *J. Bank. Financ.* **35** 1626

[37] Odean T 1998 *J. Financ.* **53** 1887

[38] Daniel K, Hirshleifer D and Subrahmanyam A 1998 *J. Financ.* **53** 1839

[39] Hirshleifer D, Subrahmanyam A and Titman S 1994 *J. Financ.* **49** 1665

[40] Jegadeesh N and Kim W 2010 *Rev. Financ. Stud.* **23** 901

[41] Tauchen G E and Pitts M 1983 *Econometrica* **51** 485

[42] Lux T 1995 *Econ. J.* **105** 881

[43] Yook S H, Jeong H, Barabási A L and Tu Y 2001 *Phys. Rev. Lett.* **86** 5835

[44] Zheng B, Ren F, Trimper S and Zheng D F 2004 *Physica A* **343** 653

[45] Avery C and Zemsky P 1998 *Amer. Econ. Rev.* **88** 724

[46] Statman M, Thorley S and Vorkink K 2006 *Rev. Financ. Stud.* **19** 1531

[47] Karpoff J M 1987 *J. Financ. Quant. Anal.* **22** 109